NEW DIRECTIONS IN APPLIED GENERAL EQUILIBRIUM MODEL CALIBRATION

Christina Dawkins

Thesis submitted in

fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Warwick

Department of Economics

August 1999

This thesis is dedicated to the memory of my father, William,

and to my mother, Eva.

NEW DIRECTIONS IN APPLIED GENERAL EQUILIBRIUM MODEL CALIBRATION

CONTENTS

List of Figures and Tables		**i**
Acknowledgements		**iii**
Declaration		**iii**
Summary		**iv**
Chapter 1	**Introduction**	**1**
1.1	Organization of the Thesis	2
Chapter 2	**What is Calibration?**	**9**
2.1	Introduction	9
2.2	The Origins of Calibration: Two Traditions	12
2.2.1	The Shoven and Whalley Simulations	13
2.2.2	The Kydland and Prescott Model	16
2.2.3	Common Motivations for Calibration	18
2.2.4	A Comparison of the Two Traditions	19
2.3	Calibration in Applied General Equilibrium Models	24
2.3.1	An Example of Calibration	27
2.4	Calibration in Context: The Modelling Process	33
2.4.1	Model Choice	34
2.4.2	Dimensionality and Functional Forms	36
2.4.3	Collection of Data	40
2.4.4	Deriving a Benchmark Equilibrium Data Set	41
2.4.5	Elasticity Specification	45
2.4.6	Calibration, the Base Case Model and Replication	47
2.4.7	Counterfactual Simulation and Interpretation of Results	47
2.4.8	Sensitivity Analysis	49
2.5	The Econometric Critique of Calibration	51
2.6.1	Responses to the Critique	54
2.6	Conclusion	56
Chapter 3	**Extended Sensitivity Analysis**	**58**
3.1	Introduction	58
3.1.1	Technical Issues	60
3.2	Calibrated Parameter Sensitivity Analysis (CPSA)	63
3.2.1	CPSA Methodology	65
3.2.2	A Formal Presentation of CPSA	72
3.2.3	An Illustration of CPSA Using a Simple Tax Model	76
3.2.4	The Mechanism of CPSA	85
3.3	Extended Sensitivity Analysis	89
3.3.1	Extended Sensitivity Analysis Methodology	90
3.3.2	Extended Sensitivity Analysis for the Côte d'Ivoire Model	92
3.4	Conclusion	101

Chapter 4 **The Significance of the Data Adjustment Algorithm Choice** **104**
4.1 Introduction 104
4.1.1 Algorithm Evaluation Experiments 105
4.1.2 Why Algorithms Perform Differently 106
4.2 The Criteria for Adjustment Algorithm Evaluation 109
4.2.1 Evaluation Criteria in Previous Work 112
4.3 Adjustment Algorithms 114
4.3.1 *Ad hoc* Adjustment Procedures 116
4.3.2 Adjustment Algorithms for Evaluation Experiments 117
4.4 Adjustment Algorithm Evaluation Experiments 122
4.4.1 Model Description 124
4.4.2 Monte Carlo Experiment Methodology 127
4.4.3 Results in the Absence of Data Reliability Information 133
4.4.4 Results with Data Reliability Information 139
4.5 The Influence of Relative Data Magnitudes on Algorithm Choice 141
4.5.1 The Effects of Alternative Data Structures 145
4.6 Conclusion 150

Chapter 5 **Decomposition Analysis Using Applied General Equilibrium Models** **152**
5.1 Introduction 152
5.1.1 Interactions Between Shocks 154
5.2 The Modelling Framework for Decomposition Analysis 156
5.2.1 Forward Simulation 157
5.2.2 Backward Simulation 159
5.2.3 Decomposition Analysis 160
5.2.4 Calibration for Decomposition Analysis 162
5.3 Implementing Decomposition Analysis 166
5.3.1 Model Structure 167
5.3.2 Calibration 169
5.4 Decomposition Analysis 185
5.4.1 Interactions Between Shocks 190
5.4.2 The Interactions Among Shocks in the Illustrative Model 194
5.5 Conclusion 199

Chapter 6 **Conclusions** **201**
6.1 Directions for Future Research 203

References **205**

List of Figures and Tables

Chapter 2

Table 2.1	An Example of a Microconsistent Data Set	29
Table 2.2	Calibrated Parameter Values for the Example Model Using the Data in Table 2.1	32
Figure 2.1	The Context for Applied General Equilibrium Model Calibration	35

Chapter 3

Table 3.1	Structure of the Tax Model Used to Illustrate CPSA	77
Table 3.2	Unbalanced Transactions Values For the Illustrative Tax Model	78
Table 3.3	Balanced Data and Central Case Model Results For the Illustrative Tax Model	81
Table 3.4	3 Point Discrete Gaussian Quadrature Approximations for the Assumed Continuous Data Distributions in the Illustrative Model	82
Table 3.5	CPSA on the Welfare Effects of Imposing a 50 Percent Tax on the Use of Capital in the Manufacturing Sector	84
Table 3.6	Sensitivity of CPSA to the Data's Probability Density Functions	86
Table 3.7	The Link Between Changes in the Raw Data Estimate of the Poor Consumer's Consumption of Non-Manufactured Goods (v_{pn}) and Changes in Model Results	88
Table 3.8	Unbalanced Household Survey Consumption Expenditure Data for the Côte d'Ivoire Model	94
Table 3.9	Aggregate Totals Used as Constraints in the RAS Adjustment Algorithm	96
Table 3.10	Extended Sensitivity Analysis Results for Personal Income Tax Incidence in a Model of Côte d'Ivoire	98
Table 3.11	Elasticities of Substitution and Bounds Used in the Extended Sensitivity Analysis	100

Chapter 4

Table 4.1	Evaluated Adjustment Algorithms	118
Table 4.2	Structure of the Model for the Monte Carlo Experiment	125
Table 4.3	True Benchmark Equilibrium Data Set Used in the Monte Carlo Experiment	126
Table 4.4	Standard Deviations Specified for the Data Generating Process	128
Table 4.5	Statistical Properties of the Algorithm-Based Sample of Hicksian Equivalent Variations (EV) in the Monte Carlo Experiment	134
Table 4.6	The Location of Bias in the Adjusted Matrix	137
Table 4.7	Statistical Properties of the Sample of Hicksian Equivalent Variations Adjusted Using Full-Information Algorithms	140
Table 4.8	Non-Zero True BED Values for the Grid of Data Structures	144
Table 4.9	Bias and *z*-Statistic Ranking of Adjustment Algorithms Under Different Data Structures	146
Table 4.10	Mean Square Error x 10^7 (level and rank) of Adjustment Algorithms Under Different Data Structures	147

Chapter 5

Table 5.1	Summary of the Model Structure Used to Illustrate Decomposition Analysis	168
Table 5.2	Raw Data Used in Constructing the Benchmark Data Sets	171
Table 5.3	Construction of the Benchmark Data Sets for Decomposition Analysis	173
Table 5.4	Derivation of Quantities and Prices Used to Calibrate Technology Parameters	176
Table 5.5	Quantities and Prices for Calibration	177
Table 5.6	Production Parameter Values	179
Table 5.7	Parameter Value Changes in the Model	181
Table 5.8	GNP in Millions of 1914 USD Arising from Changes in Parameter Configurations	187
Table 5.9	The Effects of Individual Shocks on US GNP between 1870 and 1890	188
Table 5.10	Derivation of Pairwise Interactions for the Four Model Shocks	195
Table 5.11	Derivation of Three-Way Interaction Values	197
Table 5.12	Derivation of the Four-Way Interaction Value	198

Acknowledgements

I gratefully acknowledge the financial support that has allowed me to complete this thesis. Funding was provided through a graduate research assistantship in the economics department at the University of Warwick, and from the UK Economic and Social Research Council under an award for the General Equilibrium Analysis of UK Policy Issues.

This thesis has benefitted enormously from the comments of my supervisors, Carlo Perroni and John Whalley. I am grateful to both: Carlo for his patience in understanding and proposing ways around the myriad obstacles encountered in day to day thesis research, and for being exceptionally generous with his time; and John for his invaluable suggestions of research topics, and for very useful advice on presentation and drafting.

I also thank Eva Dawkins, Julia Prescott, and all of my friends for their emotional support, without which I could never have finished this thesis. I am especially grateful to Terry Coggin who is a constant reminder to me that good things can result from taking the wrong turn at the end of a long and rainy day.

Declaration

The material in Chapter 2 of this thesis, 'What is Calibration?' is based on joint work with T. N. Srinivasan and John Whalley forthcoming (1999) as C. Dawkins, T. N. Srinivasan, and J. Whalley, "Calibration," in E. Leamer and J. Heckman (eds.), *Handbook of Econometrics* Volume 5, published by North-Holland. Unlike the original paper, which serves as an overview of calibration in economics, the discussion in Chapter 2 is intended to offer the reader a context for the issues addressed by the research in subsequent chapters. Consequently, it emphasizes the calibration of applied general equilibrium models more so than the original paper. Any errors introduced in the adaptation are, of course, my responsibility alone.

Summary

This thesis develops extensions to current techniques in applied general equilibrium (AGE) model calibration that improve existing practice and expand the use of AGE modelling to economic history applications. Chapter 1 introduces the thesis. Chapter 2 summarizes the origin and practice of calibration in economics, focussing on its role in AGE modelling.

Chapter 3 proposes two related sensitivity analysis procedures for AGE models: calibrated parameter sensitivity analysis (CPSA) and extended sensitivity analysis. Existing sensitivity techniques are incomplete because they only capture the robustness of the model's results to uncertainty in a subset of the parameters, the elasticities. The remaining parameters determine the model's static structure, but are ignored in the sensitivity literature. CPSA fills this gap. When combined with an existing elasticity sensitivity technique in 'extended sensitivity analysis,' CPSA permits sensitivity analysis with respect to uncertainty in the values of all of a model's parameters.

Chapter 4 examines the significance of the data adjustments required for calibration. It proposes that the measure of this importance should be the effect of the adjustment algorithm on the statistical properties of the model results. Simulations show that the performance of various algorithms differs significantly under such criteria, and illustrate for a specific policy experiment the link between algorithm performance and the relative magnitudes of the data. The experiments imply that the choice of data adjustment procedure is an important, if neglected, component of calibration.

Chapter 5 shows how AGE techniques can be adapted to explore decompositional issues in economic history. By incorporating information about the combined effect of several shocks to an economy in calibration, AGE models can quantify the relative contributions to change of each shock. Furthermore, the effects of shocks are non-additive, so that the marginal contribution of a shock is conditional on the presence or absence of other shocks. Chapter 6 concludes.

Chapter 1

Introduction

Since Shoven and Whalley pioneered calibration as a way of deriving applied general equilibrium model parameters in the early 1970s, the methodology has remained largely unchanged. Advances in computing have allowed more complex model structures and faster solution algorithms, but the underlying calibration procedure developed in Shoven and Whalley (1972), and discussed in Mansur and Whalley (1984) and Shoven and Whalley (1992) continues to be standard applied general equilibrium modelling practice.

This thesis proposes several extensions to the standard Shoven-Whalley methodology that improve the existing calibration practice and expand the application of applied general equilibrium modelling to include decomposition issues in economic history. Chapter 3 develops two related sensitivity analysis procedures that allow modellers to report the robustness of their model results to uncertainty in the values of all of a model's parameters, rather than to uncertainty in the values of only a subset of those parameters, as is possible under existing sensitivity procedures; Chapter 4 proposes criteria for evaluating the consistency adjustments that are made to the data during calibration and shows that for a small tax model the performance of several well known adjustment algorithms differs significantly under these criteria; and Chapter 5 modifies the standard Shoven-Whalley calibration methodology so that applied general equilibrium models can be used to decompose a known historical change into its

component causes and can also be used to explore the interactive effects of several shocks to an economy over a specified time interval.

1.1 Organization of the Thesis

The thesis begins in Chapter 2 with an overview of calibration that is intended to provide background information and a context for the methodological innovations presented in the three subsequent research chapters. Chapter 2 discusses the origins of the two calibration traditions in economics; the Kydland and Prescott calibration of aggregate macroeconomic models, and the Shoven and Whalley calibration of applied general equilibrium models. Because the calibration of applied general equilibrium models is the focus of this thesis, Chapter 2 describes the basic Shoven-Whalley methodology in detail, and then places it within the context of the wider applied general equilibrium modelling process.

Although for practical purposes calibration has been a largely unchallenged approach to applied general equilibrium model parameterization, the weaknesses in current calibration practice have long been recognized. Chapter 2 also presents the econometric critique of calibration which highlights some of these weaknesses. Elements of this critique are the motivation for the research in Chapters 3 and 4. Chapter 3 introduces sensitivity analyses which respond to the criticism that the calibrated parameters are numerically fragile, and Chapter 4 addresses the criticism that the data adjustments made during calibration introduce untraceable bias into the modelling process, by exploring the effects of such adjustments on the model results.

Chapter 3: Extended Sensitivity Analysis

Chapter 3 develops and illustrates two related sensitivity analysis procedures for applied general equilibrium models: calibrated parameter sensitivity analysis (CPSA) and extended sensitivity analysis, which combines CPSA with existing sensitivity procedures for elasticity parameters. CPSA is a procedure which generates confidence intervals for the model results based on the uncertainty in the data that are used to find values for the model's calibrated parameters. Extended sensitivity analysis produces confidence intervals for the model's results based both on the uncertainty in the values of the data used to find the calibrated parameter values and on the uncertainty in the values of the model's elasticity parameters.

Although several sensitivity analysis techniques for applied general equilibrium models have been developed in the literature, they are incomplete because they can only capture the robustness of the model results to uncertainty in the values of a subset of the model's parameters, the elasticities. The remaining parameters - the calibrated parameters - include consumers' expenditure shares and the input shares and scale parameters in the production functions. They form the static specification of the modelled economy and typically comprise the majority of the model's parameters, so that their omission from existing sensitivity analyses represents a serious gap in the modelling literature. CPSA allows modellers to undertake sensitivity analysis with respect to uncertainty in the calibrated parameters, and by combining CPSA with an existing elasticity sensitivity analysis, extended sensitivity analysis makes possible sensitivity analysis over the uncertainty in a model's complete numerical specification.

CPSA consists of perturbing the central case values of the unbalanced data from which a model's calibrated parameters are jointly determined. It uses *a priori* information about the reliability of the data to specify the perturbation in such a way that its probability of being true is known. The set of calibrated parameters associated with the perturbation is then found and used both to solve the model and to generate model results. Each model result is weighted by the probability of the perturbation used in its derivation.

The process is repeated for a sample of perturbations. The ensuing sample of model results and their associated probabilities are used to construct the expected values, standard deviations, and confidence intervals for the model results. These statistics express the robustness of the model results to uncertainty in the initial data, and hence, to the uncertainty in the values of the calibrated parameters. The methodology for extended sensitivity analysis is the same as CPSA except the perturbations include changes to the elasticity values as well as to the unbalanced data.

Chapter 4: The Importance of Adjustment Algorithm Choice

Chapter 4 focusses on the data adjustment step of the calibration process. Many algorithms exist for balancing data so that they can be used in calibration, but the applied general equilibrium modelling literature offers no guidance for choosing one adjustment algorithm over another. Chapter 4 argues that because the data used to calibrate a model are random variables, the model results are also random variables. It proposes, therefore, that the criteria for the choice of adjustment algorithm be the effect

that an adjustment procedure has on the bias of the mean values for the model results and its effect on their mean square errors. These criteria are then used to evaluate several common adjustment algorithms which are applied to data for a small tax model. The evaluation is undertaken experimentally using Monte Carlo simulations in which the true model results are known and in which a data generating process for the modelled economy is also specified.

Under the proposed criteria some of the adjustment algorithms tested in Chapter 4 perform much better than others. The performance ranking of the algorithms in the evaluation exercise, however, is specific to the model and to the tax policy experiment, so that it yields no general conclusions about which adjustment algorithm will perform best for every model. Instead, the fact that the adjustment algorithms perform so differently from one another in the small tax model is evidence that the choice of adjustment algorithm matters for the model results of at least one model, and that this choice is, therefore, also likely to be important for the results of other models. If the possibility exists that the choice of data adjustment algorithm can introduce bias to the model results, modellers have a clear incentive to evaluate alternative algorithms before choosing one. The development of a systematic algorithm evaluation technique emerges from the experiments in Chapter 4 as an important direction for future research in applied general equilibrium model calibration.

A further set of experiments in Chapter 4 explores how changing the relative magnitudes of the elements of the data set affects the performance ranking of the tested adjustment algorithms. The proposed explanation for this link is that some elements of the data are more important in determining the results of a specific policy experiment

than others. If the more important elements of the data are small relative to the less important elements, the adjustment algorithm which performs best will be the one that places a lower burden of adjustment on the smaller data elements. The results of the experiments in Chapter 4 suggest that this line of reasoning has merit.

Chapter 5: Decomposition Analysis Using Applied General Equilibrium Models

Chapter 5 shows how applied general equilibrium models can be adapted to analyze economic history problems in situations where a historian can identify the main shocks to an economy over a specified interval, and is interested in what the relative contribution of each individual shock was to some overall measure of economic change. It describes and illustrates a procedure that uses applied general equilibrium models to decompose a known historical change into its component causes.

Decomposition analysis requires the modeller to specify a pre-change and a post-change equilibrium, and to identify the main shocks to the economy in the interim. The model is calibrated so that when no shocks are introduced to the model, the model solution matches the pre-change equilibrium exactly, and when all of the shocks are introduced to the model, the model solution matches the post-change equilibrium exactly. The modeller then sequentially introduces each subset of the specified shocks to the model and solves the model for each. If, for example, an economy is modelled as facing three shocks, the model would be solved for the three cases where each shock is introduced individually, and for the three cases in which pairwise combinations of those shocks are introduced to the model.

The contribution of a specific shock to the change in a variable of interest is measured by comparing the value of that variable from the model solution that includes the shock to its value from the model solution without the shock. Consider an economy facing two shocks, A and B, in which the modeller seeks to know the contribution of shock A to a known change in GNP. In a two shock example, two such measures exist. The first gives the contribution to growth of shock A in the absence of shock B. It is found by comparing GNP when only shock A is introduced to the model with GNP in the pre-change equilibrium. The second gives the contribution of shock A to the change in GNP in the presence of shock B. It is found by comparing GNP in the post-change equilibrium with GNP when only shock B is introduced to the model.

These two measures are unlikely to be equal, because shocks to an economy are unlikely to exhibit strictly additive effects. They may be mutually enhancing so that, for example, a specific technological change may increase output more in the presence of population growth, than in its absence. Conversely, they may have offsetting effects; a technological change may increase output less in the presence of a change in consumer preferences than in its absence. The quantitative assessment of the marginal impact of a specific shock will depend on the status of the remaining shocks, and modellers should, therefore, take care to report their results conditional on the presence or absence in the experiments of other shocks to the economy.

A quantitative measure of this interaction among shocks can also be found from the decomposition analysis model simulations. It is obtained by finding the difference between the sum of the shocks' individual effects and their joint effects. In the two shock example, the joint effect of shock A and B is the difference between GNP in the

pre-change equilibrium and in the post-change equilibrium. The measure of the synergy of shocks A and B is given by adding the change in GNP attributable to shock A to the change in GNP attributable to shock B and subtracting from this sum the net change in GNP between the pre- and post-change equilibria.

Chapter 5 illustrates historical decomposition analysis using a simple applied general equilibrium model of railroads in the US in the period 1870 to 1890 in which the relative contributions to GNP growth of changes in railroad technology, other technology, factor endowments, and preferences are analyzed. While 91 percent of the change in GNP over this interval can be explained by adding together the individual effects of the four shocks, the pairwise, three-way and four-way synergistic interactions among the shocks account for the remaining 9 percent of the change, suggesting that for this application at least, the interactive effects of shocks are not negligible.

Chapter 6: Conclusions

Chapter 6 concludes. Two broad themes emerge from this thesis. The first is that the unbalanced data and the balancing adjustments made by modellers to those data are not peripheral elements of the modelling process. Instead, they are the areas in which improvements can be made to current calibration practice.

The second theme is that including a second equilibrium data observation in calibration allows historians to quantify the causes of change. By exploiting information about the combined effect of several shocks, historians can undertake a richer analysis than is possible from mechanically applying existing modelling techniques.

Chapter 2

What is Calibration?[1]

2.1 Introduction

This chapter sets the scene for the discussion in the remainder of the thesis by describing the calibration of applied general equilibrium models and placing it within the context of the wider modelling process. Calibration in economics is not, however, restricted to applied general equilibrium models. Instead, the term 'calibrated model' embraces the broad class of numerical models for which parameters are derived using methods other than econometric estimation. Two distinct methodologies share the rubric 'calibration' so that the term has been a source of much miscommunication among practitioners. One approach, originating with Kydland and Prescott (1982) parameterizes dynamic stochastic general equilibrium models. The other, which is the focus of the discussion in this thesis, was developed by Shoven and Whalley (1972) to parameterize static applied general equilibrium models. This chapter also contrasts the origins and approaches of the two methodologies.

The belief which underpins the use of calibration in economics is that numerical analysis requiring complex, econometrically untestable models is essential for the understanding of specific issues. For such issues, the task is to parameterize rather than to test a model. If the model's structure is sufficiently complex that its parameters

[1] This chapter is based on the joint work with T. N. Srinivasan and John Whalley forthcoming (1999) as C. Dawkins, T. N. Srinivasan, and J. Whalley, "Calibration," in E. Leamer and J. Heckman (eds.), *Handbook of Econometrics* Volume 5, published by North-Holland. Although most of the content of this chapter can be found in the joint paper, the emphasis here is on calibration in applied general equilibrium modelling. Any errors introduced in the adaptation of the original paper are, of course, my responsibility alone.

cannot be estimated, its parameter values must still be obtained somehow. Calibration refers to procedures that complete a model's parameterization.

The term 'calibration' in economics has been borrowed from common parlance. Its more general use includes the setting of the origin and choice of scale for a measuring instrument. For example, a thermometer is calibrated to read 0°C and 100°C when water freezes and boils. Although the meaning of calibration seems intuitive, identifying some of the characteristics of instrument calibration may be useful for understanding how the term is applied in economics.

Instrument calibration is a numerical exercise in which measurement is the central objective. Although calibration is undertaken in a controlled setting where the response to a particular stimulus is known, the goal is to produce an instrument which can be used for measurement where the response is unknown. The instrument is correctly calibrated when it reproduces the known result within an acceptable tolerance. Calibration, therefore, implies some form of result replication testing. Finally, to calibrate an instrument like a thermometer, some parameters such as the expansion coefficient for mercury and the diameter of a thermometer are taken as constant, while others are variable, such as the placement of the markings on the thermometer or the volume of mercury. The act of calibration is undertaken conditional on the values of the constant parameters, so that if, for example, another thermometer were to use coloured water in the place of mercury or had a larger bore, the temperature scale determined by the original calibration would no longer hold.

Applied general equilibrium modellers employ a process which is analogous to the calibration of the thermometer. They typically calibrate their models to a known,

single, constructed equilibrium observation, so that they generate a model specification which is capable of reproducing the input data as a model solution. Calibration, in this case, uses a single observation on the data set as a consistency check, and hence, the element of replication is present. Once calibrated, the model is used to ask how the economy might behave in an unknown, 'counterfactual' situation where, for example, tax rates are higher or a quota has been removed. Finally, in parameterizing a general equilibrium model, the model parameters are differentiated so that some, such as the elasticities of substitution in constant elasticity of substitution (CES) functional forms, are exogenously specified and are akin to the constants in case of the thermometer, while others are set through calibration. Hence, the calibrated parameter values are conditional on the specified elasticity values.

In contrast to applied general equilibrium microeconomic models, highly aggregated dynamic macroeconomic models specify a structure with a steady-state or long run joint distribution of the aggregate variables that can be described parametrically. Typically, such models include a specification of stochastic elements which influence the model behaviour. Calibration in this context consists of asking whether, for plausible values of its parameters, the steady-state distributions generated by the model correspond to those of the data. Unlike applied general equilibrium models, the term calibration in this case sometimes embraces the exogenous specification of parameter values which have no explicit consistency check with data, and which are determined independently of the model structure. Also, in contrast to applied general equilibrium models, a match between the calibrated model output and data is often interpreted as a validation of the model structure.

Calibration is a common procedure in the physical sciences and also has a natural role in economics. Policy issues need input from economics and the theoretical models required for such analyses are well developed in the literature. The sole remaining requirement for the numerical implementation of theory is to choose appropriate values for the model parameters. Where econometric parameterization procedures are not feasible, calibration is the only logical alternative.

This chapter gives an overview of calibration which provides the background information and context for the research topics addressed in the remainder of this thesis. It is organized as follows. Section 2.2 describes the origin of calibration for both static applied general equilibrium and dynamic macroeconomic models, and offers a general comparison of the two calibration traditions. Section 2.3 formalizes calibration in applied general equilibrium modelling, while Section 2.4 places the formal calibration procedure within the wider context of applied general equilibrium modelling. Section 2.5 presents the econometric critique of applied general equilibrium model calibration and shows how modellers have addressed elements of this critique. Section 2.6 concludes.

2.2 The Origins of Calibration: Two Traditions

Calibration originated as economists began to address problems which could not be handled using existing numerical techniques, particularly in the area of policy analysis. In the 1950s, for example, trade economists who were faced with the Treaty of Rome and what is today the European Union began to explore the implications of regional

trade agreements. Theoretical trade economists debated the relative importance of the trade diversion and trade creation effects stemming from a Union. These two effects clearly operated in opposite directions, so that theory could not offer guidance to an individual country which was contemplating joining. Numerical calculations with precise equilibria that reflected the structure of the modelled economy were needed to determine which effect dominated.

2.2.1 The Shoven and Whalley Simulations

One of the earliest applied general equilibrium calibration exercises was undertaken by Shoven and Whalley (1972), who attempted to refine Harberger's earlier (1962) calculations of the welfare cost associated with the differential tax treatment of capital income by sector in the US. They used Harberger's earlier model and data, averaged over years in the late 1950s and early 1960s, with 1959 as the mean, but applied Scarf's (1973) algorithm to solve the model for exact equilibria rather than the approximate equilibria that Harberger had obtained by linearizing around an initial pre-tax change equilibrium. They took Harberger's data set and extended it via a few simple adjustments into an explicit equilibrium data set.

At first they generated numerical values for the parameters of their model by adjusting initial values iteratively, and observing how closely the model solutions matched the constructed equilibrium data set. Early working paper versions included diagrams which illustrated the distance between the true solution and the model solutions arising from the use of various parameter configurations. Shoven and Whalley

then found that instead of simply applying parameter combinations, they could use the equations that characterized an equilibrium solution of the model to solve for the values of the parameters such that the equations were satisfied for the benchmark data. In essence, their procedure converted parameters into variables, data into exogenous parameters, and trivially imposed equilibrium as an identifying restriction on the model specification. This procedure became known as calibration and remains the standard practice for parameterizing applied general equilibrium models.

One reason why Shoven and Whalley adopted calibration in their work was that other fledgling numerical general equilibrium models of the time had relied exclusively on literature-based econometric estimates for all parameter values and, more importantly, had used them in their base case model specification. This practice typically generated base case representations of the modelled economy which were highly inconsistent with observation. In an economy contemplating tax reform, such a model specification might give a base case solution in which, for example, 50% of employment is in manufacturing when the data clearly showed the figure to be 25%. The actual performance of the economy in the base case, which was known from national accounts data, was in no way reflected in the base case solution of the model after parameter values were taken from literature. Shoven and Whalley's observation that model outcomes differed significantly from the data led them to reject the exclusive use of literature-based parameter values and to rely instead on parameter values generated by the model structure. Using the equilibrium solution concept of the model as an identifying restriction on parameterization led to their calibration procedure.

The calibration procedure developed by Shoven and Whalley was purely a practical response to the need for a realistic model parameterization. The paramount requirement of the calibrated model initially was, and continues to be, to reproduce the known base data as an equilibrium, either exactly or closely, in what has become known as a replication test.[2] A failure of the replication test catches many coding or other errors. Importantly, the Shoven-Whalley calibration procedure has no predictive power since a variety of models and functional forms could be calibrated to the same data.

The calibration methodology developed by macroeconomic modellers takes a different tack. Macroeconomic modellers check the value of model parameters for their ability to generate stochastic equilibrium time paths for steady-states, which are consistent with the stochastic properties of the joint distribution of the observed data on the same aggregate. Thus, they evaluate their model structure on the basis of how closely the moments of the model solution approximate the corresponding moments of the data. Observed data cannot be used to infer parameter values which exactly replicate base data, nor are equations characterizing the model solution used to solve for model parameters with the role of endogenous and exogenous parameters reversed. Unlike Shoven and Whalley, they make no model-consistent pre-adjustments to the basic data, assuming that the data represent realizations of the equilibrium path of their model. Furthermore, in contrast to microeconomic policy analyses where no stochastic disturbances are admitted, macroeconomic modellers allow stochastic shocks to enter their models.

---

[2] The replication test is undertaken assuming the absence of multiple equilibria.

2.2.2 The Kydland and Prescott Model

The earliest example of calibration in the macroeconomic tradition is Kydland and Prescott's (1982) real business cycle model. They presented a simple one sector growth model with a labour-leisure choice and non-time separable preferences, which they argued could be used to explain the autovariances of real output and the covariances of cyclical output with other aggregate time series for the post-war US economy. The crucial element of their model was the assumption that more than one time period is needed to construct newly productive capital. Kydland and Prescott then introduced various stochastic shocks into the model structure, including technology and productivity shocks.

They argued that a test of their structure was whether a set of parameters existed for which the model's co-movements for both the smoothed series (on output, investment, consumption, labour productivity, and capital stocks) and the deviations from these smoothed series were quantitatively consistent with the observed behaviour of the corresponding series for the post-war US economy. They added the further requirement that the parameters chosen should not be inconsistent with the relevant microeconomic observations, including the reported construction periods for new plants and cross-sectional observations on consumption and labour supply. They also suggested that the closeness of their specification of preferences and technology to those used in related applied work facilitated comparisons to other research.

Kydland and Prescott first specified their model so that its steady state properties were consistent with long term trend data for the US. Quantitatively

explaining the co-movements of the deviations remained as the test of the underlying theory. They emphasized some of these key co-movements; investment varied three times as much as output while for consumption the variation was only one half; variations in output largely reflected variations in hours worked per household, not capital stocks or labour productivity.

The Kydland and Prescott calibration separated parameters into those which were fixed exogenously and those which were free. The first bloc of parameters were largely chosen by appealing to plausible values for key aggregates and literature estimates. Two parameters affecting the intertemporal substitutability of leisure and three variance parameters on productivity shocks were left free, with the sum of the variance parameters restricted so that the model estimate of the variance of cyclical output equalled that of the US economy.

For each set of parameter values, the autocorrelation of cyclical output for up to six periods was computed, along with standard deviations of cyclical variables of interest and their correlations with cyclical output. These values were compared to the same statistics for the US economy. Kydland and Prescott chose what they considered to be the best fit and then examined the actual model solutions. Comparing estimated autocorrelations for real output from the model with sample values for the US economy, Kydland and Prescott concluded that the fit is surprisingly good. On this basis, they suggested that the model loosely met a goodness-of-fit criterion, and could be accepted as a reasonable structure to use to analyze macroeconomic issues in the US.

Although Kydland and Prescott provide a three page discussion of how they calibrate their model by choosing the majority of the parameter values and leaving other

parameters free to be determined by a model fit to data, calibration was not their main focus. The word itself does not appear until the latter half of the paper, and then it only appears once, in the subheading 'Model Calibration.' Nowhere does the word appear in the text of the published paper, and no explanation of the term is offered.

2.2.3 Common Motivations for Calibration

The calibration methodology developed by Kydland and Prescott, however, has been widely adopted and forms the basis of the calibration of current dynamic macroeconomic models, while the calibration of Shoven and Whalley continues to be the standard procedure in applied general equilibrium modelling. The motivation behind the development of both types of calibration was the need to have model solutions which matched data; Kydland and Prescott's desire was to have a model which was quantitatively consistent with observed time series co-variation between output, consumption and investment, while Shoven and Whalley required model solutions that matched a single disaggregated data observation of the US economy. Both sets of modellers were unable to meet this requirement with existing econometric techniques and developed their respective calibration procedures as an alternative.

Applied economists continue to be interested in numerical models with richer economic structures than are currently found in many econometric models. Calibration in economics has flourished because the economics in econometrics, and the economics in pure theory seem to have drifted apart. The tendency in econometrics has been to append increasingly sophisticated stochastic disturbance terms to relatively simple

economic models that lag behind the theoretical frontiers. For example, demand estimation has advanced from single commodity demand functions to systems of demand functions; but combined demand and supply systems are rarely estimated, and neither multi-consumer demand systems nor the two person pure exchange economy have been the topic of any econometric application. Furthermore, econometric models of demand do not incorporate features about commodities, such as product quality, which numerical modellers view as relevant for their analyses.

Thus, the shortcomings of econometric techniques in addressing specific economic problems are responsible for the origin and growth in popularity of both types of calibrated models. Yet, the two calibration traditions remain distinct.

2.2.4 A Comparison of the Two Traditions

Although the Shoven-Whalley and the Kydland-Prescott calibrations both employ techniques which fall outside the traditional econometric domain to parameterize their models, they differ in several important ways. Perhaps the most striking area of difference in the two approaches is their objectives. Applied general equilibrium models are typically constructed to answer questions about the allocative effects of economic policies, so that they require a relatively disaggregated representation of an economic system. They are regarded as simply the application of numbers to theory. Modellers make no claims about testing the theory. To them, the widespread use of a particular structure in the theoretical literature is an indication of its worth, so that they seek less to test or validate models and more to explore the numerical implications of a particular

model, conditional on having chosen it. They tend to be agnostic about particular models, accepting that many alternative structures relevant to an issue exist in the theoretical literature, sometimes producing different results. Thus, the focus of microeconomic modellers is to evaluate policies or changes conditional on a particular theoretical structure, rather than to test the theory itself. Because the complexity of the models and of the policy changes to be evaluated preclude closed form solutions, no realistic alternative to calibrated applied general equilibrium models exists.

The objective of exercises using dynamic, stochastic macroeconomic models, however, is not so much to evaluate a specific policy but to uncover the fundamental interactions in an economy. At the heart of such models is a parsimonious, idealized model derived from theory, from which the modeller extracts as much information as possible. Thus, dynamic macroeconomic modellers seek the simplest model which can explain observed trends. Where simulation outcomes differ from data, the underlying theory is not rejected, but instead the modeller adds features in an attempt to improve the match between the moments of simulation outcomes and data.[3] The relative merits of adding specific features are measured by the subsequent improvements in the model's performance.

Model evaluation is an implicit element of such an exercise: simulations that are close to the data under a specified metric validate the underlying theory. A divergence in model results and data points to the need for further theoretical research. For example, Mehra and Prescott (1985) use a calibrated model to show that under

[3] Hoover (1995) characterizes this distinction as a central difference between macro calibrators and econometricians.

reasonable restrictions, standard competitive theory cannot explain both the low average real returns to debt and the high returns to stocks. Although this so-called equity premium puzzle remains unresolved, it has provided the impetus for much of the research discussed in Kocherlakota (1996).

This difference in the underlying objectives for the two types of models is reflected in differences in their dimensionality. Applied general equilibrium models are inevitably more concerned with disaggregated representations of economies than is true of macroeconomic models. For example, the Whalley (1985) global trade model considers four trade blocs, each of which produces 33 commodities, and has a government and several household types as consumers, whereas macro models typically specify behaviour for a single representative consumer.

This scope for accommodating detail enables applied general equilibrium modellers to address specific policy questions which, by virtue of the scale of their disaggregated data requirements, are untenable within a statistical modelling framework. Consequently, the numbers of variables, data, and parameters to be calibrated are also greater than in macroeconomic models. Obtaining even a single observation for the large number of variables is a time-consuming and costly endeavour. Hence, while calibrated macro models are parsimonious in the number of variables, calibration in the applied general equilibrium case is as parsimonious as possible in its use of data; parameters are derived from a single observation of the transactions in an economic system.

Data constraints in applied general equilibrium modelling have further implications for the Shoven-Whalley calibration methodology. Calibrating from a

single observation leaves insufficient degrees of freedom to admit any stochastic structure or measurement errors. Thus, the single observation must represent an exact equilibrium solution for the model, and data adjustments are undertaken so that it does. Calibration, in this case, is deterministic in the sense that the parameters fit the data exactly.

In contrast, the real business cycle and other calibrated macroeconomic models employ a small number of aggregate variables and their dynamics include relatively few parameters. Time series observations are typically available for such models, so that the number of observations far exceeds the number of parameters to be calibrated. Calibration, in this case, finds parameters which best fit the data. Although some adjustments might be made to the data to separate their trend and cyclical components, these adjustments are made to extract the variables that are of interest from the measured data. The presence of multiple observations, however, precludes the need for the types of model-dependent data adjustments undertaken prior to applied general equilibrium model calibration.

Because the Shoven-Whalley calibration is exact, the goodness-of-fit criterion for data is obvious. This criterion for calibration in the Kydland-Prescott tradition, however, is not, since parameters are chosen to match data closely and the term 'close' must be defined somehow. Consequently, the goodness-of-fit criterion has formed a subject of debate in macroeconomic-based calibration.[4] As applied general equilibrium modellers begin to experiment with dynamic structures or with incorporating more than

[4] See, for example, Watson (1993).

one observation into elements of their calibration, the choice of goodness-of-fit criterion will become an issue for them as well.

Both methodologies set some parameters based on literature or other non model-dependent criteria, but in applied general equilibrium models the exogenous parameters are predominantly elasticities. Macroeconomic models assign values to a greater variety of parameters. One source of confusion between the two calibration traditions is that in applied general equilibrium modelling the setting of exogenous parameter values does not fall under the calibration rubric, whereas in macroeconomic model calibration it does.

In general, policy modellers employ a variety of functional forms including CES and linear expenditure system (LES), which require the specification of elasticity parameters, whereas macroeconomic modellers tend to employ simple functional forms, such as Cobb-Douglas, which do not. Because of their importance in applied general equilibrium models, the paucity of good elasticity estimates in the literature has been a major source of concern for policy modellers. Although macroeconomic modellers face fewer of these problems because they are less interested in comparative statics and restrict themselves to forms for which these elasticity issues do not arise, weaknesses in the elasticity values available to them have also become a recent source of concern. Browning *et al.* (forthcoming, 1999) highlight some of these issues.

Parameters, and hence calibration of their numerical values, are thus viewed differently in dynamic macroeconomic models and policy-oriented microeconomic models. In the former, the parameters are often the so-called 'deep' parameters of technology and tastes which are considered likely to be static over long periods. Their

values are in themselves of interest to modellers, and calibration is an attempt to recover them from aggregate data. In applied general equilibrium models, the focus is on the comparative static or, less often, comparative dynamic effects on equilibria of a single policy change or simultaneous changes in several policies. Implicitly, modellers believe that the computed effects of changes across equilibria, corresponding to pre- and post-policy change situations, would be robust to the procedures and numerical values of parameters used in replicating the pre-policy change equilibrium. The parameters themselves are, therefore, of limited interest to the analysts.

Thus, although both calibration traditions find parameters without using conventional econometric techniques, differences in the issues addressed using the models and in the nature of the available data, have led to differences in calibration methodology and in the interpretation of both the calibrated parameters and the model solutions. These inherent distinctions have seldom been explicitly identified so that they remain a potential source of confusion and barrier to communication between modellers from the two traditions.

2.3 Calibration in Applied General Equilibrium Models

Although both types of calibrated models are widely used, the research in this thesis focusses on issues surrounding the Shoven-Whalley calibration. The formal description of calibration in applied general equilibrium models offered here is intended to provide a framework for discussing these issues.

An applied general equilibrium model can be written as a system of m simultaneous equations in which a vector of parameters, $\boldsymbol{\alpha}$, and a vector of exogenous variables, $\mathbf{w}$, generate a vector of m endogenous variables, $\mathbf{Y}$. In a simple applied general equilibrium framework, the vector $\mathbf{Y}$ includes an income for each agent, a price for each commodity and factor, and an activity level for each production sector. Agents' factor and commodity endowments are included in the vector $\mathbf{w}$, while policy parameters (such as tax rates), the CES elasticities of substitution, input shares and scale parameters in utility and production functions comprise $\boldsymbol{\alpha}$. Each value in $\mathbf{Y}$ is associated with an equilibrium condition: equilibrium incomes are values that satisfy budget balance constraints for agents; equilibrium prices satisfy market clearing conditions for commodities and factors; equilibrium activity levels satisfy zero profit conditions in production sectors. These equilibrium conditions also form the basis of the more sophisticated structures discussed in Shoven and Whalley (1992), including models with taxes, joint production, nested functions, intermediate demands, decreasing returns to scale production and intertemporal frameworks.

The relationship between $\boldsymbol{\alpha}$, $\mathbf{w}$, and $\mathbf{Y}$ can be expressed in terms of a mapping, $F:\mathbf{R}^{m}\cdot\mathbf{R}^{m}$, such that[5]

$$F(\boldsymbol{\alpha}, \mathbf{w}, \mathbf{Y}) = 0. \tag{2.1}$$

[5] A general equilibrium is characterized by a set of complementary slackness conditions where, if equilibrium prices are zero, excess supply can be positive and where, if activity levels are zero, excess profits can be negative. The discussion here is restricted to the case in which prices and activity levels are strictly positive and the equilibrium conditions are satisfied with equality.

F can be considered to represent the chosen model structure and $\boldsymbol{\alpha}$ to summarize its parameterization. To parameterize a given model, modellers must specify values for the vector $\boldsymbol{\alpha}$. Ideally, they should be able to draw on econometric estimates with well defined statistical properties to assign values to these parameters, but in practice the magnitude of the data requirements make such an approach intractable. Instead, the values for parameters, $\hat{\boldsymbol{\alpha}}$, are inferred from a set of known values for $\mathbf{Y}$ and $\mathbf{w}$, $\hat{\mathbf{Y}}$ and $\hat{\mathbf{w}}$ that solve

$$F(\hat{\boldsymbol{\alpha}}, \hat{\mathbf{w}}, \hat{\mathbf{Y}}) = 0. \tag{2.2}$$

If the dimensionality of $\boldsymbol{\alpha}$ is greater than m, model parameterization becomes the two stage process developed by Shoven and Whalley (1972) and discussed in Mansur and Whalley (1984) and Shoven and Whalley (1992). This procedure partitions the vector of parameters $\boldsymbol{\alpha}$ into two subsets: $\boldsymbol{\alpha}_1$, a set of parameters which the modeller is free to specify exogenously, and $\boldsymbol{\alpha}_2$, the set of 'calibrated' parameters. If $\hat{\boldsymbol{\alpha}}_1$ is the vector of exogenously specified values for $\boldsymbol{\alpha}_1$, calibration yields values for $\boldsymbol{\alpha}_2$, $\hat{\boldsymbol{\alpha}}_2$, which ensure that for $\hat{\boldsymbol{\alpha}}_1$ and $\hat{\mathbf{w}}$, the model produces $\hat{\mathbf{Y}}$ as a solution. The vector of calibrated parameter values is a function of the exogenously specified parameters and the known solution:

$$\hat{\boldsymbol{\alpha}}_2 = G(\hat{\boldsymbol{\alpha}}_1, \hat{\mathbf{w}}, \hat{\mathbf{Y}}), \tag{2.3}$$

where the equations in G are implicit functions of the equations in F.[6]

Once values for the calibrated parameters have been found, the vector of model parameter values $\hat{\boldsymbol{\alpha}}$, and the exogenous variables $\hat{\mathbf{w}}$, can be used in (2.2) to solve for $\mathbf{Y}$ in a 'replication test.' If the solution values for $\mathbf{Y}$ are the same as $\hat{\mathbf{Y}}$, the calibration procedure has found parameters which are consistent.

Policy analysis is undertaken by perturbing some of the model parameters, computing a new equilibrium and comparing the subsequent vector of endogenous variables to the base case vector. The perturbation of the model parameters captures proposed policy changes such as a change in the tax rate, or the removal of a quota. The model's counterfactual solution is the measure of what the new policy scenario may produce. It offers a prediction of the way in which the economy is likely to respond to the change in the policy regime, while the model's base case or pre-change solution is the observed outcome from the economy under the existing policy regime.

2.3.1 An Example of Calibration

Shoven and Whalley's calibration can be illustrated using a simple general equilibrium model with consumption and production. A single consumer is endowed with two factors of production. These factors combine to produce two goods using CES technology, and the consumer has Cobb-Douglas preferences over the two goods. The

[6] Calibration can only be undertaken if the equations in G satisfy the conditions of implicit functions, that is, if the equations of F are continuously differentiable with respect to $\mathbf{Y}$, $\mathbf{w}$, and $\boldsymbol{\alpha}$ and if at $\hat{\mathbf{Y}}$, $\hat{\mathbf{w}}$, and $\hat{\boldsymbol{\alpha}}_1$, the determinant of the Jacobian matrix given by the derivatives of F with respect to $\boldsymbol{\alpha}_2$, is non-zero.

consumer generates demands for the goods by maximizing utility subject to a budget constraint.

The vector of endogenous variables in this model, **Y**, is comprised of: X_i the consumer's demand for good i; Q_i, the quantity produced of good i; P_i, the price of output i; F_i^j, the demand for factor j in the production of good i; and w^1 the price of factor 1. The vector of exogenous variables, **w**, includes the E^j, the consumer's endowment of factor j, and the price of factor 2, w^2 which is set to 1. It is chosen arbitrarily as a numeraire since only relative prices matter in the model.

Finally, the vector of parameters to be calibrated, $\boldsymbol{\alpha}_2$, is comprised of the eight parameters, β_i, a_i^j, λ_i where the β_i are the shares of goods in the consumer's utility function, the a_i^j are the CES share parameters of factor j in the production of good i, and the λ_i are the scale parameters in the production function for good i. The vector $\boldsymbol{\alpha}_1$ consists of σ_i, the two elasticities of substitution in the CES production functions.

The model, F, can be described by the eight conditions:

i)	Factor markets clear:	$\sum_i F_i^j - E^j = 0$	$(j = 1,2)$,	**(2.4)**
ii)	Goods markets clear:	$X_i - Q_i = 0$	$(i = 1,2)$,	**(2.5)**
iii)	Production sectors make zero profits:	$P_iQ_i - \sum_j w^j F_i^j = 0$	$(i = 1,2)$,	**(2.6)**
iv)	Household exhibits budget balance:	$\sum_j w^j E^j - \sum_i P_i X_i = 0$,		**(2.7)**
v)	Fixing of a numeraire:	$w^2 = 1$.		**(2.8)**

The Shoven-Whalley calibration of this model uses equilibrium data to find the values of the parameters which comprise the vector $\boldsymbol{\alpha}_2$. To be used in calibration, however, the data must represent a solution to the model, that is, they must satisfy the model's equilibrium conditions given by equations (2.4) - (2.8). Table 2.1 provides an

Table 2.1

An Example of a Microconsistent Data Set

Transactions Values in Units of Currency

			Expenditures		
	Factor 1	Factor 2	Production of Good 1	Production of Good 2	Purchases by Consumer
Receipts					
Use of Factor 1 in production (inputs)			12	10	
Use of Factor 2 in production (inputs)			8	16	
Production of Good 1 (sales)					20
Production of Good 2 (sales)					26
Consumer's endowments of factors	22	24			

example of such data. The row entries in Table 2.1 denote receipts and the column entries give expenditures, so that together the data are microconsistent in value terms: the value of inputs equals the value of outputs in each sector, the value of consumption equals that of production of each good, and the consumer is on her budget constraint.

If the units convention due to Harberger (1962) is adopted under which the quantities of both goods and services are defined as that amount which sells for one unit of currency, all base case prices in the economy can be set to 1. This convention implies that the value of transactions in Table 2.1 also denotes quantities transacted and that the market clearing conditions also hold.

Cobb-Douglas demands are given by

$$X_i = (\beta_i \textstyle\sum_j w^j E^j)\, P_i^{-1}. \tag{2.9}$$

For specified values of E^j and known solution values X_i, P_i, and w^j, the calibration of the demand parameters is undertaken by calculating

$$\beta_i = P_i X_i\, (\textstyle\sum_j w^j E^j)^{-1}. \tag{2.10}$$

On the production side, the CES factor demand functions are

$$F_i^j = \left(\frac{Q_i}{\lambda_i}\right) \left(\frac{a_i^j}{w^j}\right)^{\sigma_i} \left[\sum_j (a_i^j)^{\sigma_i} (w^j)^{(1-\sigma_i)}\right]^{\frac{-\sigma_i}{(\sigma_i - 1)}}. \tag{2.11}$$

The first step in the calibration of the production parameters is to set values for the elasticity parameters, σ_i. Suppose that either econometric estimation or a literature search yielded the elasticity values $\sigma_1 = 1.2$, and $\sigma_2 = 0.8$. First order conditions from cost minimization allow calibration of the share parameters of factors in production as

$$a_i^j = \frac{w^j F_i^{j\,(\frac{1}{\sigma_i})}}{\sum_j w^j F_i^{j\,(\frac{1}{\sigma_i})}}. \tag{2.12}$$

Substituting the a_i^j into the production function allows the calibration of the scale parameters λ_i,

$$\lambda_i = \frac{Q_i}{[\sum_j a_i^j F_i^{j\,\frac{(\sigma_i - 1)}{\sigma_i}}]^{\frac{\sigma_i}{(\sigma_i - 1)}}}. \tag{2.13}$$

The calibrated parameter values using the data from Table 2.1 and the specified elasticities, are given in Table 2.2.

A modeller would typically substitute the calibrated parameter values set out in Table 2.2 into the model given by equations (2.4) - (2.8), using the functional specifications (2.9) and (2.11), to ensure that the equilibrium solution values are the same as those given by the data in Table 2.1. This replication test provides assurance that no errors are present, either in the calibration calculations or in the model coding.

The possibility exists that the model has multiple equilibria and that the replication test might fail because the model solves for an equilibrium other than that

Table 2.2

Calibrated Parameter Values for the Example Model Using the Data in Table 2.1

utility function share parameters	β_1	$\doteq 0.43$
	β_2	$\doteq 0.57$
production function share parameters	a_1^1	$\doteq 0.58$
	a_1^2	$\doteq 0.42$
	a_2^1	$\doteq 0.36$
	a_2^2	$\doteq 0.64$
production function scale parameters	λ_1	$\doteq 1.97$
	λ_2	$\doteq 1.93.$

of the base case data. Numerical examples of multiple equilibria have been constructed by Kehoe (1985) for simple Cobb-Douglas economies with a small number of production activities. However, where smooth production functions of the Cobb-Douglas or CES variety are used, uniqueness is the more likely outcome.[7] *Ad hoc* tests, undertaken with applied models, seem to confirm this view.[8]

Although this example is simple, the same calibration approach can be used for large scale models. Piggott and Whalley (1985) use a model of the UK with 100 households, 33 productive sectors, and 29 traded goods. Including the intermediate production structure, the model uses around 20,000 parameter values. Models of these dimensions are not exceptional. An even larger model, the ORANI model of the Australian economy described in Dixon, Parmenter, Sutton and Vincent (1982), identifies 115 commodities and 113 industries in its base period input-output data.

2.4 Calibration in Context: The Modelling Process

Although a model's calibrated parameters are mechanically determined by the relationship in equation (2.3), their values depend on a wider decision process. The choice of model, functional forms, elasticity values, data, and data adjustments all contribute to the calibrated parameter values. The calibrated parameters form part of the base case model structure and partly determine the counterfactual solution. The steps

---

[7] See Kehoe and Whalley (1985).

[8] Such tests include setting the model's starting values to a slightly displaced version of the initial equilibrium solution and checking that the model calculates the initial equilibrium as a solution, and approaching equilibria at different speeds and from different starting points.

in this wider modelling process, which are summarized in Figure 2.1, are considered individually below.

2.4.1 Model Choice

The choice of model may be the major Achilles' heel in the use of calibrated models for empirical investigation, both because models are not tested against one another, and because the precise model form can have a major influence on results. Reference to widely used theoretical structures is usually an insufficient basis on which to choose models, especially since much of the theoretical discussion is oriented towards showing how changes in the model structure can change the qualitative model predictions. Model selection based on theoretical literature may sound appealing, but the literature does not offer guidance on the precise specification of the model, nor does it provide the criteria under which such a choice should be made.

An example illustrates how the conclusions of calibrated models can change substantially with the model structure. In 1962, Harberger performed some of the earliest general equilibrium simulations, implicitly calibrating a two sector model of the US economy and evaluating counterfactuals to show that a tax on one factor in one sector (the tax on capital in the corporate sector) was borne fully by that factor even if it was mobile between the two sectors. In fifteen years of subsequent literature, the addition of more sectoral disaggregation, partially mobile factors and other features failed to change the basic result; capital still bore the burden of the corporate tax.

Figure 2.1

The Context for Applied General Equilibrium Model Calibration

Model Choice
•Based on Theoretical Literature

↓

Model Specification
•Model Dimensionality
•Functional Forms

↓

Collection of Data
•National Accounts
•Surveys
•Government Publications

↓

Reconciliation into a Benchmark Equilibrium Data Set
•Formal Algorithms
•*Ad hoc* Adjustments

↓

Elasticity Specification ← ← ←
•Literature Based
•'Best Guesses'
•Estimation

↓

→ → → **Calibration**
•Given by Equation (2.3)

↓

Replication
•Absence of Coding Errors
•Calibration Check

Sensitivity Analysis
• Systematic Analysis
• Limited Analysis

← ← ← **Base Case Model**
•Initial Equilibrium
•Basis for Comparison

↓

Counterfactual
•Prospective Policy Change

↓

Interpretation of Results → → →
• Conditional on Model Assumptions

In the late 1970s, however, simulations showed that if the US economy were modelled as facing a perfectly elastic supply function for capital, instead of facing the fixed endowment, inelastic supply function scenario of Harberger, Harberger's result would reverse. Capital simply could not bear the burden of the tax in this situation, and it had to be shifted elsewhere. Two model structures could yield this feature - one with perfect international capital mobility, or one with an intertemporal structure with savings (consumption smoothing) where the savings elasticity and hence, the supply elasticity for capital within a period, is high. Modifying the original Harberger structure in either of these two directions changes the essential result.

If calibrators reject the notion of model testing, and base their model selection on theoretical literature, objective criteria for choosing models may be unattainable. Instead, modellers should qualify their results by stating more forcefully than they have, that the model output is conditional on the specific choice of model. They can use the theoretical literature to identify which features of their model results are sensitive to which structural assumptions, and then modify these assumptions to assess numerical sensitivity. Developing this direction in calibration would allow modellers to explore the structural sensitivity of model results.

2.4.2 Dimensionality and Functional Forms

The dimensionality of a model varies with the question to be answered. Typically, models built to illustrate theoretical propositions are parsimonious, capturing only the relevant economic relationships. Models designed to shed light on specific policy

questions for real economies are more detailed and their complexity reflects the nature of the specific policy questions to be analyzed. Hence, a model that addresses investment policy will typically include an intertemporal representation, whereas a model that explores interhousehold tax incidence effects requires explicit representations of household types. In such models, domestic structures are emphasized, while the rest of the world is typically presented as a single agent. Conversely, models that focus on trade policies have explicit representations of several countries or trading regions, but employ simple structures to represent their domestic economies. Model detail often centres on the sectors and agents most likely to be affected by the policy change under question, while the remainder of the economy is modelled at a relatively more aggregate level.

The dimensionality of a model is limited by data availability, since results from a model that identifies agents or sectors for which no data exist are not credible. Paradoxically, an abundance of data can also influence dimensionality, as modellers face pressure from policy makers to include economic detail because it is available and, hence, is thought to make the model more realistic. Models with too much detail, however, impede an understanding of economic processes that drive the model results. Highly detailed applied general equilibrium models have developed reputations as black boxes into which a policy change is fed as an input and from which a set of results emerges with little explanation. In these models, the interactions that drive the model results can easily become obscured. Hence, the modeller's challenge is to balance clarity with realism in the presence of data constraints.

Once the dimensionality has been determined, modellers must also specify functional forms for the behavioural relationships in the model. They typically employ the family of 'convenient' functional forms for which the solutions to optimization problems can be obtained analytically. Cobb-Douglas and CES functions are widely used. Cobb-Douglas functions are simple, but highly restrictive, since they imply unitary income and uncompensated own-price elasticities, and zero uncompensated cross-price elasticities. In contrast, CES functions relax the unitary uncompensated own-price and zero cross-price elasticities of the Cobb-Douglas functions, but do so only by adding an additional parameter - the elasticity of substitution. Modellers often have information about the structure of an economy, such as literature-based elasticity estimates, which they wish to include in their model calibration. To incorporate this information, extra parameters are often injected into the model using nested CES functions, where the elasticity parameters enter at the various nests in the structure.

However, both CES and Cobb-Douglas preferences are homothetic and so yield demand functions which have unitary income elasticities. If income elasticities are thought to be significantly different from unity, some other functional form is needed, and a Stone-Geary/Linear Expenditure System with a displaced origin for utility measurement is commonly used. The minimum consumption requirements in such a system, which can be combined with either Cobb-Douglas or CES, are typically calibrated so that they reproduce literature estimates of income elasticities of demand in the neighbourhood of the base case equilibrium.

Some modellers have moved beyond this broad class of convenient functional forms to use variants of flexible functional forms, typically trans-log. The basis for

rejecting the convenient forms lies in the empirical results of econometric studies which reject the separability implicit in Cobb-Douglas and CES functions. The major drawback to using more flexible functional forms is that they are not always globally convex. Because the policy changes analyzed in many models can lead to a counterfactual equilibrium that is far from the initial equilibrium, the use of globally convex functions is often necessary to compute a model solution.

As with the choice of model structure, the functional forms used in a model should be attuned to the issue under investigation. Consider, for example, a trade model which explores the claimed long-term decline in the terms of trade of commodity-exporting developing countries, and builds on the argument from Prebisch (1962) and Singer (1950) that developing countries export necessities and import luxuries, such as capital goods. Such a model requires income elasticities of demands that are different from unity to reflect the feature that growth in both the developed and the developing countries will adversely affect the developing country's terms of trade. This feature emerges if the income elasticities of import demand in developed countries are less than one while those in developing countries are greater than one. Using models with either Cobb-Douglas or CES preferences will not meet these conditions and a different functional form is needed. On the other hand, if the income effects from the change considered in the model are thought to be small compared to the relative price effects, a model with homothetic preferences may suffice.

As with the choice of model structure, decisions about dimensionality and functional forms are dependent on the research question. Both sets of decisions need to

balance simplicity with realism; but other pragmatic considerations such as computational feasibility also enter the choice.

2.4.3 Collection of Data

Once the model structure has been specified, modellers begin the process of data collection. The data to which model parameters are calibrated are typically derived from several sources, including household expenditure surveys, input-output tables, government administrative records, statistics from taxation departments, and national income accounts. Harnessing these data presents many challenges for the modeller, some of which are detailed in St. Hilaire and Whalley (1983). The levels of sectoral, household or product aggregation can differ among data sources. Definitions of terms can vary, and do not necessarily accord with the model requirements. Classifications in one data source may differ from those in others. For example, where one set of accounts may consider informal sector firms as those with fewer than ten employees, another may define them as firms which enter no formal contracts. Gaps can occur, with no estimate available for some components of the required data. Measurement errors abound; an estimate of the same variable in one data source may differ sharply from that in another. Data sources themselves also vary in their reliability because collection techniques and methods of analysis differ among researchers and institutions. Including more than one country in a data set compounds these consistency problems.

The issues of data reliability and compatibility mean that uncertainty surrounds the data which are used in calibration. Uncertainty in these data values translates into

uncertainty in the calibrated parameter values and, ultimately, in the model results. Chapter 3 introduces a methodology that allows modellers to report the sensitivity of the model results to the uncertainty in the initial data values.

2.4.4 Deriving a Benchmark Equilibrium Data Set

Calibration, as given by equation (2.3), requires data that represent an initial model equilibrium. The basic data, however, seldom meet the consistency requirements of an equilibrium, and modellers typically undertake adjustments to ensure that they do. In general, data adjustments involve two intertwined processes. The first, which selects single values of each data point required for model calibration, is undertaken when the data are collected. It includes the choice of one data source over another, the approximation of a desired classification with one available in the data, and the method of aggregation. By nature, this process is model and data specific.

The second is one of reconciling these point estimates into a microconsistent form so that the data meet the equilibrium conditions of the model. This process, which falls into a class of matrix adjustment problems that has been studied both in economics and in other disciplines, can be executed in a systematic way. The systematic adjustment techniques typically assume that the data are in matrix form.

The initial point estimates can be placed in a square transactions matrix such that a row, representing receipts, and a column, representing outlays, is assigned to each market, production sector, and agent defined in the model. The process of adjustment

is one in which the initial matrix is transformed into a 'biproportional' matrix. Biproportionality is a balancing condition for a square matrix $[x_{ij}]$ in which

$$\sum_j x_{ij} = \sum_i x_{ij}, \ \forall\, i = j. \tag{2.14}$$

The adjusted biproportional matrix is termed a 'benchmark equilibrium data set' (BED). If the Harberger units convention is adopted, the equilibrium conditions of the model are reflected in the BED's biproportionality condition: budget balance holds for agents (incomes equal expenditures), sectors make zero profits (sales equal production costs), and because prices are unity, markets clear (quantities demanded equal quantities supplied). Table 2.1 provides a simple example of a BED.

Although no formalized statement of the procedure exists, the reconciliation of initial data estimates into a biproportional BED for large applied models is typically undertaken in two stages. The first finds consistent values for aggregate values: total consumption, output, and intermediate demands. At this stage, matrix biproportionality is the paramount restriction on data. For example, the total supply of each good in the model must equal the total demand, typically defined as the sum of government consumption, exports, intermediate demand and private domestic consumption. The initial, unadjusted values of these aggregates rely heavily on national accounts data.

The second stage makes submatrices of data consistent with the aggregates found in the first stage. It draws on formal algorithms for balancing a matrix subject to consistency with a set of control totals. So, for example, where the model identifies

more than one private consumer, the aggregate private domestic consumption for each good can serve as the row control totals for the household consumption submatrix, and the total disposable income by household type can provide the column totals. Similarly, aggregate intermediate demand for a good gives the row totals for the intermediate demand matrix, and total expenditures on intermediate goods by sector (typically found as the residual of total receipts and expenditures on value added) provide the column control totals. Because the control totals are consistent with the biproportionality constraint, the values of the submatrices that are consistent with those control totals also fulfil the biproportionality constraint for the BED as a whole.

The information required to specify the submatrices in the benchmark data set is more detailed than is true for the aggregate values. Initial estimates for the elements of the intermediate demand matrix can be derived from input-output matrices, while those for the household consumption matrix can be derived from household expenditure surveys. Unlike national accounts, such detailed data are unlikely to be collected annually and matrix adjustment is achieved by updating earlier years' estimates so that they are consistent with the benchmark year control values.

The process of data adjustment for applied general equilibrium models has never been standardised, but most modellers follow broadly similar approaches. They typically employ *ad hoc* algorithms to derive consistent aggregate values, and resort to

formal adjustment algorithms, particularly the RAS (Row and Column Scaling) algorithm, to derive consistent consumption and production submatrices.[9]

Several formal algorithms exist to adjust an unbalanced matrix. One such algorithm is RAS, attributed to Bacharach (1970), in which the rows and columns of a matrix are scaled and sequentially updated by the ratio of the matrix row or column sum to the control total row or column sum. This adjustment algorithm allows large initial data entries to deviate more from their initial values than small entries. Other adjustment algorithms, most notably those using weighted constrained quadratic minimization, also exist. One algorithm, the Stone (1978) and Byron (1978) adjustment algorithm, is particularly appealing in that it incorporates information about the reliability of the data so that the least reliable data change more from their initial values than do the most reliable data. More recently, Golan *et. al.* (1994) have introduced an algorithm based on maximum entropy, which is related to RAS, and which can also generate a balanced matrix from incomplete data.

Although several formal adjustment algorithms exist, the applied general equilibrium modelling literature offers no guidelines about which one to choose or what the effects of that choice might be for the model results. Chapter 4 of this thesis explores some of these issues and shows that the choice of adjustment algorithm can be an important stage of the modelling process.

---

[9] See for example, Piggott and Whalley (1985), Chia, Wahba and Whalley (1992) and Brixen (1995).

2.4.5 Elasticity Specification

In non-Cobb-Douglas models, the adjusted matrix does not suffice for calibration. Calibration also requires specified values of the elasticity parameters. The values for these elasticities are obtained, where possible, from literature-based econometric estimates. Typically, literature estimates of important own-price elasticities on the demand side are used as a basis for choosing elasticities in CES preferences so that the implied point estimates of demand elasticities in the neighbourhood of the benchmark equilibrium are consistent with the literature estimates. Production elasticities are similarly obtained. Literature elasticity estimates, however, are scarce and dated, not least because the emphasis in econometric research in recent years has moved away from parameter estimation. Furthermore, the aggregation, the regional classification, and other definitions of existing elasticity estimates may not be compatible with those of the model.

The current situation with respect to literature-based elasticity estimates for use in calibrated models is poor. No estimates exist for many types of elasticities. Others have multiple, and sometimes contradictory, estimates within wide ranges. Classifications in models do not necessarily match those from which the literature-based values are derived. Furthermore, the structure of the econometric model from which the elasticity parameters have been estimated is unlikely to be similar to that of the applied general equilibrium model in which they are used, especially since the estimation procedures are unlikely to have imposed a general equilibrium structure on the data.

Elasticities that have been estimated for different classifications are routinely adopted for model use, so that for example, an estimate for the demand elasticity for food might be used to provide the demand elasticity for cheese, even though inter-food substitution is a key feature in the model. If estimates are deemed implausible, as can be the case with trade elasticities, they are often either ignored, or arbitrarily scaled, sometimes by as much as 50 percent.

Modellers occasionally undertake their own estimation for a model's elasticity values. Typically, however, the number of elasticities in an applied general equilibrium model is prohibitively large and insufficient data exist for the estimation of them all. As a result, modellers focus on estimating the elasticities which are perceived to be the most relevant for the model structure. Hence, modellers looking at trade policies would concentrate their efforts on estimating trade elasticities, while tax modellers would focus on income elasticities.

Faced with a relative absence of elasticity estimates, many of a model's elasticities are likely to be derived using 'best guesses.' Except where conventional wisdom dictates, such as the income elasticity of food being less than unity, modellers tend to follow the 'idiot's law of elasticities' - all elasticities are 1 unless evidence suggesting otherwise exists. Modellers also refer to 'coffee table elasticities' where informal discussions and opinions around the coffee table determine whether a value of, say, 0.5 or 2.0 is chosen.

2.4.6 Calibration, the Base Case Model and Replication

Once the BED and the elasticities have been specified, the mechanical process of calibration is described by equation (2.3). Calibration completes the base case model specification. In a replication test, the values for the calibrated parameters are substituted into the model and the model is solved. If calibration has been undertaken correctly and if the model is free of coding and structural errors, the base case model solution values will match the values in the BED.

The replication test, however, does not guarantee an absence of errors. Because the net-of-tax prices in the benchmark case are typically constructed to be unity, model structure errors in which prices are incorrectly multiplied will not cause the replication check to fail, but will only become evident when no reasonable solution can be found in the counterfactual simulation. Hence, a successful replication check is a necessary component of the modelling process, but is not a guarantee of model soundness.

2.4.7 Counterfactual Simulation and Interpretation of Results

Counterfactual simulations are undertaken by perturbing parameters in the base case, and solving the model with the new parameter configuration. The effects of the parameter change are gleaned by comparing the counterfactual model solution to the initial model solution. The interpretation of these results is an exercise which requires considerable caution. Modellers resort to numerical simulations with specific parameterizations when algebra and analytics fail to give clear results. These simulations represent a logical progression from theory, but they give less general

results because any numerical findings are conditional on the particular numerical specification used.

The use of data adjustments and the absence of statistical structures in deriving the numerical specification, however, precludes the use of these models for forecasting. While basing their conclusions on the best available data, modellers should not pretend that their model results yield anything other than indications of the relative orders of magnitude for possible policy adjustments in the economy.

Faced with all the arbitrariness in the model's specification, the value of modelling results lies in providing insights, rather than point estimates or forecasts. Modellers use the model results to answer broad questions and provide quantitatively informed insights. Are effects of a policy change large or are they small? Are they opposite to received wisdom, and if so why? If no previous studies of an effect exist, what might be an initial estimate? What are the relative magnitudes of effects? Paradoxically, the very framework that allows a detailed specification of the economic system introduces uncertainty into the model conclusions by virtue of its requirement for highly disaggregated and, inevitably, approximate data.

Finally, the ambiguity noted in theoretical literature, that even qualitative results depend upon assumptions, is not avoided merely by using a calibrated model. Calibration, *per se*, gives no guide as to how to choose a model, and the results of subsequent policy evaluations are all conditional on the chosen model. The calibration of applied general equilibrium models also implies no model testing because many different models with different structures could, in principle, be calibrated to the same data set. Unlike econometric exercises, these modelling efforts are simply theory with

numbers, the aim of which is to provide model conditional insights, either for policy input or for the better understanding of economic processes.

2.4.8 Sensitivity Analysis

Because of the uncertainties in the numerical specification of a model, modellers typically test the robustness of their results using some form of sensitivity analysis. The overwhelming majority of these analyses focus on the effects of the choice of model elasticities. The most common approach to sensitivity analysis is termed 'limited sensitivity analysis' by Wigle (1991). In limited sensitivity analysis, the modeller subjectively identifies the important model elasticities, so that a trade modeller might, for example, list import demand elasticities as the subjects of sensitivity analysis. The central values of these key parameters are then perturbed by a 'reasonable' amount, the model is solved, and the results are reported for the alternative elasticity configurations. The process is repeated for several values of the key elasticities. While this procedure can give some sense of whether model results are fragile, it provides no meaningful quantitative measure of robustness.

More rigorous statistical sensitivity analysis procedures have also been developed. Wigle (1991) discusses two classes of systematic elasticity sensitivity analysis used in reporting applied general equilibrium model results, both of which require the modeller to assign probabilities to alternative elasticity configurations. Conditional systematic sensitivity analysis (CSSA) infers the distribution of the model results by computing a series of solutions as each elasticity is varied while the others

remain constant. Unconditional systematic sensitivity analysis (USSA) computes model results over the entire grid of possible elasticity configurations. USSA is the most thorough and therefore, the more preferable response to criticisms of elasticity specification, but for most models the computational requirements of such a procedure are prohibitive.[10]

Pagan and Shannon (1985) develop an approximation method for performing unlimited systematic sensitivity analysis. Instead of solving the model for each point in the elasticity space explicitly, their procedure analyzes the effects of altering elasticity parameters in a region surrounding the model solution. Because their sensitivity procedure relies on calculations made using a linear approximation of the model solution, which is a function of the elasticity parameters, the computational requirements are considerably less than in unconditional systematic sensitivity analysis, while the procedure retains the flexibility to examine the effects of simultaneous elasticity variations. The Pagan-Shannon approximation procedure is applied in Pagan and Shannon (1985, 1987) and Wigle (1991).

Other sensitivity procedures, in which modellers map *a priori* information about elasticity probabilities into the model results, have also been developed. Harrison and Vinod (1992) and Harrison *et al.* (1992), develop and apply a global sensitivity analysis procedure in which the model is solved for a sample of elasticity configurations. Their procedure relies on sampling from discrete representations of what are usually continuous elasticity probability density functions. DeVuyst and Preckel (1997) argue

[10] Wigle (1991) calculates that a USSA using 5 values for each elasticity in an 18 elasticity parameter model would require more than 3 trillion model solutions.

that the methodology of Harrison and Vinod introduces an identifiable source of bias into the sampling procedure and propose an alternative way of finding discrete approximations to the continuous probability density functions, based on Gaussian quadrature. In both approaches, the model results are weighted by the probability of each elasticity configuration used in their derivation. Repeated sampling allows the modellers to build expected values and confidence intervals for the model results.

Sensitivity analysis completes the modelling process. It represents a way for modellers to address some of the weaknesses inherent in a methodology that requires subjective judgement at many junctures, and which employs estimates for a large number of diverse data points.

2.5 The Econometric Critique of Calibration

The calibration methodology used for applied general equilibrium models has been criticized in Jorgenson (1984) and more recently by McKitrick (1995, 1998) on several grounds. They argue that the data pre-adjustments in the process of implementing calibration introduce untraceable bias into the data and hence, into the parameters and ultimately the model results. The use of a benchmark year for calibration also enters their critique since any anomalies in the economy for that year can be transmitted to the calibrated parameter values, and hence, to the model results. They highlight the inadequacies of the elasticity estimates in applied models, and argue that the reliance on CES and Cobb-Douglas functional forms is restrictive and unrealistic. This restrictive class of functional forms precludes complementarities, and incorporates

elasticities of substitution which are independent of prices and which thus, unrealistically constrain behavioural responses in counterfactual simulations. McKitrick's (1998) illustration that a model's functional structure has a large effect on results highlights the shortcomings of relying on these functional forms.

Jorgenson and McKitrick's proposed alternative is the simultaneous estimation of all of a model's elasticities and share parameters using time series data. This approach allows elasticity estimation which is fully consistent with the definitions of variables employed in the model, and does not require the use of restrictive functional forms. The statistical basis of estimation isolates systematic effects from random noise, and the use of unadjusted time series data precludes the introduction of pre-adjustment bias.

Explicit econometric approaches to applied general equilibrium modelling have thus far been limited to a handful of papers: Clements (1980), Jorgenson (1984), Jorgenson, Slesnick and Wilcoxen (1992), McKitrick (1995), and McKitrick (1998). Most of these econometric general equilibrium models, however, estimate model subsystems rather than incorporating the full set of cross-equation equilibrium restrictions.

If estimation is superior to calibration in so many ways, why it has not been more widely adopted? One issue is the difficulty in imposing the equilibrium solution concept, which is central to general equilibrium analysis, as a series of cross-equation restrictions in estimation. Another is the paucity of time series data on the variables of interest for the questions that are addressed in calibrated models. The estimation of large dimensional models, or models which focus on variables that are not measured in

national accounts data, may be intractable. The effort required to generate the single observation required for calibration can itself be formidable, and extending the process to include time series observations may be close to impossible. For example, modellers must frequently update an earlier year's input-output matrix as an approximation to that of the benchmark year, because most countries do not produce annual input-output tables.[11]

The econometric approach also precludes the use of some simplifying techniques commonly employed in applied general equilibrium models. One such technique is the Harberger (1962) convention, whereby the units of quantities defined in the model are given by that quantity which sells for one unit of currency in the base period. This convention allows the modeller the simplification of representing heterogenous quantities in a homogenous manner, both in data and in the model. For example, if labour inputs were to be measured as hours worked, some correction would have to be made for different levels of labour efficiency and skill. The use of this assumption also reduces the number of variables required in the model; the modeller need only collect data in value terms, rather than in separate price and quantity terms. Such a convention, however, creates time-dependent units that make the interpretation of the results of counterfactual policy simulations a somewhat delicate issue. How, for example, should the modeller interpret a 10% increase in the price of a non-electrical machinery aggregate in the counterfactual equilibrium? How should labour of different

[11] Furthermore, where they are produced, they are often generated by updating a previous year's table rather than by undertaking new production surveys.

skills be aggregated, when compared to a cost of one unit in the benchmark equilibrium? Such a convention makes time series estimation virtually impossible.

Faced with the weaknesses presented in the econometric critique, why do policy modellers persist with their work? The answer lies in the lack of practical alternatives. Policies will be decided with or without numerical input. Modellers' underlying belief is that imperfect analysis is better than no analysis. To contribute to debate on the social issues of the day a modeller must make the best use of the available information, rather than refraining from any analysis until every parameter is definitively tied down. Modelling is a way of harnessing available information to contribute to policy making by raising the level of debate - an argument which would clearly be rejected by those whose advocate an exclusively positivist approach to research in the social sciences and to producing policy recommendations.

2.6.1 Responses to the Critique

Instead of adopting econometric estimation, modellers have responded to certain aspects of the Jorgenson-McKitrick critique within the calibration paradigm. The weakness of the elasticity estimates has been addressed via the sensitivity analysis procedures in Wigle (1991), Pagan and Shannon (1985, 1987), DeVuyst and Preckel (1997), Harrison and Vinod (1992), and Harrison *et al.* (1992), discussed earlier. Modellers need no longer rely on restrictive functional forms; a fully flexible, globally regular functional form, has been developed by Perroni and Rutherford (1998).

The issues of the adjustments made to data for calibration purposes have been largely ignored in the modelling literature, but the possible pitfalls of drawing conclusions from a single and possibly unrepresentative, single year benchmark observation have been explored in several papers. Roberts (1994) examines the significance of the choice of benchmark year in a model of Poland by calibrating a model to BEDs for five different years, and concludes that model results are robust to the choice of year. Adams and Higgs (1990) also address this problem, arguing that year-specific effects can be mitigated by averaging several years' data to create a synthetic benchmark data set. Using the Australian ORANI model, they illustrate how agricultural data from an abnormal 'year of record' can affect policy conclusions.

The introduction of untraceable bias to the model parameterization through pre-adjustments remains a largely unaddressed issue. One exception is Wiese (1995), who derives two BEDs using alternative accounting assumptions for employer contributions to health insurance and traces the effects of these assumptions on model results. His experiments indicate that the model results are affected by the accounting conventions used in the data. Different accounting conventions could, in principle, also affect econometric estimates since such conventions serve as identifying assumptions.

Thus the Jorgenson-McKitrick critique has provided some impetus for modellers to improve the current calibration procedures. The alternative which they propose - econometric parameterization- certainly responds comprehensively to their critique, but on practical grounds it remains outside the reach of most applied modellers.

2.6 Conclusion

Calibration is a term used to describe the non-econometric parameterization of economic models. Although calibrated models are common in the analysis of macroeconomic issues, the most prevalent use of calibration is to parameterize applied general equilibrium models. The calibration of applied general equilibrium models is, however, an imperfect procedure, but the alternative, econometric estimation, is seldom feasible for the types of questions that modellers wish to examine. The choice for modellers, then, is to either abandon numerical analysis of the problem at hand, or to proceed with a flawed technique. Since governments debate and implement policies with or without technical input, and since most modellers believe that imperfect numerical analysis can contribute to the policy debate better than no analysis, they continue in their endeavours. Instead of abandoning calibration, modellers seek ways to mitigate the effects of the weaknesses in their technique. The research presented in Chapters 3 and 4 of this thesis forms part of this ongoing search.

This thesis also explores modifications to the standard calibration technique that allow modellers to address a wider range of issues. Calibration enables modellers to undertake economic analysis in situations where the time series required for estimation do not exist, but for which a single benchmark observation can be constructed: highly disaggregated economies, economies with informal sectors, countries with poor quality data, and historical economies. Although the relevant time series for the estimation of such models may be unavailable, a historian may be able to obtain a second data observation. Chapter 5 shows how the calibration of applied general equilibrium models

can be adapted to incorporate the information in a second observation so that historians can decompose the individual and interactive effects of several simultaneous shocks to an economy.

Chapter 3

Extended Sensitivity Analysis

3.1 Introduction

The values of the parameters used in economic models are typically surrounded by uncertainty, and one technique that allows modellers to quantify the robustness of their model results to this uncertainty is sensitivity analysis. This chapter proposes two related sensitivity analysis procedures for applied general equilibrium models: 'calibrated parameter sensitivity analysis' (CPSA), and 'extended sensitivity analysis.' CPSA is a procedure that produces confidence intervals for the model results based on the uncertainty in the data used to generate a model's calibrated parameters. Extended sensitivity analysis combines CPSA with existing sensitivity procedures for elasticities to generate confidence intervals that reflect the uncertainty in the values of all of a model's parameters.

Several sensitivity analysis techniques for applied general equilibrium models have been developed in the literature (Pagan and Shannon, 1985; Pagan and Shannon, 1987; Wigle 1991; Harrison and Vinod 1992; Harrison, Jones, Kimbell and Wigle 1992; DeVuyst and Preckel 1997), but these procedures are incomplete because they can only capture the robustness of the model results to uncertainty in a subset of the model's parameters. They focus on the values of exogenously assigned elasticity parameters,[12] while the calibrated parameters - those that are obtained from combining

[12] Elasticities are not the only exogenous parameters for which sensitivity analysis has been undertaken. Rutström (1991), for example, conducts sensitivity analysis over the values of the minimum requirement parameters in a linear expenditure system.

elasticity information with flow or stock data - are excluded.[13] This omission stems partly from the perception that whereas a model's elasticity values are often obtained through informed guesswork and can therefore be very uncertain, the calibrated parameter values have a more solid empirical foundation in data. However, the considerable uncertainty that typically surrounds the data used for calibration introduces uncertainty into the calibrated parameter values, making them also candidates for sensitivity analysis. This uncertainty arises initially through measurement error and is augmented by the consistency adjustments made to the data so that they meet the equilibrium conditions of the model.

Because the calibrated parameters determine the static specification of the modelled economy and because they typically comprise the majority of the model's parameters, their omission from previous sensitivity analysis procedures represents a serious gap in the modelling literature. This chapter fills this void by developing and illustrating a 'calibrated parameter sensitivity analysis' methodology, termed CPSA. It then proposes an 'extended sensitivity analysis' procedure which integrates CPSA with DeVuyst and Preckel's (1997) elasticity sensitivity analysis methodology, and allows the modeller to measure the robustness of the model's results to uncertainty in the model's full numerical specification.

[13] In so far as the calibrated parameters are functions of the exogenously specified parameters, previous sensitivity analyses capture some of the uncertainty in the calibrated parameters. The approach here, however, provides a framework for addressing the *full* uncertainty in the calibrated parameter values.

3.1.1 Technical Issues

One of the difficulties in developing CPSA is that the basic approach adopted in elasticity sensitivity analysis cannot be applied to the calibrated parameters. This approach involves perturbing the central model elasticity values, solving the model using those perturbed values, and comparing the ensuing model results to the central case results. Elasticity sensitivity analysis requires each elasticity perturbation to be associated with a unique model result.

Unlike the exogenously specified elasticities, however, the calibrated parameters cannot be individually perturbed. A given perturbation to one calibrated parameter would require changes to other parameters to maintain the consistency conditions of the base equilibrium. No such realignment of the remaining parameters is unique, however, and therefore no single change to the model results can be determined from a given perturbation.

For example, consider a model with a fixed labour endowment and several production sectors. If the modeller wishes to observe the effects on model results of changing the input share of labour in one production sector, the input share of labour in at least one other sector would also have to change to maintain the base-period equilibrium condition that the labour market clears. The modeller, however, has no way of determining which of the remaining production sectors should absorb this change. Because several options exist for meeting the model's consistency requirements, and because each could lead to a different model result, the initial perturbation does not lead to a unique change in the model results. Sensitivity analysis for the input share of labour

in a single production sector is, therefore, impossible. A similar argument holds for any individual calibrated parameter value.

The innovation in CPSA, however, is to conduct sensitivity analysis over entire configurations of calibrated parameters instead of individual parameter values. CPSA hinges on the fact that the calibrated parameters are the end product of a longer process: raw data are adjusted into a benchmark equilibrium data set (BED) that meets the equilibrium conditions of the model; and together with the specified elasticity parameters, the BED determines the joint values of the calibrated parameters.

At the heart of CPSA is the insight that if the data adjustment procedure and the values of the elasticities are kept constant, a given collection of raw data will lead to a single configuration of calibrated parameter values and a unique model solution. Under these two constancy conditions, sensitivity analysis with respect to uncertainty in the values of the model's calibrated parameters is equivalent to sensitivity analysis with respect to uncertainty in the raw data values. Unlike the elements of the BED and the calibrated parameters, the unadjusted raw data have no consistency restrictions on the values they can assume, so that they can be individually perturbed. Their perturbation is the essential feature of CPSA.

The CPSA methodology requires the modeller to specify a plausible perturbation of each element of the raw data from its central case value. Together, these perturbed elements form a perturbed raw data set, which is balanced with the same adjustment procedure that was used to balance the original, unperturbed raw data. The ensuing BED leads to a set of calibrated parameter values that are then used to solve the model. *A priori* information about the data's reliability allows the modeller to assign a

probability of being true to the perturbed raw data set, and hence, to both the set of calibrated parameters that is derived from that data set, and to the ensuing model results.

This process is repeated for a series of perturbations. The resulting series of solution values and their associated probabilities allow the modeller to build a picture of the sensitivity of the model results to uncertainty in the raw data. This sensitivity is expressed using confidence intervals for the model results. Thus, CPSA translates the modeller's knowledge of uncertainty in the raw data, through the calibrated parameters and into a measure of robustness for the model results. In doing so, it completes the framework for reporting the model's sensitivity to its full numerical specification. Extended sensitivity analysis simply combines CPSA with existing sensitivity analysis procedures for elasticity parameters so that the modeller can report a measure of the robustness of the model results to the joint uncertainty in the raw data and the elasticity values.

This chapter is organized as follows. Section 3.2 presents and illustrates the CPSA methodology using a simple applied general equilibrium model. Section 3.3 proposes and applies an extended sensitivity analysis procedure in which CPSA is combined with the elasticity sensitivity analysis procedure of DeVuyst and Preckel (1997). The application examines the sensitivity of personal tax incidence results in a model of Côte d'Ivoire due to Chia, Wahba, and Whalley (1992), to the parameters calibrated from the household consumption expenditure data and to selected elasticities. Section 3.4 concludes with comments on the implications of the procedure.

3.2 Calibrated Parameter Sensitivity Analysis (CPSA)

The basic approach of the CPSA procedure proposed in this section is to perturb the initial values of the unadjusted data from which the calibrated parameters are ultimately derived, and to observe the effect of the perturbation on the model results.[14] Despite the absence of consistency restrictions on the values the data can assume, the perturbations in CPSA are not arbitrary. CPSA uses information about the reliability of the individual data elements to perturb the data in such a way that the probability that the collection of perturbed data is true, can be identified.

This attachment of a probability to a perturbation is an indispensable component of any sensitivity analysis. Without some measure of the likelihood of the specified change, the modeller cannot determine the significance of that change for the model results, and the sensitivity analysis becomes meaningless.

As an example, consider a model for an economy where the best initial estimate for GNP is $747.3 billion and in which a tax change leads to a welfare loss of $1.4 billion. Suppose that the modeller wishes to observe the sensitivity of this model result to the initial value of GNP. A sensitivity experiment is performed in which the initial value of GNP is changed to $740 billion, and this change yields a welfare loss of $1.39

[14] Although the idea of systematically perturbing the unadjusted data to observe the effect on model results is absent from the applied general equilibrium modelling literature, several modellers have explored the sensitivity of model results to alternative BEDs: Roberts (1994) examines the effects the choice of benchmark year for the BED; Adams and Higgs (1990) argue for the use of a synthetic 'typical' BED rather than one derived from a particular 'year of record'; Wiese (1995) derives two BEDs using alternative accounting assumptions for employer contributions to health insurance and traces the effects of these assumptions on model results. These exercises all argue for particular incarnations of the BED, rather than proposing a systematic analysis of the effects of uncertainty in the data from which the BED is derived, as is the case here.

billion. What conclusions, then, can the modeller make about the robustness of the welfare results to the value of GNP?

The sensitivity interpretation depends on the probability that the perturbed value of GNP is the true, but unobservable, value of GNP. If the modeller thinks that the true value of GNP lies between $747.0 billion and $747.6 billion with a probability close to one, the probability that the true value of GNP is $740 billion is very small. In probability terms, the simulated perturbation represents a large deviation from the initial value, but only generates a small change in the welfare loss, leading the modeller to conclude that the results are robust to the uncertainty in the value of GNP.

On the other hand, suppose that the modeller thinks that the true value of GNP could lie anywhere between $500 billion and $900 billion with approximately equal probability. In such a case, the perturbation from $747.3 billion to $740 billion represents a very likely change in the initial data value so that the associated welfare change becomes much more significant and confidence in the model results would be considerably lower than in the initial case. Thus, information about the probability of a perturbation is essential for the modeller to interpret the outcome from a sensitivity exercise.[15]

---

[15] 'Limited sensitivity analysis' in which the model is solved for an arbitrary change in the values of key model elasticities, is a prevalent form of elasticity sensitivity analysis in applied general equilibrium modelling. It is used, for example, in Chia, Wahba, and Whalley (1992). In one sense, it suffers from this interpretation problem because no probabilities are explicitly assigned to the elasticity perturbations. On the other hand, these probabilities are implicit: most modellers would agree, for example, that the CES elasticities used in these models are bounded between zero and three, and the choices of perturbations in the limited sensitivity analysis would have reasonably high probabilities of being true. From a systematic or formal perspective, however, the implicit probabilities in limited sensitivity analysis are inadequate.

3.2.1 CPSA Methodology

Since assigning a probability to a given perturbation of the initial data is fundamental for a meaningful sensitivity analysis, the CPSA procedure takes care to identify perturbed data sets in such a way that each is associated with a probability of being the unobservable, true data set.

CPSA is undertaken subsequent to a central case model simulation. Before the sensitivity analysis, the modeller has already adjusted an initial data set and used it to calibrate and solve the model. The initial, unadjusted collection of data is termed the raw data set, the unperturbed data set, or the central case data set. It is comprised of many data elements, and the value of each element in this central data set is referred to as the central case, initial, or unperturbed value.

CPSA uses information about the reliability of the individual initial data values to construct several possible perturbations for each individual data element, and to attach a probability of being true to each possible perturbation. A single perturbed data set can then be constructed using a single value for each data element, where this value may be either the central case value of the element or one of its possible perturbed values. The probability that the perturbed data set is true is derived from the individual probabilities of its constituent data points.

The collection of all possible perturbed data sets, arising from using all combinations of possible (perturbed and central case) values for the individual data elements is also specified in CPSA. This collection is an exhaustive representation of the uncertainty in the data: the sum over all the data sets in the collection, of the

probability associated with each data set, is one. A true picture of the effect of uncertainty in the data on the model results could be gathered from adjusting each of the perturbed data sets in the collection, and then using each resulting balanced data set to calibrate and solve the model, but the number of data sets in the collection is too large for this approach to be practical. Instead, the final step of CPSA is to sample from the collection of perturbed data sets and to use the sample to infer the sensitivity of the model results to the uncertainty in the data.

The CPSA procedure is comprised of the four following steps, which are formally presented in Section 3.2.2.

Step 1. Specification of a priori distributions for the data elements

In the first step of CPSA, the modeller subjectively specifies the probability distribution for each of the initial data elements. Assuming that the modeller has used the best available estimate for the central modelling exercise, the distributions are defined so that the expected values of the data are the same as their values in the central case raw data set. The subsequent moments of these distributions will be largely based on subjective beliefs about the reliability of the data, although some data publications include estimates of dispersion which can be incorporated into the specification.[16] For the sake of simplicity in exposition, the individual data elements are assumed to be

[16] For example, Crossman (1988) categorizes elements of the Australian National Accounts as being of poor reliability with error margins greater than 10 percent, medium reliability with error margins of 3-10 percent, and good reliability with error margins of 0-3 percent.

independently distributed, but CPSA is a sufficiently general procedure that it can include jointly distributed data.

One restriction for CPSA is that the model must be solvable over the supports of the distributions - the values which the data variables are allowed to assume. For example, if the model structure does not admit a negative value for the endowment of labour, then the distribution for the labour endowment cannot include a non-zero probability of being negative.

Given the nature of the data used in applied general equilibrium models, these distributions will typically be bounded: the value of transactions within an economy is usually bounded from below by zero and from above by some finite value such as the value of GNP.[17] A bounded specification of the distribution is not, however, strictly necessary for CPSA.

Step 2. Discrete specification of the distributions

The second step in CPSA is to define possible perturbed values of each data element and to associate each of these perturbed values with a non-zero probability of being true. These perturbed values must be defined in such a way that they capture all of the uncertainty in the data element; the sum of the probabilities associated with each value

[17] One technical restriction on the specification of the continuous distributions for CPSA is that they have finite moments, and bounded distributions fulfil this requirement. Clearly, if (for example) the variance is infinite, assigning a meaningful probability to a given perturbation from the central value of a variable is impossible. Again, the nature of the data used in applied general equilibrium models makes this restriction strictly technical.

must be one. If the specified distribution for a data element is discrete, this second step is unnecessary since a discrete distribution already meets this criterion.

The more likely case, however, is that the modeller specifies a continuous probability density function for one or more of the elements in the initial data set. This specification creates a problem for sensitivity analysis. Applied general equilibrium models require discrete data for their calibration, but the probability that any single value drawn from the support of a continuous distribution and used in the model is true is zero. Since sensitivity analysis requires that the data perturbations have non-zero probabilities, values drawn from the supports of continuous distributions cannot be used for CPSA. To circumvent this difficulty, CPSA uses discrete approximations to the continuous distributions.

Finding discrete approximations to continuous probability density functions is a problem that also arises in the elasticity sensitivity analysis procedures of Harrison and Vinod (1992) and Harrison, Jones, Kimbell and Wigle (1992). The elasticities for which they develop their sensitivity analysis procedures are also assumed to be continuously distributed. Their approach to finding discrete approximations is first to decide on the number of points which will comprise the discrete approximation for a particular elasticity. Suppose that this number is three. They divide the support of the distribution for that elasticity into equi-probable intervals corresponding to the number of points they have specified for the approximation, so that in this example, each interval would contain a probability of one-third. Within each interval, they then find the mid-probability point and use this point to represent the interval in the approximate distribution. For a three-point discrete approximation to the continuous function, then,

the first point would be the point to the left of which one-sixth of the distribution's probability lies, the second point would be the central case value, and the final point would be the point to the right of which one-sixth of the distribution's probability lies. In this discrete approximation, each of the three points would be associated with a probability of being true of one-third.

Miller and Rice (1983), however, show that the higher order moments of a discrete approximation to a continuous function derived in this way do not match the higher order moments of the original distribution, and thus this methodology introduces an identifiable source of error. Instead, they propose a Gaussian quadrature approximation procedure which avoids this error. In Gaussian quadrature, the discrete approximation is specified so that its moments mimic those of the original continuous distribution. The innovation of DeVuyst and Preckel (1997) is to introduce Gaussian quadrature into elasticity sensitivity analysis for applied general equilibrium models. Given its theoretical superiority to the Harrison and Vinod procedure, Gaussian quadrature is also adopted in CPSA to find the discrete approximations to the continuous distributions of the data elements.

By the end of the second step of CPSA, the modeller should be able to identify a series of possible perturbed values and probabilities for each element in the initial data set. As an example, suppose that the central case raw data is given by the vector

$$\bar{\mathbf{A}} = [1 \;\; 2],$$

with elements $\bar{\mathbf{a}}_1$ and $\bar{\mathbf{a}}_2$. Suppose that in Step 1 of CPSA, element $\bar{\mathbf{a}}_1$ is specified as being distributed $N(1, 0.02)$, and element $\bar{\mathbf{a}}_2$ is specified as being distributed $N(2, 0.04)$. A three point Gaussian quadrature would approximate the continuous distributions for $\bar{\mathbf{a}}_1$ by the three discrete point and probability pairs

$$(\hat{a}_1^1 = 0.755, \qquad p_1^1 = 0.1667)$$
$$(\hat{a}_1^2 = 1.000, \qquad p_1^2 = 0.6666)$$
$$(\hat{a}_1^3 = 1.245, \qquad p_1^3 = 0.1667)$$

and $\bar{\mathbf{a}}_2$ by

$$(\hat{a}_2^1 = 1.654 \qquad p_2^1 = 0.1667)$$
$$(\hat{a}_2^2 = 2.000 \qquad p_2^2 = 0.6666)$$
$$(\hat{a}_2^3 = 2.346 \qquad p_2^3 = 0.1667).$$

Step 3. Construction of a joint distribution for the data elements

Once discrete representations of the distributions for the individual data elements have been constructed (or specified, in the case of discrete distributions), the third step in CPSA is to use them to specify the collection of all possible perturbed data sets and to assign a probability of being true to each data set in the collection, that is, to specify the joint distribution for the data elements. This joint distribution is constructed by using all combinations of possible values for the individual data elements. The probability that each perturbed data set in the collection is true is given by the product of the probabilities of its constituent points. Because these perturbed data sets represent all the possible unbalanced data sets, their cumulative probability of being true is one.

The joint distribution for the above example would then consist of the nine vector and probability pairs:

$([\hat{a}_1^1, \hat{a}_2^1] = [0.755, 1.645], \quad p_1^1 p_2^1 = 0.0278)$
$([\hat{a}_1^1, \hat{a}_2^2] = [0.755, 2.000], \quad p_1^1 p_2^2 = 0.1111)$
$([\hat{a}_1^1, \hat{a}_2^3] = [0.755, 2.346], \quad p_1^1 p_2^3 = 0.0278)$
$([\hat{a}_1^2, \hat{a}_2^1] = [1.000, 1.645], \quad p_1^2 p_2^1 = 0.1111)$
$([\hat{a}_1^2, \hat{a}_2^2] = [1.000, 2.000], \quad p_1^2 p_2^2 = 0.4444)$
$([\hat{a}_1^2, \hat{a}_2^3] = [1.000, 2.346], \quad p_1^2 p_2^3 = 0.1111)$
$([\hat{a}_1^3, \hat{a}_2^1] = [1.245, 1.645], \quad p_1^3 p_2^1 = 0.0278)$
$([\hat{a}_1^3, \hat{a}_2^2] = [1.245, 2.000], \quad p_1^3 p_2^2 = 0.1111)$
$([\hat{a}_1^3, \hat{a}_2^3] = [1.245, 2.346], \quad p_1^3 p_2^3 = 0.0278).$

Step 4. Sampling

Typically, however, the number of data sets in the joint distribution is too large for the modeller to use each to solve the model.[18] The final step in CPSA is to sample from the joint distribution. If the number of collections of data in the joint distribution is sufficiently small that each can be labelled, the modeller labels each perturbed data set in the joint distribution and then selects a simple random sample directly. If, as is likely, the number of perturbed data sets is too large to label each, a randomized factorial sampling procedure can be employed. In this sampling procedure, the modeller constructs a single, random, perturbed data set by sequentially choosing a random value for each individual data point from its individual (actual or approximate) discrete distribution.

In the above example, the modeller would construct a random perturbed data set by first choosing randomly from the three possible values for the first data element: $\hat{a}_1^1 = 0.755$, $\hat{a}_1^2 = 1.000$, and $\hat{a}_1^3 = 1.245$, and then choosing randomly from the three possible values for the second data element: $\hat{a}_2^1 = 1.654$, $\hat{a}_2^2 = 2.000$, and $\hat{a}_2^3 = 2.346$.

---

[18] For example, the BED for the Cote d'Ivoire model of Chia, Wahba and Whalley (1992) has 700 data elements. The support of the joint distribution formed from a 2-point Gaussian quadrature discrete approximation to their distributions would have 2^{700} points!

A sample of perturbed data sets is generated by repeating this process. The possibility exists that the same perturbed data set might be chosen twice, so that this process is one of sampling with replacement. Once a sample of perturbed data sets has been generated, each data set in the sample is placed into a matrix format and adjusted into a BED using the same algorithm as was used to adjust the initial, unperturbed collection of data. The ensuing BEDs are employed to calibrate and solve the model, and each model result in the sample is weighted by the probability that the data used in its derivation are true. Finally, the sample is used to find expected values, standard deviations and confidence intervals for the model results.

3.2.2 A Formal Presentation of CPSA

Let the vector **X**, with elements x_j, $j = 1,..., N$, be the vector of data elements required to calibrate an applied general equilibrium model, so that this vector includes all the data variables for which the modeller must specify values. Let the vector $\bar{\mathbf{A}}$ with elements $\bar{a}_j$ be the best initial estimate of **X**. CPSA is a procedure in which the x_j are viewed as random variables, and the $\bar{a}_j$ as realizations from the probability density functions of the x_j. The CPSA methodology is comprised of the following four steps.

Step 1. Specification of the a priori distributions for the x_j

The modeller specifies an *a priori* distribution for each x_j, denoted here by $\{x_j\}$, where $\{x_j\}$ is the probability density function if x_j is a continuous random variable, and the probability mass function if x_j is a discrete random variable. For the purposes of

simplicity, the x_j are assumed to be independently distributed. The random variable x_j must have finite moments, and the support of $\{x_j\}$ must be consistent with the model structure. Because the $\bar{a}_j$ are assumed to be the best initial estimates for x_j in the specified distribution, $E(x_j) = \bar{a}_j$. The variance, $E(x_j - \bar{a}_j)^2$, will be informed by the reliability of the data sources as well as the prior modifications undertaken to generate the unadjusted data.

Step 2. Discrete specification of the continuous $\{x_j\}$

In Step 2, a discrete approximation is found to each continuous $\{x_j\}$, where the discrete approximation is comprised of K pairs of points, $\hat{a}_j^k$, $k = 1, ..., K$, and probabilities p_j^k, such that $\Sigma_k p_j^k = 1$. A discrete approximation is obtained using Gaussian quadrature. For each x_j, Gaussian quadrature chooses K pairs $(\hat{a}_j^k, p_j^k)$ such that

$$\sum_{k=1}^{K} p_j^k \, (\hat{a}_j^k)^l = E\,(x_j)^l, \tag{3.1}$$

where $l = 0, 1, ..., 2K\text{-}1$ are the moments of $\{x_j\}$.

The Gaussian quadrature approximation is found as follows (see Miller and Rice (1983); Preckel and DeVuyst (1992)). For each j, the modeller first solves the linear system of K equations, where the mth equation, $m = 1,...,K$, (and dropping the j subscript) is given by

$$\sum_{l=0}^{K-1} c_l \; E\,(x^{l+m-1}) = -\; E\,(x)^{K+m-1}, \tag{3.2}$$

for the coefficients c_l. The solution values for the c_l are then substituted into the polynomial

$$\sum_{l=0}^{K} c_l \, E\,(x)^l = 0, \qquad \textbf{(3.3)}$$

and its roots are found. These roots are the $\hat{a}_j^k$ points for discrete approximation. The final step is to find the probabilities for the $\hat{a}_j^k$. These are given by solving equation 3.1 for the values of the p_j^k.

Step 3. Construction of a joint distribution

The joint distribution for the elements of **X**, denoted here by {**X**}, is derived from probability mass function representations of the elements in **X**. If the *a priori* distributions are discrete, these probability mass functions are simply the $\{x_j\}$, whereas if the *a priori* distributions are continuous, the probability mass functions are given by the Gaussian quadrature discrete approximations to the $\{x_j\}$. Let each x_j have a probability mass function representation with K point and probability pairs. The joint distribution (see Preckel and DeVuyst, 1992) is given by the N^K vector and probability pairs:

$$\{\mathbf{X}\} = \quad ([\hat{a}_1^{k_1}, \hat{a}_2^{k_2}, \ldots, \hat{a}_N^{k_N}], \prod_{j=1}^{N} p_j^{k_j}) \;\forall\; k_1 = 1,\ldots,K,\; k_2 = 1,\ldots,K,\; \ldots,\; k_N = 1,\ldots,K. \qquad \textbf{(3.4)}$$

Where the x_j are discretely distributed, $\{\mathbf{X}\}$ is the true joint distribution. If $\{\mathbf{X}\}$ is formed from Gaussian quadrature approximations to continuous probability density functions for the x_j, the joint distribution also preserves up to and including the $2K$-1 moments of the original, continuous joint distribution. Because this joint distribution is formed under the assumption of stochastic independence of the x_j, the covariances and higher order cross-moments are zero.[19]

Step 4. Sampling

If N^K is sufficiently small, a simple random sample is taken from the points in the support of $\{\mathbf{X}\}$ by labelling each point with an integer in the interval $[1, 2,..., N^K]$ and drawing a simple random sample from that interval. If N^K is large and a simple random sample cannot be easily generated, random sampling from the support of $\{\mathbf{X}\}$ is achieved by the completely randomized factorial sampling design used in Harrison and Vinod (1992): each point in the sample is generated by randomly selecting its elements from the supports of the discrete representations of the $\{x_j\}$, so that a sample data set is generated by randomly selecting from the values $[\hat{a}_j^1, \hat{a}_j^2, ..., \hat{a}_j^K]$ for each $j = 1,...,N$.

Let $\mathbf{S}_i$ be a random vector from the support of $\{\mathbf{X}\}$, and let P_i be the probability mass of $\mathbf{S}_i$. The modeller applies the same adjustment algorithm to $\mathbf{S}_i$ as was applied to

[19] The assumption of stochastic independence for such data is supported in applications of the Stone-Byron adjustment algorithm in the social accounting literature, which requires an *a priori* specification of the variance-covariance matrix for a social accounting matrix. For an example, see Crossman (1988). CPSA can, in principle, be extended to the case where elements of **X** are jointly distributed. Preckel and DeVuyst (1992) give a Gaussian quadrature joint distribution for the case in which the x_j are joint normally distributed.

$\bar{\mathbf{A}}$, to generate a BED. The BED is then used to calibrate and solve the model. Let $\mathbf{R}_t$ denote the vector of model results arising from the unbalanced data vector $\mathbf{S}_t$.

The process is repeated T times to generate a sample of model results. T is chosen to be sufficiently large that the sample moments are consistent estimators of the population moments. To ensure that all vectors in the support of $\{\mathbf{X}\}$ have the same probability of being sampled, sampling is undertaken with replacement allowing the possibility that the same vector may be drawn more than once. Each $\mathbf{R}_t$, $t = 1,...,T$, is weighted by P_t to find the expectations, standard deviations, and confidence intervals for the model results.

3.2.3 An Illustration of CPSA Using a Simple Tax Model

The Shoven and Whalley (1984) simple 2x2x2 model, with two consumers (rich and poor), two factors of production (capital and labour), and two commodities (manufactured and non-manufactured goods), is used to illustrate CPSA. Table 3.1 summarizes the model structure. The base case version of the model has no taxes. In the counterfactual experiment, a 50 percent tax is levied on the use of capital in the manufacturing sector, resulting in welfare changes for both consumers. These welfare changes, measured by the Hicksian equivalent variation as a proportion of base income, provide the basis for the sensitivity analysis.

The initial, unbalanced data set used for this model is given in Table 3.2. It is derived by choosing a random value from a uniform distribution in which the expected value of each data point is the value used in Shoven and Whalley (1984), and the

Table 3.1

Structure of the Tax Model Used to Illustrate CPSA

Production

- Output is produced using capital and labour combined in proportions implied by CES technology in each sector.
- The elasticity of substitution in the production of manufactured goods is 2.0 and in that of non-manufactured goods, 0.5.
- Share parameters for the CES function are calibrated from the BED.

Consumption

- The utility of each consumer is a CES function of manufactured and non-manufactured goods.
- The rich consumer's utility function has an elasticity of substitution of 1.5 and the poor consumer's has one of 0.75.
- Share parameters for the CES function are calibrated from the BED.

Endowments

- The rich consumer is endowed with capital and the poor consumer with labour.

Equilibrium Conditions

- Markets clear for all goods and factors.
- Zero profits are made in each sector.
- Each consumer's expenditures equals his/her income.

Counterfactual

- A 50 percent tax is levied on the use of capital in the production of manufactured goods.
- The rich consumer receives 40 percent of tax revenues and the poor consumer receives 60 percent.
- Welfare changes for each consumer are measured by equivalent variation as a proportion of base income: $EV^i = (U^i_c - U^i_b) / U^i_b$ where U^i_b is the utility of consumer i , i = {rich, poor}, in the base case and U^i_c is utility after the imposition of the tax.

Table 3.2

Unbalanced Transactions Values For the Illustrative Tax Model[1]

(in units of currency)

1. Consumption by Households

	Goods	
	Manufactures	Non-Manufactures
Rich	17.2	25.8
Poor	22.0	52.7

2. Factor Demands by Sector

	Sectors	
	Manufactures	Non-Manufactures
Capital	7.1	30.4
Labour	34.0	56.6

3. Factor Endowments

	Factors	
	Capital	Labour
Rich	48.3	0
Poor	0	59.0

note 1: Values were derived as random numbers drawn from uniform distributions with means equal to the Shoven and Whalley (1984) balanced values and standard deviations equal to 10 percent of those balanced values.

Known Totals (values used in the Shoven and Whalley (1984) model)

Rich Household's Endowment of Capital	34.3
Poor Household's Endowment of Labour	60.0
Total Demand for Capital	34.3
Total Demand for Labour	60.0
Total Output of Manufactured Goods	34.9
Total Output of Non-Manufactured Goods	59.4

---

standard deviation is ten percent of the Shoven and Whalley value.[20] The adjustment algorithm used to balance the data is the commonly used constrained quadratic minimization algorithm in which each term is weighted by the unadjusted data value. Thus, if $\bar{a}_j$, j=1,..,10, denotes each of the ten unbalanced, non-zero data values in Table 3.2, this algorithm finds balanced values, $\bar{q}_j$, such that the expression $\sum_j ((\bar{q}_j-\bar{a}_j)/\bar{a}_j)^2$ is minimized subject to the constraints of the specific experiment.

Two sets of experiments, each of which imposes different constraints on the adjustment algorithm, are performed to illustrate the CPSA procedure. The first set assumes that the modeller knows with certainty the aggregate incomes, demands, and outputs in the economy, and that they are the values used by Shoven and Whalley. These 'known' control totals are given in the final section of Table 3.2. In this case, the constraints on the adjusted data are that the adjusted endowments equal the known endowments, that the sum of demands for each good equals the known value for the output of each sector, and that the sum of the input demands for each factor equals the value of the known total endowment for each factor.

The second set of experiments assumes that these totals are unknown. In this case, the adjustment constraints are simply that the data meet the equilibrium conditions of the model: markets clear, sectors make zero profits, and the households exhibit

[20] The choice of a uniform distribution is arbitrary, but the value for the standard deviation is roughly consistent with data. A time series of annual values for value added in manufacturing for the United States was found to have a standard deviation of 10.2 percent. The time series was constructed using annual data for 1970 to 1992 taken from the International Bank for Reconstruction and Development (1993) data base. The series "value added in manufacturing" given in current USD was deflated by the ratio of current USD to constant 1985 USD GDP at factor cost to generate a constant value series.

budget balance. Both sets of balanced data are given in Table 3.3, together with the central case welfare results. The robustness of these equivalent variations to uncertainty in the initial data is the focus of the CPSA exercise.

In the first step of the CPSA, uniform distributions are specified for the initial data elements, where the expected value of each data element is given by the central unadjusted data value in Table 3.2 and its standard deviation is ten percent of that value.[21] The second step of CPSA uses Gaussian quadrature to find discrete approximations to these continuous distributions. In this example, they are represented by the three-point approximations given in Table 3.4, which preserve up to and including the fifth moments of the original distributions.

The discrete approximations to the distributions for the individual data elements are then used to characterize the joint probability density function. The support of this joint distribution is given by the combinations arising when each of the ten data elements assumes one of the three values in the support of its discrete distribution. The result is a set of 10^3 possible configurations, each of which has a probability of being true given by the product of the probabilities of its ten constituent points.

A random, unadjusted data configuration is drawn from this support. This data configuration is derived by choosing randomly from the three point discrete distribution for each data element. For example, the three points in the distribution for the rich

[21] Although the data generating process for the economy in this example has been specified to generate the unbalanced data set in Table 3.2, it would be unknown for a modeller undertaking CPSA. Here, the modeller has correctly specified the shape of the distribution, and the proportional magnitudes of the variances, but has the expectation that the error of the initial estimate is zero, which of course, it is not.

Table 3.3

Balanced Data and Central Case Model Results For the Illustrative Tax Model

Case 1: Benchmark Values for Data Balanced Using Known Totals

	value in units of currency
Rich Consumption of Manufactured Goods	15.1
Rich Consumption of Non-Manufactured Goods	19.2
Poor Consumption of Manufactured Goods	19.8
Poor Consumption of Non-Manufactured Goods	40.2
Input of Capital to the Manufacturing Sector	7.7
Input of Labour to the Manufacturing Sector	27.2
Input of Capital to the Non-Manufacturing Sector	26.6
Input of Labour to the Non-Manufacturing Sector	32.8
Capital Endowment of the Rich Consumer	34.3
Labour Endowment of the Poor Consumer	60.0

Hicksian Equivalent Variation from a 50 % tax on Capital in Manufacturing

	proportion of base income
Rich Consumer's EV	-0.1223
Poor Consumer's EV	0.0610

Case 2: Benchmark Values for Data Balanced Using Equilibrium Constraints

	value in units of currency
Rich Consumption of Manufactured Goods	16.3
Rich Consumption of Non-Manufactured Goods	26.0
Poor Consumption of Manufactured Goods	20.5
Poor Consumption of Non-Manufactured Goods	52.1
Input of Capital to the Manufacturing Sector	8.3
Input of Labour to the Manufacturing Sector	28.5
Input of Capital to the Non-Manufacturing Sector	34.0
Input of Labour to the Non-Manufacturing Sector	44.1
Capital Endowment of the Rich Consumer	42.3
Labour Endowment of the Poor Consumer	72.6

Hicksian Equivalent Variation from a 50% tax on Capital in Manufacturing

	proportion of base income
Rich Consumer's EV	-0.1126
Poor Consumer's EV	0.0572

Table 3.4

3 Point Discrete Gaussian Quadrature Approximations for the Assumed Continuous Data Distributions in the Illustrative Model

values in units of currency

	Expected Value and Bounds of the Assumed Uniform Distribution	1st Point (probability: 0.278)	2nd Point (probability: 0.444)	3rd Point (probability: 0.278)
Rich Consumption of Manufactured Goods	17.2 [14.2, 20.1]	14.8	17.2	19.5
Rich Consumption of Non-Manufactured Goods	25.8 [21.3, 30.3]	22.3	25.8	29.3
Poor Consumption of Manufactured Goods	22.0 [18.2, 25.8]	19.1	22.0	25.0
Poor Consumption of Non-Manufactured Goods	52.7 [43.6, 61.9]	45.7	52.7	59.8
Input of Capital to Manufacturing Sector	7.1 [5.8, 8.3]	6.1	7.1	8.0
Input of Labour to Manufacturing Sector	34.0 [28.1, 39.9]	29.5	34.0	38.6
Input of Capital to Non-Manufacturing Sector	30.4 [25.2, 35.7]	26.3	30.4	34.5
Input of Labour to Non-Manufacturing Sector	56.6 [46.8, 66.4]	49.0	56.6	64.2
Capital Endowment of Rich Consumer	48.3 [39.9, 56.6]	41.8	48.3	54.7
Labour Endowment of Poor Consumer	59.0 [48.8, 69.2]	51.1	59.0	66.9

consumer's endowment of capital are 41.8, 48.3 and 54.7. The data configuration is constructed by randomly choosing one of these three values, then randomly choosing one of the three possible values for the poor consumer's endowment of labour, and so on until a random value has been chosen for each of the ten data elements. The probability associated with the configuration is given by the product of the probabilities of its ten constituent data elements.

This process is repeated to generate a sample of fifty unadjusted data configurations. In the first experiment, each configuration is then adjusted into a BED using the known totals as constraints, and the model is calibrated and solved. Attached to each result, is the probability that the configuration used in its derivation is true. The process is the same in the second experiment, except the data are adjusted using only the model's equilibrium conditions as the balancing constraints.

Table 3.5 presents the summary statistics that characterize the output of the CPSA sensitivity procedure. It gives the means, standard deviations and 95 percent confidence intervals for the results of both experiments undertaken in this illustrative example. From Table 3.5, the modeller could conclude that the model results are robust to the uncertainty in the initial data values: the signs of the welfare changes are preserved, and the central case variants lie well inside the 95 percent confidence interval. Where the control totals are known, the standard deviations of the results are lower than where the model's equilibrium conditions alone provide the underlying adjustment consistency constraints. This result is consistent with the additional information introduced into the system by known totals.

Table 3.5

CPSA on the Welfare Effects of Imposing a 50 Percent Tax on the Use of Capital in the Manufacturing Sector

Hicksian Equivalent Variations measured as a proportion of base income

Case 1: Data Balanced Using Known Control Totals

	Central Case[1]	Mean	Standard Deviation	95% Confidence Interval[2]
EV Rich	-0.1223	-0.1219	0.0095	[-0.1644, -0.0794]
EV Poor	0.0610	0.0609	0.0050	[0.0385, 0.0833]

Case 2: Data Balanced Using Equilibrium Constraints

	Central Case[1]	Mean	Standard Deviation	95% Confidence Interval[2]
EV Rich	-0.1126	-0.1117	0.0097	[-0.1551, -0.0683]
EV Poor	0.0572	0.0572	0.0055	[0.0326, 0.0818]

note 1: The central case uses the raw data given in Table 3.2.
note 2: Confidence intervals are derived using Chebychev's Theorem.

These sensitivity results are, of course, dependent on the *a priori* specification for the distributions of the data elements. If the modeller's *a priori* information about the data had led to a specification in which the standard deviations of the uniform distribution for the data were 300 percent of their base values, the modeller could no longer feel confident about the signs of welfare effects; the 95 percent confidence intervals presented in Table 3.6 include positive and negative equivalent variations for both consumers.

3.2.4 The Mechanism of CPSA

CPSA is based on changing the values of the initial data, and the mechanism by which those changes affect model results is specific to the model's structure. One of the reasons for choosing a very simple model to illustrate CPSA is that this mechanism is transparent. In the illustrative example, the introduction of a tax on the use of capital in the manufacturing sector causes the price of manufactured goods to rise relative to that of non-manufactured goods, and subsequently causes a net decrease in the demand for manufactured goods. The shift in production towards the more labour-intensive, non-manufactured good forces up the price of labour relative to capital. The rich consumer, who is endowed only with capital and who receives 40 percent of tax revenues, experiences a loss in income, and hence, a loss in utility. This welfare loss is compounded by the increase in the price of manufactured goods which figure more prominently in the rich consumer's utility function than in the poor consumer's. In the sensitivity analysis, a higher share of manufactured goods in the constant elasticity of

Table 3.6

Sensitivity of CPSA to the Data's Probability Density Functions

Standard Deviation of Distributions are 300 Percent of the Raw Data Value

Hicksian Equivalent Variations measured as a proportion of base income

Data Balanced Using Known Totals:
Uniform Distributions with Standard Deviations of 300 percent of the base value

	Central Case[1]	Mean	Standard Deviation	95% Confidence Interval[1]
EV Rich	-0.1223	-0.1099	0.0565	[-0.3626, 0.1428]
EV Poor	0.0610	0.0657	0.0350	[-0.0908, 0.2222]

note 1: The central case uses the raw data given in Table 3.2.
note 2: Confidence intervals are derived using Chebychev's Theorem.

substitution (CES) utility function of the rich consumer and/or a higher share of capital in the production function for manufactured goods should, therefore, result in a greater decrease in the welfare of the rich consumer as the tax on the use of capital in the manufacturing sector is imposed.

This expectation is supported by evidence from a simple experiment. If all but one of the data estimates are held constant at their central case value and the poor consumer's consumption of non-manufactured goods, v_{pn}, is allowed to vary, the way in which the changes in one data value lead to variations in the model welfare effects can be traced. This link is shown in Table 3.7 for three values of v_{pn}. The first section in Table 3.7 demonstrates the link between the initial value of v_{pn} and final benchmark consumption values. The lowest value for v_{pn} is 50 percent of the true value of the poor consumer's consumption of non-manufactured goods (used in Shoven and Whalley, 1984), the highest is twice the true Shoven and Whalley value and the middle value is the true value. In each case, the other nine data values are those given in Table 3.2. The same constrained quadratic minimization algorithm is used to adjust the data as in the central case, using the known totals of Table 3.2 as constraints. These changes in the adjusted value for the poor consumer's consumption of non-manufactured goods, together with changes in the remainder of the adjusted elements in the BED, have consequences for the values of the calibrated parameters. As v_{pn} increases, the changes in the BED imply that the calibrated share parameter for manufactured goods in the rich consumer's CES utility function increases, while that in the poor consumer's utility function decreases. The greater weight on the rich consumer's share parameter for manufactured goods yields the higher equilibrium consumption values given in the third

Table 3.7

The Link Between Changes in the Raw Data Estimate of the Poor Consumer's Consumption of Non-Manufactured Goods (v_{pn}) and Changes in Model Results[1]

1. BED Consumption Levels[2]
(values are in units of currency)

v_{pn}	Rich Consumption Manufactures	Rich Consumption Non-Manufactures	Poor Consumption Manufactures	Poor Consumption Non-Manufactures
20.6	8.88	25.42	26.02	33.98
41.2	13.82	20.48	21.08	38.92
82.4	16.88	17.42	18.02	41.98

↓

2. Utility Function CES Share Parameters Implied by Alternative BEDs[3]

v_{pn}	Manufactures' Share in Rich Utility	Non-Manufactures' Share in Rich Utility	Manufactures' Share in Poor Utility	Non-Manufactures' Share in Poor Utility
20.6	0.331	0.669	0.412	0.588
41.2	0.435	0.565	0.306	0.694
82.4	0.495	0.505	0.245	0.755

↓

3. Counterfactual Demands

v_{pn}	Rich Consumption Manufactures	Rich Consumption Non-Manufactures	Poor Consumption Manufactures	Poor Consumption Non-Manufactures
20.6	6.80	23.70	26.00	37.44
41.2	10.84	19.42	21.01	42.66
82.4	13.41	16.68	17.95	45.89

↓

4. Equivalent Variation as a Proportion of Base Income

v_{pn}	Rich EV	Poor EV	Rich Base Utility	Poor Base Utility	Rich Counterfactual Utility	Poor Counterfactual Utility
20.6	-0.1127	0.0556	18.65	30.35	16.55	32.04
41.2	-0.1202	0.0598	17.37	31.82	15.28	33.72
82.4	-0.1251	0.0627	17.15	33.36	15.01	35.45

note 1: The poor consumer's consumption of Non-Manufactured Goods is 41.2 in the 'true' case. The values for v_{pn} are chosen here to be 0.5 (low), 1 (intermediate) and 2 (high) times this value.

note 2: The values shown here are those which change as a result of altering only v_{pn} in the raw data. The remaining elements of the BED are the same throughout.

section of Table 3.7. These higher share parameters lead to a greater disutility from the increase in the price of the manufactured good and the consequent greater loss in welfare, as reflected by the values of the rich consumer's equivalent variation in the final section of Table 3.7.

The opposite effect is evident for the poor consumer whose share of manufactured goods in utility decreases as v_{pn} increases, and whose subsequent counterfactual demand for the manufactured good decreases with higher values of v_{pn}. Increases in the value of the initial estimate for the poor consumer's consumption of non-manufactured goods thus lead to higher welfare gains for the poor consumer, from the imposition of the tax.

3.3 Extended Sensitivity Analysis

While CPSA allows modellers to undertake sensitivity analysis with respect to the values of the hitherto neglected calibrated parameters, the elasticity parameters remain a highly uncertain component of the modelling process. The 'extended sensitivity analysis' proposed in this section combines CPSA with the existing elasticity sensitivity analysis methodology advocated in DeVuyst and Preckel (1997), so that modellers can report the sensitivity of their model results to uncertainty in all of the model's parameters.

3.3.1 Extended Sensitivity Analysis Methodology

The extended sensitivity analysis methodology requires a simple modification to the CPSA procedure described in Section 3.2. Instead of specifying *a priori* distributions just for the initial data elements, distributions are specified for both the initial data and for the exogenous elasticity parameters. Thus, if N is the number of data elements, and J is the number of exogenously specified parameters, the modeller must specify $(N+J)$ probability distributions. The conditions that apply to the distributions for the data in CPSA also apply to the elasticities in extended sensitivity analysis: they must have finite moments, and the model must be solvable over their supports.

The remaining steps in extended sensitivity analysis follow those in CPSA, except they apply to both the initial data and to the elasticities. Gaussian quadrature is used to construct discrete approximations to the continuous distributions of both the data and the elasticities, and a joint distribution of the data and elasticities is created from these discrete approximations. If K is the number of points in the support of each discrete distribution, the joint probability density function contains $(N+J)^K$ points. Each point in the support of the joint distribution is comprised of an unadjusted data set and a set of elasticity values. Its probability mass is given by multiplying the product of the N probabilities of its unadjusted data values with the product of its J elasticity probabilities.

Random samples, each of which is comprised of an unbalanced data and a set of elasticity values, are then drawn from the joint distribution. As in CPSA, the data component of each random sample is constructed by sequentially choosing a random

value for each data element from the K values in the support of its individual discrete distribution. Similarly, the elasticity component of the random sample is derived by sequentially selecting a random value for each elasticity from the K values in the support of its distribution.

The data in each sample are balanced by applying the same adjustment algorithm as was applied to the central case data. Together with the elasticities in the sample, the balanced data are then used to calibrate and solve the model. Means, standard deviations, and confidence intervals for the true model results are calculated from the probability weighted sample model results, as in CPSA.

This extended sensitivity analysis methodology is illustrated using an existing model developed by Chia, Wahba, and Whalley (1992) for tax incidence analysis in Côte d'Ivoire. While the simple Shoven and Whalley example was chosen to illustrate CPSA in Section 3.2 on the basis that its small dimensionality offers transparency, the Côte d'Ivoire model is used to illustrate extended sensitivity analysis because it typifies the policy modelling exercises for which such sensitivity analyses are important.

The attempt here at realism is hampered by a lack of knowledge about the unadjusted data, and the reliability of the elasticities actually used in the Côte d'Ivoire model. The lack of such information means that several assumptions about the data, elasticities and data adjustments are made in the illustration of extended sensitivity analysis that follows.

This obstacle, however, is not unique to the Côte d'Ivoire model. In practically all cases, the information required to undertake extended sensitivity analysis is not available to anybody other than the original modeller, and usually this information has

been discarded early in the modelling process. Typically, modellers have no use for unadjusted data; they report only the adjusted version of the data and the central case elasticities. Hence, extended sensitivity analysis has normative implications for modellers. They must maintain a version of the unadjusted data, record their assessment of the reliability of both the unadjusted data and of the elasticities, and report their adjustment procedure in detail if extended sensitivity analysis is ever to be undertaken.

3.3.2 Extended Sensitivity Analysis for the Côte d'Ivoire Model

The incidence analysis of Chia, Wahba, and Whalley is undertaken for six taxes/subsidies by replacing each with an equal yield, neutral tax on consumption, and finding the associated welfare change for each of seven household types.[22] The exercise that follows examines the sensitivity of the personal income tax incidence results to uncertainty in the consumption expenditure data and in the values of the consumption and production elasticities of substitution.

The welfare changes on which the Chia, Wahba and Whalley tax incidence results are based, derive from household utility functions that are defined over the consumption of goods and services in the model.[23] The data-based component of the

---

[22] The original model is calibrated to a 1986 BED and solved using MPS/GE (Rutherford, 1989), but it has been rewritten in GAMS (see Brooke, Kendrick and Meeraus, 1988) to allow a simple incorporation of data adjustment, and to facilitate the CPSA component of extended sensitivity analysis.

[23] The Côte d'Ivoire model identifies seven socio-economically based household types, each of which receives utility from the consumption of ten goods and services. Incomes derive mainly from capital and labour endowments, as well as interhousehold transfers. Households pay personal income tax and make social security transfers to the government, but also receive income from the government in the form of education and other transfers. The model

extended sensitivity analysis is undertaken for this consumption expenditure data. Changes in utility arise directly from changes in consumption levels, but the extent to which a change in the consumption of a particular good translates into a change in utility is determined by the share parameter of that good in the CES utility function. Through calibration, the values of the consumption expenditure data (together with the elasticity of substitution in consumption) determine the values of these share parameters.

The consumption data are assumed to have been obtained from a household survey that reports mean consumption by household type, and Chia, Wahba, and Whalley are assumed to have derived an unbalanced estimate of total consumption expenditure by each household type from scaling the survey data by the number of households in each group.[24] Because the actual household survey data are unknown, they are approximated by the artificially constructed household survey data given in Table 3.8. The elements of this artificial data set are randomly drawn from a normal

---

distinguishes fifteen productive sectors, each of which produces output using value added and intermediate goods. All twelve formal sectors pay production taxes and all formal sectors, except the government services sector and the gas, electricity and water sector, also receive subsidies. Eight of the formal productive sectors trade internationally, and since Côte d'Ivoire is modelled as a small, open, price-taking economy, exporters face a perfectly elastic demand function for their output. Traditional exports and exports of primary processed goods are taxed. Imports, used in the production of intermediate goods and in household consumption, are subject to tariffs. The Ivorian price stabilization policy for coffee, cocoa and other exports is captured in the model. In 1986, the benchmark year, the fund experienced a net inflow of revenues and thus the traditional export sector pays into the stabilization fund, while the non-traditional export sector receives only a proportion of those revenues.

[24] The number of households by type is as follows; export croppers: 2.436 million; savannah food croppers: 1.320 million; other food croppers: 1.524 million; government employees: 1.416 million; formal sector households: 0.912 million; small businesses: 2.580 million; inactive households: 1.812 million.

Table 3.8

Unbalanced Household Survey Consumption Expenditure Data for the Côte d'Ivoire Model

Annual Consumption Expenditure[1] in CFA Francs

Consumption Good	Household Type						
	Export Croppers	Savannah Croppers	Other Croppers	Gov't Workers	Formal Sector	Small Business	Inactive
Rice	3260	4054	1970	10000	18014	8879	10635
Other Subsistence Agr.	35669	52039	47922	35616	40742	35082	27816
Traded Agr. Products	218	0	307	7815	8790	4265	8796
Primary Processed	35651	36759	41063	56167	105131	58828	51389
Manufactures	20209	13847	12817	40251	64624	29694	21507
Electricity, Gas, Water	1817	2530	1659	5401	5965	1951	1884
Construction	2460	1871	1784	6340	5103	3956	2214
Transport	6171	0	9450	34028	30515	12630	34329
Financial Services	576	353	385	4757	2916	1214	1165
Non-Financial Services	6392	7723	6904	18904	34329	1086	13319

note 1: The absence of the actual unbalanced data used for the Côte d'Ivoire model means that these values have been artificially constructed. Unbalanced data were derived as random numbers drawn from a normal distribution with mean equal to the balanced value in the Chia, Wahba, and Whalley model and standard deviation as the following: Rice, Construction and Financial Services, 10% of the balanced value, Other Subsistence Agricultural Products, Traded Agricultural Goods, Primary Processed Goods, Manufactures, and Electricity, Gas, Water, 20% of the balanced value, Transport and Non-Financial Services, 30 % of the balanced value.

distribution with an expected value equal to the known, adjusted value used by Chia, Wahba, and Whalley, and a standard deviation equal to the proportions of the base case given in note 1 of Table 3.8.[25]

The balanced values of the consumption expenditure data for the Côte d'Ivoire model are assumed to have been derived in a two stage process. In the first stage, aggregate values for the total final household consumption of each good consistent with the values for total production, exports, government consumption and intermediate demand would have been found. These values are assumed to be the totals used by Chia *et al.* and are given in the first section of Table 3.9. Similarly, aggregate household consumption expenditure would have been specified. These values are presented in the second section of Table 3.9, and are also the values used in the original model. In the second stage, an adjustment algorithm would have been applied to the initial, unbalanced consumption data under consistency conditions implied by the aggregate values from the first stage.

The unbalanced data in Table 3.8 are scaled by the number of households of each type (given in footnote 24), and are then adjusted using the prevalent RAS adjustment algorithm, where the consistency constraints are that i) the total consumption of each good by each household type, summed across household types is equal to the aggregate final household consumption for that good from section 1 of

---

[25] Chia, Wahba, and Whalley list their primary data sources as the national accounts, the Banque de données financières (from which balance of payments data was obtained), tax data and household budget survey data, but do not state explicitly which elements of the BED derive from which source. As a result, the sensitivity analysis presented here provides an illustration of the methodology rather than informed insight into the Côte d'Ivoire model results.

Table 3.9

Aggregate Totals Used as Constraints in the RAS Adjustment Algorithm

1. Known Row Totals:

Aggregate Consumption by Product	**(million CFA francs)**
Rice	86484
Non-Rice Subsistence Agricultural Products	516210
Traded Agricultural Products	50164
Goods from Primary Processing	617750
Manufactured Goods	341565
Electricity, Gas, Water	30864
Construction	37600
Transport	201072
Financial Services	17509
Non-Financial Services	153498

2. Known Column Totals:

Aggregate Consumption Expenditure by Household	**(million CFA francs)**
Export Croppers	296186
Savannah Food Croppers	157369
Other Food Croppers	185213
Government Employees	375647
Formal Sector Households	285345
Small Businesses	418530
Inactive	334426

Table 3.9 and ii) the sum across goods of total consumption expenditure by household type is equal to the disposable income of each household type, net of interhousehold transfers and savings, given in section 1 of Table 3.9.[26]

Together with the original elasticity values, the resulting balanced matrix is used to calibrate and solve the model to obtain incidence results for the removal of the personal income tax. These central case results are given in column (1) of Table 3.10. Their robustness to uncertainty in the values of the initial household consumption data in Table 3.8 and to uncertainty in the central case values of selected production and consumption elasticity values is the focus of the subsequent sensitivity analysis.

The illustration of extended sensitivity analysis presented here considers the uncertainty in the calibrated parameters given by the consumption expenditure matrix together with uncertainty in the values of three sets of elasticities used in CES functions in the model; the elasticity of substitution of consumption goods in preferences,[27] the Armington elasticity of substitution between domestic and imported goods in consumption, and the elasticity of substitution between capital and labour in production.

---

[26] Let the unbalanced data be represented in the matrix form of Table 3.8, so that the element $\bar{a}_{ij}$ denotes the expenditure by household j on good i, and let the total consumption of each product be given in the first section of Table 3.9, and the total expenditure by household be given in the second section of Table 3.9. The RAS algorithm, attributed to Bacharach (1970) is a scaling algorithm in which each row of the initial matrix is scaled by the ratio of the known row total (section 1 of Table 3.9) to the actual total. The columns of the ensuing updated matrix are scaled by the ratio of the known column totals (section 2 of Table 3.9) to the updated matrix column totals. This process is applied iteratively until the deviation of the updated matrix totals from the control totals is deemed to be sufficiently close to zero.

[27] In the central case, these are all 1 implying Cobb-Douglas preferences for households. Sensitivity analysis with respect to this value can therefore also be interpreted as sensitivity over the choice of functional form.

Table 3.10

Extended Sensitivity Analysis Results for Personal Income Tax Incidence in a Model of Côte d'Ivoire

Hicksian Equivalent Variations expressed as a percentage of benchmark gross income

Unbalanced Data Adjusted Using the RAS Algorithm

Household Type	Central Case (1)	Expected Value (2)	Standard Deviation (3)	95 % Confidence Interval[1] (4)
Export Croppers	-0.224	-0.214	0.013	[-0.272, -0.156]
Savannah Croppers	-0.685	-0.690	0.019	[-0.775, -0.605]
Other Food Croppers	-1.600	-1.603	0.020	[-1.692, -1.514]
Government Employees	3.493	3.490	0.009	[3.450, 3.530]
Formal Households	-0.605	-0.607	0.007	[-0.638,-0.576]
Small Businesses	-1.617	-1.614	0.006	[-1.641, -1.587]
Inactive	2.666	2.661	0.015	[2.594, 2.728]

note 1: Confidence intervals are derived using Chebychev's Theorem.

The elements of the consumption data matrix are assumed to be uniformly distributed where the standard deviation differs by good: rice, construction, and financial services are assumed to be the most reliably reported goods with standard deviations of 10 percent of their base value; transportation and non-financial services the least reliably reported with standard deviations of 30 percent of their base value; and the data on the remaining goods are assumed to be of intermediate reliability with standard deviations of 20 percent of their base values. Thus, the distribution for the data element, export croppers' rice consumption, has a lower bound of 2696 CFA francs, an expected value of 3260 CFA francs, and an upper bound of 3825 CFA francs.

Likewise, the elasticities are also assumed to be uniformly distributed. The bounds for the distributions of the production elasticities are assumed to be the central values +/-0.35, while those of other elasticities are assumed to be +/- 40 percent of their initial values. The central case values and bounds of these elasticities are given in Table 3.11.

The uniform distributions for the data and the elasticities are then approximated with three-point discrete approximations obtained from Gaussian quadrature. The support of each approximate distribution has a low, middle and high value. The low value in each approximation derived through Gaussian quadrature is given by the lower bound of the distribution plus 11.27 percent of the range and is associated with a probability of 0.28. The middle value is the lower bound plus 50 percent of the range (the central case value) with a probability of 0.44, and the high value, the upper bound minus 11.27 percent of the range, is associated with a probability of 0.28.

Table 3.11

Elasticities of Substitution and Bounds Used in the Extended Sensitivity Analysis

1. Elasticity of Substitution Between Capital and Labour in Production Sectors (bounds are central case value +/- 0.35)

	Central Value	Lower Bound	Upper Bound
Food	0.4	0.05	0.75
Traditional Exports	0.4	0.05	0.75
Non-Traditional Exports	0.5	0.15	0.85
Formal Sector Primary Processing	0.8	0.45	1.05
Formal Sector Manufacturing	0.8	0.45	1.05
Gas and Electricity	0.8	0.45	1.05
Transportation	0.5	0.15	0.85
Formal Sector Services	0.8	0.45	1.15
Financial Services	0.8	0.45	1.15
Informal Sector Services	0.9	0.55	1.25
Informal Sector Primary Processing	0.9	0.55	1.25
Informal Sector Manufacturing	0.9	0.55	1.25
Informal Sector Construction	0.4	0.05	0.75
Formal Sector Construction	0.4	0.05	0.75

2. Elasticity of Substitution Between Goods in Utility (bounds are central case +/- 40 percent)

	Central Value	Lower Bound	Upper Bound
All Households	1	0.6	1.4

3. Elasticity of Substitution Between Imports and Domestic Goods in Consumption (bounds are central case +/- 40 percent)

	Central Value	Lower Bound	Upper Bound
All Goods	2	1.6	2.4

With 31 elasticities and 70 data elements in the consumption expenditure matrix, the support of the discrete joint probability distribution approximation has 3^{101} points. The probability associated with any one of those points is given by the product of the probabilities of its data and elasticity components. A random sample of 500 points is drawn from the joint probability distribution on the assumption that this number is sufficiently large that the sample mean and standard deviation can be used to derive confidence intervals for the model's welfare results.

The sensitivity results are reported in columns (2), (3), and (4) of Table 3.10. The confidence intervals in Table 3.10 suggest that if the specified distributions for the data and the elasticities are true, the model results are robust to uncertainty in the parameters, in the sense that the signs of the welfare effects do not change. Furthermore, at the 95 percent confidence interval, the ranking of the incidence of the Ivorian personal income tax among household groups remains the same as in the central case: government employees bear most of the burden of the tax with inactive households assuming a secondary burden. Thus, if the many assumptions made about the source and nature of uncertainty in the data for the Côte d'Ivoire model hold, the central case model results could be confidently presented as inputs into a debate on tax policy reform.

3.4 Conclusion

Among the criticisms levelled against applied general equilibrium models is one of empirical weakness - model parameterization relies on point observations which lack

the statistical rigour of time series data. One means of addressing this criticism is for modellers to incorporate whatever information they do have about the quality of those single observations in the modelling process, via sensitivity analyses. This chapter has developed a sensitivity analysis procedure, CPSA, that produces measures for the robustness of the model results to uncertainty in the raw data used in calibration, and hence, to uncertainty in the values of the calibrated parameters. When combined with an existing sensitivity analysis methodology for elasticities to form the 'extended sensitivity analysis' procedure, CPSA allows the modeller to generate confidence intervals for the model results based on the uncertainty in the model's full numerical specification.

Both extended sensitivity analysis and CPSA require the modeller to incorporate information that is used in the modelling exercise, but which is often discarded early on: the values of the unadjusted data, assessments of the reliability of those data, and a record of the data adjustment process. As the procedures proposed here illustrate, retaining and exploiting this information allows modellers to enrich their numerical analyses.

A central feature of the methodologies described in this chapter is that the perturbed data are always adjusted using the same adjustment algorithm that was used to balance the central case raw data. The two adjustment procedures used in the examples, the weighted quadratic minimization algorithm in Section 3.2 and the RAS algorithm in Section 3.3, are formal, well known adjustment algorithms.

In practice, however, modellers make limited use of such algorithms. Much of the adjustment to the values found in primary data sources occurs in the *ad hoc*

adjustment procedures used to derive consistent aggregate totals. These *ad hoc* adjustments should also be held constant in the sensitivity analyses procedures presented here, but they are seldom recorded, and thus, unreproducible. The normative implication of the CPSA and extended sensitivity analysis procedures is that modellers be precise about recording all of their data adjustments. Paradoxically, these sensitivity analyses require modellers to take greater notice of how they adjust their data and of the uncertainties in the initial data values, but they dispense with the need to describe them to the final model users in detail by summarizing their effects via terse confidence intervals over the model results.

Chapter 4

The Significance of the Data Adjustment Algorithm Choice

4.1 Introduction

The problem of adjusting a matrix so that it meets consistency criteria arises not only in the derivation of input-output and social accounting matrices, but also in the derivation of the benchmark equilibrium data sets (BEDs) used to calibrate applied general equilibrium models. Many matrix adjustment algorithms exist, and while their characteristics are the focus of much of the input-output and social accounting literature, the basis for choosing an adjustment algorithm and the consequences of that choice are issues which have been unexplored by applied general equilibrium modellers.

This chapter proposes criteria for the adjustment algorithm choice in applied general equilibrium models. It argues that the basis for algorithm choice should be the effect of an adjustment algorithm on the statistical properties of the model results rather than its effect on the adjusted data, as has been the criterion in the input-output and social accounting literature. Although applied general equilibrium models are considered deterministic, the unbalanced data that are used as inputs to the modelling process are random variables, and hence, the model results are also random variables. The model structure and adjustment algorithm together form a complex rule for mapping the random input variables - the unbalanced data - into the random output variables - the model results. If the model structure component of this rule is constant, the statistical properties of the distribution of the model results vary only with the choice of the adjustment algorithm. An adjustment algorithm which leads to a

distribution for the model results with an unbiased mean and a low dispersion performs better than one which leads to a distribution with a biased mean and a high dispersion.

4.1.1 Algorithm Evaluation Experiments

Having proposed criteria for choosing adjustment algorithms, this chapter then uses them to evaluate several well known adjustment algorithms for a small tax model. The complexity of applied general equilibrium models is too great to allow an analytical evaluation of the adjustment algorithms. Instead, this evaluation is undertaken experimentally. It uses Monte Carlo simulations in which the data generating process for the unbalanced input data is specified, and in which the true model results are also known. In the Monte Carlo simulations, repeated samples of the unbalanced input data are generated, and each is balanced using all of the adjustment algorithms being tested to create a series of algorithm-dependent adjusted data matrices. Each balanced data matrix is then used to solve the model and to find model results.

The process is repeated for the sample of unbalanced data sets, so that each adjustment algorithm produces a sample of model results. This sample is used to find the expected values and standard deviations of the model results, for each algorithm. The bias associated with each adjustment algorithm is given by comparing the expected values of the model results generated by that algorithm with the known, true model results, and this bias is tested to determine if it is statistically different from zero. To capture both the statistical bias and efficiency implications of the adjustment algorithms for the model results, the bias and the variance of the sample are also used to construct

a second measure of adjustment algorithm performance, the mean square error of the model results.

The performance ranking of the adjustment algorithms that emerges from this experiment is specific to the small tax model, and therefore offers no general conclusions about the preferability of one adjustment algorithm to another. However, the fact that the algorithms perform quite differently from each other in this model, implies that the choice of adjustment algorithm cannot be overlooked as a significant factor affecting the results of other models. Thus, modellers have a clear incentive to evaluate alternative algorithms before choosing one. The simulation results in this chapter suggest that the development of an adjustment algorithm evaluation technique that allows modellers to choose the optimal adjustment algorithm for the model structure and policy experiment under consideration is an important direction for further research in applied general equilibrium model calibration.

4.1.2 Why Algorithms Perform Differently

Having established that adjustment algorithms perform differently from one another, this chapter uses a further set of simulations to explore why one algorithm may perform better than another. One explanation is that some of the elements of the data matrix are more important in determining the model results for a specific policy experiment than other elements, and that the preferred algorithm should be the one that minimizes the bias in those matrix elements which are expected to be relatively more important for the model results.

For example, the policy experiment in the small tax model is to impose an *ad valorem* tax on the use of capital in the manufacturing sector and to observe the effect of this tax on the welfare of the single consumer identified in the model. The *ex ante* expectation is that the welfare result in this model is affected most by the matrix elements that determine the input shares in manufacturing production. In the modelled economy, the manufacturing sector is small relative to the non-manufacturing sector, and hence the matrix elements for factor inputs to manufacturing are small relative to the other data elements. Under this argument, an adjustment algorithm that minimizes the bias in smaller data elements by allowing them to remain close to their initial values and shifts the relative burden of adjustment to the larger data elements, which are less important for the policy experiment, should perform better than one that does not.

The results from the experiments support this line of reasoning. One of the algorithms tested - a quadratic minimization algorithm that is weighted by the value of the initial matrix element squared - adjusts smaller elements of the unbalanced matrix relatively less than any of the other algorithms tested and it emerges as the preferred adjustment algorithm.

One implication of this argument is that if the elements of the data that determine input shares in manufacturing production are not small relative to the other data, this adjustment algorithm should no longer perform better than the other algorithms under consideration. A further set of experiments is undertaken to explore the link between the relative magnitudes of elements in the data set and the performance of adjustment algorithms. In this set of experiments, the relative magnitudes of the elements of the data set are altered and the adjustment algorithm evaluation exercise is

repeated. The results of this set of experiments indicate that as the magnitudes of elements of the data that determine input shares in manufacturing production change relative to the magnitudes of the other data elements, the performance ranking of the tested algorithms also changes.

One result which emerges from previous adjustment algorithm evaluations in the input-output and social accounting literature is that algorithms which incorporate additional information, such as the variances of the data or values for the row and column totals of a matrix, perform better than those which do not (Khan 1993; Harrigan, McGilvray and McNicoll 1980; Stone 1978; and Byron 1978; 1996). The experiments undertaken here compare the performances of adjustment algorithms which use a uniform amount of information, since a modeller is faced with fixed information about the data before choosing an algorithm.

Four information classes of algorithm which reflect the information a modeller is likely to have about the data are tested - those which incorporate information on i) only the unbalanced matrix, ii) the unbalanced matrix and the true value of the row and column totals, iii) the unbalanced matrix and its variance-covariance matrix, iv) the unbalanced matrix, its variance-covariance matrix and the true row and column totals. The unsurprising conclusion that adjustment algorithms using full-information are preferred to their partial information counterparts is the only generalization which emerges from these simulations. Within an information class, however, the best choice of adjustment algorithm remains model specific.

This chapter is organized as follows. Section 4.2 argues that the criteria for choosing an adjustment algorithm in applied general equilibrium modelling should be

its effect on the statistical properties of the model results, while Section 4.3 discusses the adjustment algorithms to be evaluated in this chapter. Section 4.4 presents the algorithm evaluation experiment methodology and results. Section 4.5 examines the sensitivity of the adjustment algorithm performance ranking from Section 4.4 to changes in the relative magnitudes of the elements of the data set. Section 4.6 concludes.

4.2 The Criteria for Adjustment Algorithm Evaluation

This chapter argues that the choice of adjustment algorithm should be based on its effect on the statistical properties of the model results, rather than on criteria relating to the adjusted data matrix.[28] Model results are the heart of the modelling exercise: they are often used to inform the policy process, and erroneous model results can have potentially serious repercussions in the real economy. Thus, modellers are concerned with minimizing the likelihood of such errors, and the choices made in the modelling process, including the choice of adjustment algorithm, are undertaken to meet this objective.

Underlying the wish of modellers to minimize the error in their model results is the idea that some 'true' model result exists, but its value is unknown and unobservable. Instead, the modeller estimates the value of the model results from a single unbalanced raw data set and a set of elasticities as inputs. This process is often

[28] Others (see, for example, Schneider and Zenios 1990) also consider practical criteria such as the ease of implementation, and the computer time required to balance a matrix. As computing technology improves, these practical considerations are likely to diminish in importance.

considered to be deterministic since applied general equilibrium models admit no explicit stochastic structure. If the adjustment algorithm, the elasticity specification, the model structure and the policy experiment are all held constant, a given set of raw data will yield a unique model solution (assuming an absence of multiple equilibria) and a unique set of model results. Because only one raw data set is ever balanced and then used to calibrate the model, no statistical considerations arise.

The problem with this view is that it ignores the fact that the data inputs are random variables which are subject to measurement errors. The unbalanced data set used by the modeller is generated by a random realization from the probability distribution for each random element. If the modeller were to undertake the production and consumption surveys again so that the values for the raw data would be measured a second time, the second raw data set would very likely be different from the first - it would be generated by a second random realization from each variable's probability density function. In principle, the modeller could generate a large sample of random unbalanced raw data sets. If each of these data samples were to be balanced with a specific adjustment algorithm and then used to calibrate the model, each would lead to a set of model results so that the sample of input data would lead to a sample of model results. Because the model results are generated using random variables as inputs, they too are random variables.

The process of deriving estimates of the model results from the sample of unbalanced data sets is one of applying a mapping from one set of random variables, the unbalanced data, into another set, the model results. In this case the mapping rule is complex. It consists of the data adjustment procedure, the model, and the policy

experiment. If the model and policy experiment components of this rule are held constant, however, the choice of an adjustment algorithm determines the statistical properties of the distribution of the model results derived from a given sample of unadjusted data. The distributions of the model results derived using various adjustment algorithms will each have means that are estimators of the true model results, and variances that capture the dispersion in those model results. An adjustment algorithm which performs well will yield a distribution for the model results that has an unbiased mean and a low dispersion.

If **R** denotes the true model result and $\overline{\mathrm{R}}_\beta$ is the model result from using adjustment algorithm β, then the bias of $\overline{\mathrm{R}}_\beta$, denoted by $\mathbf{B}_\beta$, is given by the difference between the expected value of the model result and the true model result,

$$\mathbf{B}_\beta = E(\overline{\mathrm{R}}_\beta) - \mathbf{R}. \tag{4.1}$$

If two adjustment algorithms yield unbiased, or low-bias estimates of the mean value for the model results, the one which yields a lower dispersion in the distribution of the model results will be preferred. The mean square error of the model results derived using adjustment algorithm β, $\mathrm{MSE}(\overline{\mathrm{R}}_\beta)$, accommodates a tradeoff between dispersion and bias and is given by the expectation of the square of the difference between the estimate of the model result and the true model result,

$$\mathrm{MSE}(\overline{\mathrm{R}}_\beta) = E[(\overline{\mathrm{R}}_\beta - \mathbf{R})^2]. \tag{4.2}$$

Thus, a preferred adjustment algorithm yields distributions for the model results that have unbiased mean values and low mean square errors.

The question which then arises is whether the adjustment algorithms that are commonly used by applied general equilibrium modellers differ significantly in their effect on the statistical properties of the model results. If they do not, then the choice of algorithm is peripheral to the modelling process, but if they do, then modellers have an incentive to exercise care in their choice of adjustment algorithm. The experiments in Section 4.4 of this chapter show that adjustment algorithms can differ substantially in their performance. Given a choice of adjustment algorithms and a desire to minimize the likelihood of error in the model results, the modeller should thus choose the adjustment algorithm which minimizes the bias and/or the mean square error in the distributions of the model results. Even though the modeller uses only one unadjusted data set, the probability of minimizing the error in the model results from that single data set is smaller using an adjustment algorithm that performs well than one that performs poorly.

4.2.1 Evaluation Criteria in Previous Work

The criteria for choosing an adjustment algorithm here contrast with previous work in the social accounting and input-output literature, where the focus has been the effect of the adjustment algorithm on the properties of the balanced matrix, instead of the model results (see, for example, Khan, 1993; Schneider and Zenios, 1990; Günlük-Şenesen and Bates, 1988; Parikh, 1979; Lynch, 1976; Malizia and Bond, 1974). Modellers are, of

course, more concerned with how the choice of algorithm affects the model results than with its implications for the adjusted matrix *per se*. Furthermore, this work in the social accounting literature is of limited value for applied general equilibrium modellers because the statistical properties of the adjusted matrix arising from the use of a particular adjustment algorithm do not necessarily transmit to the model results derived from that adjusted data - an adjustment algorithm that yields an unbiased adjusted data matrix may not yield unbiased estimates for the model results.

This lack of equivalence exists because the adjusted data are mapped through the model into the model results. It is evident from the meaning of statistical bias. Consider an algorithm, β, that yields an adjusted, unbiased matrix, $\hat{\mathbf{A}}^{\beta}$. If $\mathbf{A}$ is the true matrix of balanced data, then by definition, $E(\hat{\mathbf{A}}^{\beta}) = \mathbf{A}$. Under the expectation operator, this unbiasedness does not imply that the expectation of a *function* of $\hat{\mathbf{A}}^{\beta}$ will be unbiased. Specifically, it does not imply that for all functions F, $E\,(F(\hat{\mathbf{A}}^{\beta})) = F((\mathbf{A}))$, particularly when F represents the equations of a highly non-linear general equilibrium model.

Social accountants have been forced to adopt model-neutral evaluation criteria because input-output tables and social accounting matrices have many potential applications, and are not associated with a specific set of model results. In contrast, the adjusted data matrix in an applied general equilibrium model, the BED, is constructed as an input to a specific model so that a mapping exists between an adjustment algorithm and the final model results. This link allows the effects of the adjustment algorithm on the statistical properties of the model results to form the basis of the algorithm choice in applied general equilibrium modelling.

4.3 Adjustment Algorithms

Ideally, the adjustment algorithm evaluation experiments undertaken in this chapter should reflect the adjustment algorithms used in current modelling practice, but no universally accepted or recognized adjustment methodology exists. The adjustment of the raw data matrix into a BED can, however, be characterized as falling into a general class of matrix adjustment problems for which several well-known algorithms have been developed. Surveys of these matrix adjustment algorithms form part of the discussion in Günlük-Şenesen and Bates (1988) and Schneider and Zenios (1990). These surveys identify two classes of matrix adjustment problem based on the required consistency conditions. Schneider and Zenios characterize these two problems as follows:

"**Problem 1.**

Given an m x n nonnegative matrix $\bar{\mathbf{A}}$ and positive vectors **u** and **v** of dimensions m and n, respectively, determine a nearby m x n nonnegative matrix $\hat{\mathbf{A}}$ such that $\sum_{j=1}^{n} \hat{a}_{ij} = u_i$ for $i = 1, 2,..., m$, $\sum_{i=1}^{m} \hat{a}_{ij} = v_j$ for $j = 1, 2,..., n$, and $\hat{a}_{ij} > 0$ only if $a_{ij} > 0$.

Problem 2.

Given an n x n nonnegative matrix $\bar{\mathbf{A}}$, determine a nearby n x n nonnegative matrix $\hat{\mathbf{A}}$ such that $\sum_{i=1}^{n} \hat{a}_{ij} = \sum_{i=1}^{n} \hat{a}_{ji}$, for $i = 1, 2,..., n$, and $\hat{a}_{ij} > 0$ only if $a_{ij} > 0$. " (*source*: Schneider and Zenios 1990, p.440.)[29]

---

[29] Matrix notation has been changed to be consistent with Section 4.2.

Thus, the first class of problems finds a balanced matrix such that the rows and columns sum to predetermined values whereas the second imposes the 'biproportionality' restriction, discussed in Chapter 2.[30] The emphasis on finding a *nearby* matrix contains the implicit expectation that the errors associated with each matrix element are zero.

Both problems arise in the derivation of microconsistent benchmark data sets for applied general equilibrium models. The first stage in adjusting data typically finds consistent aggregate values. Because, at this stage, matrix biproportionality is the main restriction on data, the modeller solves Problem 2. The second adjustment stage uses the aggregate values derived in the first stage as constraints for the adjustment of submatrices in the model, thereby solving Problem 1.

Although no accepted universal methodology exists for the derivation of a model's BED, in practice the use of formal adjustment algorithms tends to be limited to Problem 1. Modellers are more likely to employ *ad hoc* algorithms to derive consistent aggregate values, and to resort to formal adjustment algorithms, particularly the RAS algorithm, to derive consistent consumption and production submatrices. Part of the rationale for this distinction is that modellers typically take the initial values for the submatrices from a single source and have no *a priori* information on which to distinguish the individual elements in the adjustment. In such a case, the use of a formal, mechanical adjustment algorithm is appealing.

---

[30] In the literature on the derivation of balanced input-output matrices (for example, Lecomber 1975; Günlük-Şenesen and Bates, 1988), part of the evaluation criteria for adjustment algorithms is whether they preserve the signs of the initial data. The presence of negative elements in the matrix affects the performance of some of these algorithms. Because a BED presents transactions values in an economy, it can be formulated as a non-negative matrix, and hence the Schneider and Zenios characterization of matrix adjustment is used here.

Aggregate totals, however, are often derived from various sources, and modellers are likely to develop subjective assessments of their reliability and model suitability. In this case, formal, mechanical algorithms which do not incorporate such information may be regarded as inadequate. *Ad hoc* adjustments are one means of including these assessments. The Stone-Byron adjustment algorithm, discussed below, provides a formal alternative for their inclusion.

4.3.1 *Ad hoc* Adjustment Procedures

From a clarity standpoint, an *ad hoc* balancing of the aggregate data is unsatisfactory since it is seldom documented. The absence of documentation means that the procedure cannot be reproduced and that the modeller cannot undertake sensitivity analysis with respect to the initial data, as discussed in Chapter 3, or with respect to the algorithm itself. It is also reminiscent of the attitude challenged in this chapter, that the way in which the BED is derived is unimportant for the model results.[31]

On the other hand, the adjustments made in these *ad hoc* procedures may reflect an intuitive recognition of the link between the data adjustment and the model results. For example, a modeller who is exploring changes in trade policy is likely to retain the

[31] Although modellers differ in the care with which they undertake data adjustments, they can be cavalier about the derivation of the BED, (or equivalently the social accounting matrix (SAM) which forms the BED in some models). De Melo (1988) p. 323 summarizes this characteristic by stating that in the applied general equilibrium literature "the accounting framework receives attention, although the main concern is with modelling and policy issues rather than with the necessary, underlying SAM." In a footnote, he reiterates this point with reference to specific modellers: "...Dervis, de Melo and Robinson (1982), for example, view the construction of a balanced SAM for calibration as important, but auxiliary and somewhat separate from modelling." (de Melo, (1988) p. 323, footnote 3.)

raw data on aggregate trade flows on the basis that they are the most important elements of data for the specific modelling exercise, and allow other data to absorb the changes required for consistency. Faced with the same initial data, a modeller looking at fiscal policy issues may hold data from the domestic accounts as close to their raw values as possible and allow the trade flows data to absorb imbalances. In such cases, the choice of *ad hoc* adjustment procedures over formal, mechanical algorithms, allows the modeller the flexibility to introduce an implicit judgement about the relative significance of different data elements for the model results. The problem remains, however, that these judgements are implicit rather than explicit and the modeller has no way of testing their effects on the model results.

4.3.2 Adjustment Algorithms for Evaluation Experiments

Because *ad hoc* algorithms are largely undocumented and hence not reproducible, the evaluation experiments undertaken here are limited to published, formal adjustment algorithms. The adjustment algorithms to be tested solve both Problems 1 and 2. Since modellers typically employ formal adjustment algorithms where row and column totals are known, the results of any algorithm evaluation experiments undertaken for Problem 1 will be germane to current modelling practice, whereas those directed towards Problem 2 will underline the need for modellers to be explicit in reporting the derivation of the aggregate totals.

The algorithms to be evaluated are formally described in Table 4.1. Ideally, the choice should reflect algorithms which are in common use among modellers, but such

Table 4.1

Evaluated Adjustment Algorithms

Let $\hat{a}_{ij}$ be the ijth element of the adjusted matrix $\hat{\mathbf{A}}$ with variance v_{ij} and $\bar{a}_{ij}$ be the ijth element of the initial matrix $\bar{\mathbf{A}}$.
Let $\boldsymbol{Z}$ be the set of strictly positive elements in $\bar{\mathbf{A}}$ and let $\boldsymbol{X}$ be its zero elements.[1]

Algorithm Name[2]	Minimand	Constraints: Problem 1 — Given a matrix with row control totals R_i and column control totals C_j	Constraints: Problem 2 — Given a matrix without known totals
1. Unweighted quadratic minimization	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^2$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
2. $\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^2 (\bar{a}_{ij}^{-0.5})$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
3. $\bar{a}_{ij}$ -weighted quadratic minimization	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^2 (\bar{a}_{ij}^{-1})$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
4. $\bar{a}_{ij}^{2}$-weighted quadratic minimization	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^2 (\bar{a}_{ij}^{-2})$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
5. RAS (row and column scaling)	$\sum_{(ij) \in Z} \hat{a}_{ij} \ln (\hat{a}_{ij} [\bar{a}_{ij}^{-1}])$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	not applicable
6. DSS (diagonal similarity scaling)	$\sum_{(ij) \in Z} \hat{a}_{ij} [\ln (\hat{a}_{ij} [\bar{a}_{ij}^{-1}]) - 1]$	not applicable	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
7. Stephan	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^2 (v_{ij}^{-0.5})$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
8. Stone-Byron	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^2 (v_{ij}^{-1})$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$
9. Alternative polynomial	$\sum_{(ij) \in Z} (\hat{a}_{ij} - \bar{a}_{ij})^4 (v_{ij}^{-1})$	$\sum_j \hat{a}_{ij} = R_i$, $\sum_i \hat{a}_{ij} = C_j$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$	$\sum_j \hat{a}_{ij} = \sum_i \hat{a}_{ij}$, $\hat{a}_{(ij) \in X} = 0$, and $\hat{a}_{ij} \geq 0$

Note 1: All BEDs can be constructed with non-negative flows and thus negative matrix elements are not considered here.
Note 2: Except for RAS, DSS, Stephan and Stone-Byron, these adjustment algorithm names have been created for the purposes of this chapter.

information is scant. The RAS algorithm, due to Bacharach (1970), is included among those evaluated, partly because some modellers report its use,[32] and partly because it has been widely used and analyzed in the input-output and social accounting literature.[33,34] The choice of adjustment algorithms also includes the DSS algorithm discussed in Schneider and Zenios (1990), since it provides a logarithmic counterpart to RAS for solving the matrix balancing problem in the absence of known control totals.[35]

Anecdotal evidence suggests that modellers employ both the unweighted quadratic minimization algorithm, attributed to Almon (1968, cited in Lecomber, 1975) and the $\bar{a}_{ij}^{2}$-weighted quadratic minimization algorithm, and hence they are also included in this evaluation exercise. The $\bar{a}_{ij}$-weighted quadratic minimand proposed by Friedlander (1961, cited in Lecomber 1975) and the $\bar{a}_{ij}^{0.5}$-weighted quadratic minimization algorithm are evaluated so that the choice of algorithms in this chapter parallels that of Günlük-Şenesen and Bates (1988) more closely.

---

[32] See for example, Piggott and Whalley (1985), Chia, Wahba and Whalley (1992) and Brixen (1995).

[33] For example, Khan (1993) compares gross sectoral output from a 1984-85, survey-based input-output matrix for Pakistan with gross sectoral output implied by matrices updated from the 1974-75 input-output table using RAS and two proportions updating methods. Lynch (1986) examines the reliability of the RAS adjustment procedure using base matrices for UK 1963 data, updating them to 1968 and comparing the results with 1968 census matrices. Parikh (1979) evaluates RAS by updating the 1959 absorption flow matrices of nine European countries with the 1965 absorption flow technologies. Similarly Malizia and Bond (1974) evaluate RAS by comparing a survey-based 1967 input-output table for Washington State with an RAS-updated 1963 input-output matrix.

[34] Although RAS is presented here as an optimization algorithm, part of its early appeal was its much less computer intensive manifestation as a simple scaling algorithm. The description of RAS as a scaling algorithm is given in Section 2.4.4 and in footnote 26.

[35] To my knowledge, no applied general equilibrium modeller has used DSS.

The value of the exponent in the weights used in the quadratic minimization adjustment algorithms affects how matrix elements of different magnitudes are adjusted. Higher exponents mean that larger matrix elements bear proportionally more of the adjustment burden than do smaller elements. When this exponent is zero, as is the case in the unweighted quadratic minimization algorithm, the adjusted matrix will be relatively sparse since more of the smaller elements will be adjusted to zero. For a modeller, such sparseness represents no economic activity in smaller sectors and is likely to generate biased model results if such sectors are operational in the 'true' economy. Consequently, this adjustment algorithm is unlikely to perform as well as other algorithms in the experiments that follow.

Harrigan and Buchanan (1984) illustrate that a Taylor series expansion of the RAS minimand around the initial data values approximates the $\bar{a}_{ij}$-weighted quadratic minimand. Since both adjustment algorithms generate similar adjusted matrices from a given unadjusted matrix, they also lead to similar model results. Consequently, they should perform similarly.

A final set of adjustment algorithms incorporates additional information about the reliability of the individual matrix elements, allowing less reliable elements to be adjusted further from their initial values than more reliable elements. The central adjustment algorithm evaluated in this full information class is the Stone-Byron algorithm which, under the assumptions that the matrix elements are independently distributed and the expectations of their errors are zero,[36] can be represented as a

[36] This assumption is frequently employed in applications of the Stone-Byron algorithm. Stone (1978) makes this assumption in the illustrative example of the algorithm, as does Byron in the

weighted quadratic minimization algorithm where the weights are given by the variances of the matrix elements. This adjustment algorithm has risen to prominence in recent years among social accountants because of its desirable properties: the Stone-Byron adjustment algorithm yields the best linear unbiased estimator of the 'true' matrix (Byron, 1978) and, if the errors associated with each of the matrix elements are normally as well as independently distributed with expected values equal to zero, it also yields the maximum likelihood estimator of the true matrix (Weale, 1985). Its popularity among social accountants is the justification for including it here.

Two variations of full information adjustment algorithms are also included to contrast with the Stone-Byron algorithm, rather than because they receive much attention in the current literature. One variation, due to Stephan (1942, cited in Lecomber 1975) uses the standard deviations rather than the variances of the matrix elements as weights in the quadratic minimization. The second variation, termed here the 'alternative polynomial' algorithm, uses the variances as weights, but the exponent on the distance component of the minimand is 4 instead of 2 as in all other cases. For an adjustment algorithm to find a meaningful minimum distance between the adjusted and the unadjusted matrix, the exponent in the numeraire of the minimand must be even and strictly positive. The use of quadratic (or least squares) minimization algorithms

---

exposition of an alternative procedure for solving the optimization problem. Crossman (1988) in an application of the Stone-Byron algorithm to the derivation of an Australian SAM also makes this assumption, as does Byron (1996) in a Monte Carlo simulation exercise. Günlük-Şenesen and Bates (1988) also assume zero covariances in the Stone-Byron algorithm for their evaluation experiments.

predominates in the literature and the 'alternative polynomial' algorithm with the higher-valued exponent is employed here for comparative purposes only.

The evaluation exercises that follow are undertaken on the assumption that the problem facing modellers is to choose an adjustment algorithm in light of *fixed* information about the data. Hence the performance of adjustment algorithms is compared within an information class. Previous evaluations (Khan, 1993; Harrigan, McGilvray, and McNicoll, 1980; Günlük-Şenesen and Bates, 1988) show that increasing the information input to the adjustment problem leads to better estimates of the adjusted matrix. This result carries through to the argument that if all elements of the BED under one adjustment algorithm exhibit relatively less bias than under another, then the more policy relevant data elements, and hence the model results, are also likely to be less biased. That modellers should include as much information as possible in their adjustment algorithm is undisputed here, but the problem remains that for fixed information about the data, several adjustment options are available and the modeller must choose among them.

4.4 Adjustment Algorithm Evaluation Experiments

The approach to adjustment algorithm evaluation undertaken here is experimental rather than analytical because the mapping of an adjusted matrix through an applied general equilibrium model is sufficiently complex for an analytical evaluation of the effects of adjustment algorithm choice on model results to be impractical. Instead, Monte Carlo experiments are used which generate random data from a specified data generating

process. Because the data generating process is known, the true values of the data and of the model results are also known.

The experiments evaluate the adjustment algorithms described in Table 4.1 by considering their effects on the statistical properties of the welfare result from a simple applied general equilibrium tax model. Under the Monte Carlo methodology, the true BED is perturbed by attaching a random error from a specified data generating process to each matrix element, resulting in an unbalanced raw data matrix. The raw data matrix is adjusted using each of the algorithms under consideration, and each of the ensuing balanced matrices is then employed to calibrate and solve the model. A series of model results emerges (in this case a welfare change measured by Hicksian equivalent variation in money metric terms for the single model consumer), where each result is associated with the adjustment algorithm used in its derivation. The process is repeated for a sample of unbalanced raw data matrices to generate a sample of equivalent variations associated with each adjustment algorithm. The sample mean of the equivalent variation associated with each adjustment algorithm and the sample standard deviation are then found. Together with the equivalent variation arising from calibrating and solving the model using the true BED, these values are used to find the statistical bias of the equivalent variation associated with each adjustment algorithm and to test the hypothesis that its value is zero.

4.4.1 Model Description

The experiments employ a 1x2x2 model with a single consumer, two factors of production, capital and labour, and two commodities, manufactured and non-manufactured goods. The consumer is endowed with both factors. In the base case version of the model, all taxes are zero. Table 4.2 summarizes the model structure and its elasticity specification, while the true benchmark data used to calibrate this model for the Monte Carlo experiments are given in Table 4.3. [37]

The true benchmark data set has been constructed so that its elements sum to 100, and so that the manufacturing sector uses capital and labour in a 1:4 ratio and the non-manufacturing sector in the ratio 4:1. The value of non-manufactured goods in consumption is four times that of manufactured goods. These ratios and the normalization provide a way of standardizing the relative magnitudes of the data, so that they can be systematically altered in subsequent experiments.

Policy analysis in applied general equilibrium models is undertaken through counterfactual analysis - a model parameter is perturbed, a new equilibrium is computed, and the subsequent vector of endogenous variables is compared to the benchmark equilibrium. In the counterfactual experiment for this model, the perturbed parameter is a tax rate. A 50 percent tax is levied on the use of capital in the manufacturing sector. The tax revenues are returned to the consumer, resulting in a welfare change, measured in money metric terms by the Hicksian equivalent variation.

---

[37] This model is the single consumer version of the model used for the illustration of CPSA in Chapter 3, but the consumer here has Cobb-Douglas rather than CES preferences and the production elasticities differ.

Table 4.2

Structure of the Model for the Monte Carlo Experiment

Production

- Output is produced using capital and labour combined in proportions implied by CES technology in each sector.
- The elasticity of substitution in the production of manufactured goods is 0.75 and in that of non-manufactured goods, 0.75.
- Share parameters for the CES function are calibrated from the BED.

Consumption

- The utility of the sole consumer is a Cobb-Douglas function of manufactured and non-manufactured goods.
- Share parameters for the Cobb-Douglas function are calibrated from the BED.

Endowments

- The consumer is endowed with all capital and labour.

Equilibrium Conditions

- Markets clear for all goods and factors.
- Zero profits are made in each sector.
- The consumer's expenditures equal his/her income.

Counterfactual

- A 50 percent tax is levied on the use of capital in the production of manufactured goods.
- All tax revenues are returned to the consumer.
- Welfare changes for the consumer are measured by equivalent variation: $EV = [(U_c - U_b) / U_b]I_0$ where U_b is the utility of the consumer in the base case, U_c is utility after the imposition of the tax, and I_0 is the benchmark income.

Table 4.3

True Benchmark Equilibrium Data Set Used in the Monte Carlo Experiment[1,2]

Transactions by agents are in units of currency

	Expenditure				
Income	Capital	Labour	Manufactures Sector	Non-Manufactures Sector	Consumer
Capital	0	0	1.33	21.33	0
Labour	0	0	5.33	5.33	0
Manufactures Sector	0	0	0	0	6.67
Non-Manufactures Sector	0	0	0	0	26.67
Consumer	22.67	10.67	0	0	0

note 1: Rows denote incomes and columns denote expenditures.
note 2: Values are rounded and hence biproportionality is not exact.

This equivalent variation is taken to be the model result of greatest interest and is the basis of the adjustment algorithm evaluations. The true value of the equivalent variation, derived using the true benchmark data, is -0.00078 units of currency.

4.4.2 Monte Carlo Experiment Methodology

The first step of the experiment is to specify the data generating process for the modelled economy. The data generating process is assumed to be one in which a random error is attached to each non-zero element of the 'true' benchmark data set. Let matrix **A** with elements a_{ij} be the true balanced benchmark data set given in Table 4.2. The unbalanced raw data set, $\bar{\mathbf{A}}$, with elements $\bar{a}_{ij}$ is derived by attaching a random error term, ϵ_{ij}, where $E(\epsilon_{ij}) = 0$, to each a_{ij} so that $\bar{a}_{ij} = a_{ij} + \epsilon_{ij}$. For the purposes of this experiment, the data generating process in the economy is assumed to be such that $\epsilon_{ij} \sim N(0, \sigma_{ij}^2)$, $E(\epsilon_{ij}\epsilon_{vs}) = 0$ for all $i \neq v$ and/or $j \neq s$[38] and $\epsilon_{ij} = 0$ for all $a_{ij} = 0$.

The values of the standard deviations for the normal distribution, the σ_{ij}, are assumed to be proportional to the a_{ij}, and are given in Table 4.4. To find a rough, but somewhat realistic approximation for the orders of magnitude of the standard deviations, a time series was constructed from UK data, that corresponds to the matrix elements in the model's BED. The estimate for the standard deviation of each matrix element in the BED is obtained by multiplying the true BED value estimate by the ratio of the standard deviation to the mean for the corresponding UK data element. The UK

[38] The assumption that the error terms are independently distributed is a simplification which is frequently employed in the derivation of social accounting matrices. See Crossman (1988) for an example.

Table 4.4

Standard Deviations Specified for the Data Generating Process

Matrix Element	Standard Deviation (proportion of the mean)
Capital Endowment	0.20
Labour Endowment	0.13
Labour in Manufacturing	0.10
Capital in Manufacturing	0.11
Labour in Non-Manufacturing	0.18
Manufacturing in Consumption	0.08
Non-Manufacturing in Consumption	0.21

data series run from 1970 to 1994, and their standard deviations include technological change as well as measurement error.[39] These values are, therefore, likely to be somewhat greater than the true values of the standard deviations sought, but because only rough orders of magnitude are required for this simulation exercise, no correction is made for the technological change.[40]

Once the matrix has been perturbed, it is adjusted using each of the algorithms in Table 4.1. The experiments are undertaken first for the case in which the modeller knows the value of the control totals (Problem 1), and repeated later for the case where biproportionality forms the main constraint (Problem 2). The control totals for Problem

---

[39] The time series were constructed as follows using current values. The value of manufacturing in total consumption is given by the series: "Total Manufacturing: value added in factor values," and the use of labour in manufacturing is given by the series: "Total Manufacturing: wages and salaries paid to employees in manufacturing." Both series are UNIDO data taken from the International Statistical Yearbook database. The value of non-manufacturing in total consumption is taken as the residual of GDP when manufacturing has been taken into account. GDP is taken from the UK Office of National Statistics series "Gross Domestic Product SA/ National Product PD BLN," also found in the International Statistical Yearbook database. The use of capital in manufacturing is the difference between the total value of manufacturing and the use of labour in manufacturing. The labour endowment in the economy is given by "GDP: all industries income from employment," a UK Central Statistical Office, Quarterly National Accounts series found in the Navidata database. Subtracting the use of labour in manufacturing from the labour endowment, gives the use of labour in the non-manufacturing sector, which together with the value of output from the non-manufacturing sector, gives the use of capital in the non-manufacturing sector. The capital endowment in the economy is the residual from subtracting the labour endowment from GDP. Finally, to remove the effects of inflation from the variability in the time series, all series are adjusted by a GDP deflator constructed from ratio of the current GDP series above and the series: "Gross Domestic Product SA/ National Product 1990 PD BLN" taken from the same source as the current GDP series.

[40] The assumption of normality for the errors means that the probability that $\tilde{a}_{ij} < 0$ is non-zero. The standard deviations given in Table 4.4 mean that this probability is very low and the random sample of matrices for the simulations in this chapter included no negative elements. Thus, although it is not a problem for the experiments in this chapter, the possibility of negative data elements might cause problems in a repetition of these experiments, so that a truncated normal distribution(truncated at zero) would have been a preferable specification for the data generating process.

1 are given by the true row and column totals in Table 4.3. If $\beta = 1,...,8$ indexes the eight adjustment algorithms applied to each problem (recalling that RAS can only be applied to Problem 1 and DSS to Problem 2), the balancing process for each results in a series of eight BEDs, $\hat{\mathbf{A}}^{\beta}$. Each $\hat{\mathbf{A}}^{\beta}$ is then employed to calibrate and solve the model, yielding a series of equivalent variations, denoted here by $\hat{w}^{\beta}$.

The Hicksian equivalent variation is a standard welfare measure in tax models. It measures the amount of money that must be transferred to the consumer before the imposition of the tax so that he/she would be indifferent between membership in the tax and in the no-tax economies. Since the tax is welfare reducing, the consumer would be willing to pay money to prevent the tax and the equivalent variation is negative. If, under the adjustment algorithm β, U_1^{β} is the consumer's utility after the imposition of the tax, U_0^{β} is his/her utility in the absence of the tax, and I^{β} is the consumer's income in the absence of the tax, then the Hicksian equivalent variation is given by

$$\hat{w}^{\beta} = [(U_1^{\beta} - U_0^{\beta})/U_0^{\beta}]\, I^{\beta}. \tag{4.3}$$

To generate insight about the expected value and dispersion of the equivalent variations generated by each adjustment algorithm, the process is repeated using 500 randomly generated unadjusted data sets. If $k = 1,.., 500$ indexes the repetitions, then each unbalanced matrix in the sample, $\bar{\mathbf{A}}^{k}$, is a random realization from the specified data generating process. Each $\bar{\mathbf{A}}^{k}$ is balanced by each of the eight adjustment algorithms, resulting in the 4000 BEDs, $\hat{\mathbf{A}}^{k\beta}$, which are employed to calibrate and solve the model for the welfare values $\hat{w}^{k\beta}$.

The 500 values of the Hicksian equivalent variation arising from each adjustment algorithm m, are used to find the mean for the equivalent variation associated with that algorithm. If this sample value is denoted by μ^{β}, then

$$\mu^{\beta} = (\sum_k \hat{w}^{k\beta})/500. \tag{4.4}$$

The standard deviation of the sample equivalent variation, s^{β}, yields a measure of dispersion in the sample of equivalent variations associated with each adjustment algorithm,

$$s^{\beta} = (\sum_k (\hat{w}^{k\beta} - \mu^{\beta})^2/499)^{0.5}. \tag{4.5}$$

One of the performance measures for an adjustment algorithm is whether the expected value of the equivalent variation that is generated by using that algorithm is biased. The sample bias, denoted in this case by B^{β}, is

$$B^{\beta} = \mu^{\beta} - w, \tag{4.6}$$

where w is the equivalent variation derived from using the true benchmark data set to calibrate and solve the model. Although the value of an estimator's bias can be used for comparative purposes, the question remains as to whether or not a given level of bias is meaningful - if an algorithm returns an average equivalent variation of -2.001 percent of GDP and the true value is -2.000 percent of GDP, is -0.001 a large or a small

difference? The answer to this question depends on the dispersion in the estimates of welfare; if all of the welfare estimates lie between -2.0015 and -2.0005, then this bias will be significant but if they lie between -6.0 and 2.0, it may not.

A hypothesis test determines whether or not the bias associated with a specific adjustment algorithm that has been calculated from the 500 welfare estimates is statistically different from zero. This hypothesis test is a two-tailed test large sample test in which a z-statistic, z, is formed such that[41]

$$z = (\mu^{\beta} - w)/[s^{\beta}/(500)^{0.5}]. \tag{4.7}$$

If z is greater than 1.96 and less than -1.96, the hypothesis that the bias is statistically different to zero cannot be rejected at the 5% significance level.

Bias, however, is not the only important characteristic of the adjustment algorithm-based sample of welfare results. The mean of this sample may be unbiased, but it may have a large dispersion. An alternative statistic for evaluating the performance of the adjustment algorithms is given by the mean squared error of the equivalent variations, the MSE^{β}, defined as

$$MSE^{\beta} = E[(\mu^{\beta} - w)^2]. \tag{4.8}$$

[41] The 500 samples are sufficient that the Central Limit Theorem, under which the sampling distribution is approximately normal, can be invoked.

The mean square error incorporates the principle that the modeller wishes to minimize the error in the expected value of the model result, but tempers this criterion with a desire to minimize the dispersion in that expected value.

4.4.3 Results in the Absence of Data Reliability Information

Because a modeller is faced with a fixed amount of information about the data before choosing an adjustment algorithm the evaluation of adjustment algorithm performance here is undertaken for four separate information classes: where the modeller knows i) only the unbalanced matrix, ii) the unbalanced matrix and the true value of the row and column totals, iii) the unbalanced matrix and its variance-covariance matrix, iv) the unbalanced matrix, its variance-covariance matrix and the true control totals. The adjustment algorithms evaluated in this initial set of simulations do not use information about the reliability of the data, and thus fall into the information classes i) and ii).

The sample mean, standard deviation, bias, test statistic and mean square error associated with each adjustment algorithm in these two information classes are given in Table 4.5. What emerges from Table 4.5 is that the adjustment algorithms tested clearly differ in their implications for the statistical properties of the model result. Under all the evaluation criteria, the a_{ij}^2-weighted quadratic minimization performs best, both when applied to solve Problem 1 with known control totals and Problem 2 where biproportionality forms the adjustment constraint. It also yields the only unbiased expectation of the equivalent variation. Hence, a modeller who is faced with the policy experiment, model, and data generating process presented in these experiments would

Table 4.5

Statistical Properties of the Algorithm-Based Sample of Hicksian Equivalent Variations (EV) in the Monte Carlo Experiment

(in the absence of data reliability information)

1. Adjustment algorithms applied to Problem 1 where control totals are known.

Adjustment Algorithm	EV Expected Value (true value = -0.00078) (x 10^2)	Bias (x 10^2)	Standard Deviation (x 10^2)	z-Statistic	Mean Squared Error (x 10^7)
Unweighted quadratic minimization	-0.096	-0.018	0.076	-5.33	6.158
$\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	-0.088	-0.011	0.046	-5.24	2.257
$\bar{a}_{ij}$ -weighted quadratic minimization	-0.084	-0.006	0.024	-5.83	0.623
$\bar{a}_{ij}^{2}$-weighted quadratic minimization	-0.078	-0.001	0.010	-1.20*	0.108
RAS (row and column scaling)	-0.083	-0.005	0.022	-5.62	0.491

2. Adjustment algorithms applied to Problem 2 without control totals.

Adjustment Algorithm	EV Expected Value (true value = -0.00078) (x 10^2)	Bias (x 10^2)	Standard Deviation (x 10^2)	z-Statistic	Mean Squared Error (x 10^7)
Unweighted quadratic minimization	-0.099	-0.021	0.083	-5.73	7.299
$\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	-0.084	-0.006	0.039	-3.50	1.522
$\bar{a}_{ij}$ -weighted quadratic minimization	-0.079	-0.002	0.015	-2.64	0.225
$\bar{a}_{ij}^{2}$-weighted quadratic minimization	-0.077	0.001	0.009	1.33*	0.078
DSS (diagonal similarity scaling)	-0.080	-0.003	0.016	-3.98	0.251

* denotes that the sample mean is unbiased at a 5% significance level.

likely overstate the welfare loss associated with the imposition of the tax, if he/she were to use any of the evaluated adjustment algorithms other than the $\bar{a}_{ij}^2$-weighted quadratic minimization.

Three other features of Table 4.5 are worth noting. In both Problems 1 and 2, the logarithmic optimization algorithm performs like the $\bar{a}_{ij}$-weighted quadratic minimization algorithm, supporting the claim of approximate similarity in Harrigan and Buchanan (1984). Second, the unweighted quadratic minimization algorithm performs notably more poorly than the others, except in the case of the *z*-statistic for the algorithms applied to Problem 1. Its poor performance is consistent with the expectation that relative to the other algorithms, it will introduce sparsity into more of the balanced matrices, reducing the opportunities for substitution into the non-operational sectors in response to the tax levy, and hence lead to bias in the model results. The bias associated with this adjustment algorithm is large, but the bias magnitude is offset by a large dispersion, so that the *z*-statistic is not larger than that of other biased estimators in the case of Problem 1.

Finally, the bias, the standard deviation, the mean square error, and in the case of Problem 2, the *z*-statistic are monotonic in the exponent on the weight of the weighted quadratic minimand. An explanation for this pattern is as follows. The relatively small elements of the true matrix are the capital and labour inputs in manufacturing and the labour input in the production of non-manufactured goods. Because the data generating process is such that the expected values of the elements of the randomly perturbed matrix are the values of the true matrix, and because their variance is proportional to the magnitude of their true values, the relatively smaller

elements in the *perturbed* matrices are also the capital and labour inputs into manufacturing and the capital input in the production of non-manufactured goods.

The smaller the value of exponent in the minimand, the larger is the proportional burden of adjustment for the small elements of the perturbed matrices. That is, smaller elements of the initial data matrix will adjust more from their initial values under the unweighted quadratic minimization algorithm (which could be expressed using the exponent zero as the $\bar{a}_{ij}^{0}$-weighted quadratic minimization algorithm), than under the $\bar{a}_{ij}^{2}$-weighted algorithm. The expected error of the initial value of the raw data is zero so that where these smaller elements adjust more, they are likely to have greater bias in their adjusted value. Table 4.6 gives the average proportional bias in the adjusted values for the matrix elements under each of the adjustment algorithms. As anticipated, the unweighted quadratic minimization algorithm leads to greater bias than the other adjustment algorithms, particularly with respect to the smaller matrix elements.

Several interesting features emerge from Table 4.6. First, in the case of algorithms solving Problem 1 with control totals, the ranking of bias for all of the adjusted matrix elements is the same as the ranking of the bias in the model results. This feature is consistent with the argument that as the overall bias in the adjusted matrix decreases, so will the bias in those elements which are the most important for the policy under consideration.

Where the algorithms solving Problem 2 are employed, more adjustment occurs in the data and the exact correlation of the results and the bias rankings no longer

Table 4.6

The Location of Bias in the Adjusted Matrix

Average Bias in the Adjusted Matrix Elements for the 500 Samples under Each Adjustment Algorithm

(percent of true value for non-zero matrix elements)

Algorithm	Matrix Element							
	capital in m'factures	labour in m'factures	capital in non-m'factures	labour in non-m'factures	capital endowment	labour endowment	m'factures consumed	non-m'factures consumed
Base Value	**1.33**	**5.33**	**21.33**	**5.33**	**22.67**	**10.67**	**6.67**	**26.67**
1. Adjustment algorithms applied to Problem 1 where control totals are known.[1]								
Unweighted quadratic minimization	84.2	21.1	5.3	21.1	0.0	0.0	0.0	0.0
$a_{ij}^{0.5}$-weighted quadratic minimization	46.7	11.7	2.9	11.7	0.0	0.0	0.0	0.0
a_{ij} -weighted quadratic minimization	22.8	5.7	1.4	5.7	0.0	0.0	0.0	0.0
a_{ij}^{2}-weighted quadratic minimization	10.2	2.6	0.6	2.6	0.0	0.0	0.0	0.0
RAS (row and column scaling)	21.2	5.3	1.3	5.3	0.0	0.0	0.0	0.0
2. Adjustment algorithms applied to Problem 2 without control totals.								
Unweighted quadratic minimization	94.2	15.5	12.5	29.9	11.0	11.0	14.7	10.3
$a_{ij}^{0.5}$-weighted quadratic minimization	41.7	9.3	12.6	19.5	11.4	8.4	8.2	10.1
a_{ij} -weighted quadratic minimization	15.7	7.1	13.4	14.3	12.4	7.0	5.5	10.7
a_{ij}^{2}-weighted quadratic minimization	9.2	5.9	15.5	12.2	14.6	6.2	4.7	12.5
DSS (diagonal similarity scaling)	16.2	7.1	12.6	14.1	11.7	6.9	5.6	10.0

note 1: Under the consistency constraints, the bias in the case where control totals are known is zero for data elements that are the sole occupants of a row or column.

holds[42] - the preferred adjustment algorithm introduces proportionally higher bias into three of the matrix elements than does any of the other algorithms. Because these elements are by far the largest in magnitude, the average absolute bias in the adjusted matrices is unlikely to be less than for the lower ranked algorithms.

The explanation for the better performance of the $\bar{a}_{ij}^2$-weighted quadratic minimization despite its introduction of larger absolute bias lies in *which* of the adjusted matrix values are the least biased. A policy change induces strong initial substitution and income responses that are mitigated through secondary general equilibrium interactions in the remainder of the economy. The response of the targeted agents is determined, in part, by the calibrated parameter specifications for their demand and supply functions, which in turn, are derived from the BED.

In the specific counterfactual experiment, where the capital used for manufacturing is the target of the tax, the values of the input share parameters in manufacturing will govern the initial response to the tax. These parameters are calibrated from the manufacturing sector factor inputs in the adjusted matrix. The capital and labour inputs into manufacturing are thus the matrix elements which contribute relatively more to the sensitivity of the model results. The algorithm which moves these more important matrix elements proportionally less from their initial and - under the expectation that the error of each of the matrix elements is zero- on average, true values, will result in less biased values for these elements and hence, less biased

[42] Because the benchmark data set contains four elements that are the sole occupants of a row and column (the two endowments and the two consumption expenditures), algorithms solving Problem 1 return unbiased values for these elements.

model results. Table 4.6 shows that of the algorithms tested, the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm introduces the least bias to the adjusted values for factor inputs in manufacturing. It does so because the values of the inputs into the manufacturing sector are smaller than the remaining (non-zero) matrix elements, and the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm moves smaller elements less from their initial values than do the other adjustment algorithms which have lower exponents in their weights.

4.4.4 Results with Data Reliability Information

In many cases, the modeller has *a priori* information about the relative reliability of the data which can be incorporated into the adjustment procedure. Instead of systematically differentiating among elements on the basis of their magnitudes as do the other adjustment algorithms the full information adjustment algorithms evaluated here allow less reliable matrix elements to adjust more than their more reliable counterparts.

Table 4.7 reports the results for those experiments which evaluate the full information algorithms in Table 4.1 applied to both Problems 1 and 2. A comparison with the results in Table 4.5 shows that, in general, full information algorithms are preferable to limited information algorithms, so that a modeller who has full information should always use a full information algorithm. Although the Stone-Byron algorithm is the only well known and studied algorithm in this class, it is not the only option. The argument that a modeller should base the choice of adjustment algorithm on its implications for the statistical properties of the model results also applies to the class of full-information adjustment algorithms.

Table 4.7

Statistical Properties of the Sample of Hicksian Equivalent Variations Adjusted Using Full-Information Algorithms

1. Adjustment algorithms applied to Problem 1 where control totals are known.

Adjustment Algorithm	EV Expected Value (true value = -0.00078) ($\times 10^2$)	Bias ($\times 10^2$)	Standard Deviation ($\times 10^2$)	z-Statistic	Mean Squared Error ($\times 10^7$)
Stephan	-0.0772	0.0003	0.0087	0.80*	0.0759
Stone-Byron	-0.0773	0.0002	0.0085	0.61*	0.0727
Alternative Polynomial	-0.0785	-0.0016	0.0161	-2.24	0.1141

2. Adjustment algorithms applied to Problem 2 without control totals.

Adjustment Algorithm	EV Expected Value (true value = -0.00078) ($\times 10^2$)	Bias ($\times 10^2$)	Standard Deviation ($\times 10^2$)	z-Statistic	Mean Squared Error ($\times 10^7$)
Stephan	-0.0768	0.0007	0.0088	1.76*	0.0782
Stone-Byron	-0.0772	0.0003	0.0087	1.68*	0.0764
Alternative Polynomial	-0.0776	<0.0001	0.0094	-0.10*	0.0891

* denotes that the sample mean is unbiased at a 5% significance level.

The results in Table 4.7 show that in the experiments applied in the presence of control totals (Problem 1), not all full information adjustment algorithms yield statistically unbiased mean values of the model results, and the Stone-Byron algorithm outperforms both of the other algorithms under all of the evaluation criteria. Where the adjustment algorithms have been applied to balance the matrices using the biproportionality constraint (Problem 2), the 'alternative polynomial' algorithm yields the least bias in the mean value for the equivalent variation. Because all three adjustment algorithms lead to statistically unbiased estimates of the equivalent variation, however, the Stone-Byron algorithm which yields the least dispersion in the values of the equivalent variation, and also the lowest mean square error would be preferred.

As with the $\bar{\mathbf{a}}_{ij}^2$-weighted quadratic minimization algorithm in the absence of reliability information, the superior performance of the Stone-Byron algorithm here is not a general result - it depends on the model, policy and data specifications. But again, the fact that a difference in performance exists in this model is evidence that the choice of adjustment algorithm is an important component of the modelling process.

4.5 The Influence of Relative Data Magnitudes on Algorithm Choice

The experiments that are undertaken in the previous section evaluate the performance of various adjustment algorithms in an economy where the manufacturing sector is small relative to the non-manufacturing sector. In the absence of information about the variances of the data, the $\bar{\mathbf{a}}_{ij}^2$-weighted quadratic minimization algorithm performs best.

The reason given for its superior performance is that it displaces the small-valued matrix elements (the factor inputs into the manufacturing sector) that are the most important for the tax policy experiment, less than the larger matrix elements that are less important for the policy experiment. The question then arises of how the performance of the adjustment algorithms will change if the factor inputs into the manufacturing sector are large relative to the other matrix elements, or more generally, the question arises of how the adjustment algorithms perform when the relative magnitudes of the data change.

To capture the effect of the data structure on the performance of adjustment algorithms, the experiments which follow are undertaken for a grid of 'true' matrices each with elements of different relative magnitudes. This grid is constructed by systematically altering the relative magnitudes of the ratios in the data. The model structure given in Table 4.2 contains three key ratios: the capital to labour ratio in the production of manufactures, the capital to labour ratio in the production of non-manufactures, and the ratio of manufactures to non-manufactures in consumption. To maintain a consistent basis for comparison, total factor income is held constant throughout. Fixing total factor income, however, leaves only two ratios in the data to be independently specified: the ratio of manufactures to non-manufactures in consumption, denoted by χ, and the ratio of capital to labour in the production of manufactured goods, denoted by ξ. By residual, the ratio of capital to labour in the production of non-manufactured goods is $1/\xi$, the reciprocal of that for manufactured goods.

Three values (small, medium and large) of each χ and ξ are considered here, resulting in a grid of nine 'true' data sets. The various data sets arise from the permutations of setting χ and ξ to 1/4, 1, and 4 so that, for example, the economy represented when $\chi = 1/4$ and $\xi = 4$ is one in which the consumer has a strong preference for non-manufactured goods, manufactures are produced using capital intensive technology, and non-manufactured goods are produced using labour intensive technology. The case considered in the previous experiments sets $\chi = 1/4$ and $\xi = 1/4$.

These ratios characterize the nine BEDs presented in Table 4.8. In each data set, the size of the economy is maintained constant across all permutations by normalizing the elements in the benchmark data set so that they sum to 100, and holding total factor income constant at 100/3. The 'flattest' BED - the one with the least dispersion in its elements - is the case in which χ and ξ are both unity. Here, both sectors use the same production technology and the consumer demands equal quantities of the two goods in the initial equilibrium.

The Monte Carlo simulation methodology is as before: each true BED forms the basis of a sample of 500 raw data sets, where the perturbations of the initial data elements are found by drawing randomly from a normal distribution. Each of the sample raw data sets is adjusted using each of the algorithms under consideration. The model is then calibrated using each balanced matrix, and solved to find the equivalent variation. In this case, the experiments are limited to the partial information algorithms in the presence and absence of known control totals since these provide sufficient evidence of the changing link between algorithm performance and the relative magnitudes of the true data.

Table 4.8

Non-Zero True BED Values for the Grid of Data Structures

Matrix Structure	**χ=0.25 ξ=0.25**	**χ=0.25 ξ=1**	**χ=0.25 ξ=4**	**χ=1 ξ=0.25**	**χ=1 ξ=1**	**χ=1 ξ=4**	**χ=4 ξ=0.25**	**χ=4 ξ=1**	**χ=4 ξ=4**
Matrix Element									
Capital Endowment	22.67	16.67	10.67	16.67	16.67	16.67	10.67	16.67	22.67
Labour Endowment	10.67	16.67	22.67	16.67	16.67	16.67	22.67	16.67	10.67
Capital in Manufacturing	1.33	3.33	5.33	3.33	8.33	13.33	5.33	13.33	21.33
Labour in Manufacturing	5.33	13.33	21.33	3.33	8.33	13.33	1.33	3.33	5.33
Capital in Non-Manufacturing	21.33	13.33	5.33	13.33	8.33	3.33	5.33	3.33	1.33
Labour in Non-Manufacturing	5.33	3.33	1.33	13.33	8.33	3.33	21.33	13.33	5.33
Manufacturing Consumption	6.67	16.67	26.67	6.67	16.67	26.67	6.67	16.67	26.67
Non-Manufacturing Consumption	26.67	16.67	6.67	26.67	16.67	6.67	26.67	16.67	6.67
Sum of Manufacturing inputs	6.66	16.66	26.66	6.66	16.66	26.66	6.66	16.66	26.66
Sum of non-Manufacturing inputs	26.66	16.66	6.66	26.66	16.66	6.66	26.66	16.66	6.66

Table 4.8 also gives the values of the inputs into the manufacturing and non-manufacturing sectors. In the three cases where $\xi = 1/4$, the manufacturing sector inputs are small relative to the non-manufacturing sector inputs, and the expectation is that the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm will perform relatively better than the other algorithms. In contrast, where $\xi = 4$, the manufacturing sector inputs are large relative to the non-manufacturing sector inputs, and the unweighted quadratic minimization algorithm, or one of the weighted algorithms with a lower exponent in the weight, should perform best.

4.5.1 The Effects of Alternative Data Structures

Table 4.9 presents the bias and the corresponding test statistic for the equivalent variations arising from using each of the tested adjustment algorithms to balance the random data samples generated under each of the nine different matrix structures, while Table 4.10 gives the value of the mean square error and the adjustment algorithm ranking in terms of the mean square error criterion. The results indicate unambiguously that the relative magnitudes of the data affect the performance ranking of the tested algorithms in terms of their implications for both the bias and the mean square error of the model results. Overall, however, the performance of the tested algorithms is poor under the bias criterion.

The one adjustment algorithm which yields the greatest number of unbiased estimates of the model inferences is the unweighted quadratic minimization algorithm. The reason for this result is that it produces a large dispersion in the values of the

Table 4.9

Bias and *z*-Statistic Ranking of Adjustment Algorithms Under Different Data Structures

Bias ($x10^2$) and *z*-Statistic of Hicksian Equivalent Variation Estimates Arising from Introducing a 50 Percent Tax on Manufacturing Capital

Matrix Structure	χ=0.25 ξ=0.25		χ=0.25 ξ=1		χ=0.25 ξ=4		χ=1 ξ=0.25		χ=1 ξ=1		χ=1 ξ=4		χ=4 ξ=0.25		χ=4 ξ=1		χ=4 ξ=4	
True Values of EV	-0.000775		-0.001712		-0.001889		-0.001804		-0.002975		-0.002025		-0.001881		-0.001886		-0.000890	
1. Adjustment algorithms applied to Problem 1 with known control totals.																		
Unweighted quadratic minimization	-0.0182	-5.33	-0.0020	-1.02*	0.0032	15.43	-0.0037	-1.63*	0.0027	14.62	0.0025	1.91*	0.0062	17.38	0.0050	3.41	0.0003	0.13*
$\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	-0.0108	-5.24	-0.0054	-3.73	0.0015	12.59	-0.0054	-3.33	0.0030	14.47	0.0043	3.59	0.0023	15.34	0.0040	3.39	0.0012	1.01*
$\bar{a}_{ij}$-weighted quadratic minimization	-0.0063	-5.83	-0.0059	-5.21	0.0008	9.07	-0.0055	-4.52	0.0036	12.52	0.0065	5.18	0.0010	12.82	0.0051	4.52	0.0025	2.57
$\bar{a}_{ij}^{2}$-weighted quadratic minimization	-0.0006	-1.20*	-0.0027	-3.00	0.0005	5.02	-0.0035	-4.16	0.0055	9.98	0.0099	6.79	0.0004	7.19	0.0091	7.24	0.0035	3.37
RAS	-0.0054	-5.62	-0.0049	-4.61	0.0007	9.57	-0.0037	-3.40	0.0031	13.98	0.0062	5.03	0.0007	13.83	0.0041	3.91	0.0024	2.47
2. Adjustment algorithms applied to Problem 2 without control totals.																		
Unweighted quadratic minimization	-0.0212	-5.73	-0.0018	-0.75*	0.0040	3.40	-0.0020	-0.80*	0.0049	2.77	0.0032	1.21*	0.0109	6.63	0.0066	2.14	-0.0089	-2.54
$\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	-0.0060	-3.50	-0.0012	-0.75*	0.0050	4.66	-0.0007	-0.43*	0.0092	5.34	0.0081	3.92	0.0084	6.44	0.0097	4.05	0.0034	1.83*
$\bar{a}_{ij}$-weighted quadratic minimization	-0.0018	-2.64	0.0005	0.43*	0.0072	7.11	0.0018	1.85*	0.0137	7.90	0.0121	6.34	0.0096	8.04	0.0116	5.75	0.0050	4.27
$\bar{a}_{ij}^{2}$-weighted quadratic minimization	0.0005	1.33*	0.0039	4.76	0.0119	11.24	0.0058	7.25	0.0206	11.12	0.0162	8.47	0.0123	10.04	0.0114	5.98	0.0042	3.87
DSS	-0.0028	-3.98	-0.0014	-1.24*	0.0053	5.36	0.0001	0.12*	0.0107	6.32	0.0098	5.15	0.0076	6.51	0.0089	4.42	0.0035	2.98

* denotes that the sample mean is unbiased at a 5% significance level.

Table 4.10

Mean Square Error x 10^7 (level and rank) of Adjustment Algorithms Under Different Data Structures

MSE Level and Ranking of Hicksian Equivalent Variation Estimates Arising from Introducing a 50 percent Tax on Manufacturing Capital

Matrix Structure	χ=0.25 ξ=0.25		χ=0.25 ξ=1		χ=0.25 ξ=4		χ=1 ξ=0.25		χ=1 ξ=1		χ=1 ξ=4		χ=4 ξ=0.25		χ=4 ξ=1		χ=4 ξ=4	
True Values of EV	-0.000775		-0.001712		-0.001889		-0.001804		-0.002975		-0.002025		-0.001881		-0.001886		-0.000890	
1. Adjustment algorithms applied to Problem 1 with known control totals.																		
Unweighted quadratic minimization	6.1576	5	1.9112	5	0.0326	5	2.5354	5	0.0239	1	0.8804	4	0.1005	5	1.1150	5	1.9239	5
$\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	2.2573	4	1.0600	4	0.0089	4	1.3651	4	0.0298	2	0.7277	1	0.0165	4	0.7119	3	0.7231	4
$\bar{a}_{ij}$-weighted quadratic minimization	0.6229	3	0.6650	3	0.0043	2	0.7677	3	0.0529	4	0.8187	3	0.0038	3	0.6499	2	0.4825	2
$\bar{a}_{ij}^{2}$-weighted quadratic minimization	0.1082	1	0.4215	1	0.0047	3	0.3631	1	0.1792	5	1.1608	5	0.0015	1	0.8816	4	0.5471	3
RAS	0.4913	2	0.5761	2	0.0032	1	0.5910	2	0.0347	3	0.7871	2	0.0019	2	0.5737	1	0.4724	1
2. Adjustment algorithms applied to Problem 2 without control totals.																		
Unweighted quadratic minimization	7.2994	5	3.0324	5	0.7225	5	3.1568	5	1.6071	3	3.4606	5	1.4805	5	4.7504	5	6.2131	5
$\bar{a}_{ij}^{0.5}$-weighted quadratic minimization	1.5215	4	1.3949	4	0.5910	3	1.1847	4	1.5655	2	2.2141	4	0.9284	4	2.9500	4	1.7085	4
$\bar{a}_{ij}$-weighted quadratic minimization	0.2245	2	0.6134	2	0.5661	2	0.4907	2	1.6842	4	1.9589	2	0.8100	2	2.1647	3	0.7086	3
$\bar{a}_{ij}^{2}$-weighted quadratic minimization	0.0775	1	0.3426	1	0.7043	4	0.3474	1	2.1358	5	2.0895	3	0.8982	3	1.9545	1	0.5969	1
DSS	0.2506	3	0.6370	3	0.5166	1	0.5077	3	1.5491	1	1.8938	1	0.7429	1	2.0897	2	0.7054	2

equivalent variation, so that the hypothesis of unbiasedness cannot be rejected. This reasoning is evident from Table 4.10, which indicates that despite the desirable bias properties for the model results of the unweighted quadratic minimization algorithm, in most cases the mean square error associated with its use is more than double that associated with each of the other tested algorithms.

The only case in which the unweighted quadratic minimization algorithm performs adequately under the mean square error criterion (ranks first where control totals are known and third where they are not) is where $\chi = 1$ and $\xi = 1$. This case represents the matrix structure in which the dispersion between the lowest and highest data values is the least. Where discrepancies in the magnitudes of data are large, the unweighted adjustment algorithm is more likely than the other adjusted algorithms to set the smallest valued matrix elements to zero, creating the conditions for zero-valued production parameters, and hence large dispersions in the model results. This mechanism is clearly less operative for 'flatter' data structures where the absolute deviation of adjusted data from their initial values will be similar for all of the elements of the matrix.

The *ex ante* expectation that the optimality of the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm should persist where the inputs into the manufacturing sector are small relative to the other matrix elements is only partially met. Of the three cases in Table 4.10 where the manufacturing inputs are small ($\xi = 1/4$), the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm only generates an unambiguously unbiased model result where $\chi = 1/4$. In the cases where $\chi = 1$ and $\chi = 4$, it provides the least biased result with known control totals. Under the BED structures in Table 4.8, the only

elements of the matrix which can possibly exhibit bias in their adjusted values when control totals are known are the four factor inputs, because the control totals guarantee that the balanced endowment and consumption values will equal their true values. Hence, in the case with known control totals, the importance of the relative magnitudes of the factor inputs is isolated, and the algorithm which adjusts the manufacturing inputs the least, leads to less bias in the model results. In the absence of control totals, all the matrix elements adjust and this effect is no longer isolated.

Although the bias criterion does not yield the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm as the preferred algorithm as often as expected, the experiment results under the mean square error criterion presented in Table 4.10 meet expectations better. The $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm yields the least mean square error in five of the six cases where $\xi = 1/4$. In the remaining case, where $\chi = 4$ (under the biproportionality constraint), it ranks third.

The expected poor performance of the $\bar{a}_{ij}^2$-weighted quadratic minimization algorithm relative to the algorithms with lower exponents in their weights when $\xi = 4$ is met in the bias results of five of the six cases, the exception being where $\chi = 1/4$ (under known control totals). In terms of the mean square error criterion, it is also met in all but one case. These results indicate that although the experiment outcomes could not be predicted exactly, in general they conform to prior expectations.

Finally, RAS and DSS consistently emerge as strong middle of the road options. They also mimic the results from the $\bar{a}_{ij}$-weighted quadratic minimization algorithm, confirming that the Harrigan and Buchanan (1984) result also holds for the grid of data structures tested here. While a least squares minimand may exist which outperforms the

logarithmic minimand, for any given structure in the experiments here, at least one quadratic minimand is a poorer option than its logarithmic counterpart. These results indicate that modellers' traditional reliance on RAS as an adjustment algorithm is not misguided.

4.6 Conclusion

This chapter proposes that the criteria for choosing a procedure to adjust data prior to calibration be the implications of that procedure for the statistical properties of the model results. Monte Carlo simulations show that several common adjustment algorithms differ in their implications for the bias and dispersion of the welfare results from a small tax model. Although the ranking of the adjustment algorithms is specific to the tax model, the fact that the choice of adjustment algorithm matters for this model is evidence that the choice of adjustment algorithm may affect the statistical properties of the results from other applied general equilibrium models.

One reason why adjustment algorithms perform differently from one another is that they differ in how they assign the burden of adjustment among the various unbalanced data elements. The data elements differ in their significance for the model results, and the adjustment algorithm which performs best in the simulations is the one that places the burden of adjustment on the less significant elements. A further set of simulations in which the relative magnitudes of the data are systematically varied, indicates that this line of reasoning has merit.

More generally, these experiments show that the choice of adjustment algorithm is not an auxiliary component of the modelling process, but should be considered by modellers as central to proper calibration. The results of the experiments in this chapter underline the need for a more integrated calibration process which features a systematic and model-specific evaluation procedure for choosing an adjustment algorithm. Although no such methodology is proposed here, finding one provides a future direction for research in applied general equilibrium model calibration.

Finally, the results of these experiments have wider implications for the reporting of national statistics. Rather than presenting a system of fully reconciled national accounts as an end product, statisticians could also provide unbalanced data together with their assessment of the reliability of those data, so that modellers can implement their own adjustment procedure in light of the specific problem on which they wish to focus. Such a practice could be undertaken at a reasonable cost since no information beyond that which statisticians actually employ is required.

Chapter 5

Decomposition Analysis Using Applied General Equilibrium Models

5.1 Introduction

One of the challenges facing economic historians is to unravel the relative individual contributions made by several simultaneous shocks to a known historical change. Applied general equilibrium modelling is a technique which has been adopted from contemporary policy analysis to this end. Unlike policy analysts, however, economic historians can know the net effect of the combined shocks from data. This chapter proposes a modification to the standard applied general modelling methodology that makes use of the information available to historians and offers a more systematic way of decomposing historical change into its component causes than is possible using the traditional modelling techniques.

Although economic theory provides a framework for assessing the qualitative implications of a specific historical event, historians must turn to numerical techniques to gain insight about its quantitative effects. Economic historians, including Temin (1971), James (1984), and Thomas (1987), have long advocated the use of applied general equilibrium analysis as one such numerical technique. The advantage that general equilibrium models have over partial equilibrium techniques is that they capture an economy's interactive responses to a shock. Furthermore, for economic historians faced with a paucity of time series data, applied general equilibrium modelling has the practical advantage over statistical techniques of requiring only a single observation for each variable.

These features of general equilibrium models have led some economic historians to adopt applied general equilibrium modelling in their analyses of economic history questions. In doing so, they have largely maintained the methodology of contemporary policy analysts. This traditional methodology, discussed in Chapter 2, is to specify a model which, when solved, yields a base period equilibrium as a model solution. The effects of a prospective policy change are then gleaned by introducing it to the model, solving for the hypothetical, 'counterfactual' equilibrium, and comparing this counterfactual solution to the base year equilibrium. Although modellers stress that the resulting insights are conditional on the model structure, they are also conditional on the assumption that the economy will only be subjected to the simulated changes. The possibility of unforeseen shocks diminishes the predictive power of such simulations.

Economic historians, however, are in a unique position relative to contemporary policy analysts because they can, in principle, identify the major shocks and policy changes that affected an economy over a specific interval. Furthermore, they can know the combined effect of those shocks on the economy from data so that for historians, unlike policy analysts, the characteristics of the final equilibrium hold little mystery. Of greater interest are the relative contributions of various shocks to economic changes. Thus, economic historians have used applied general equilibrium models to address questions such as the relative contribution of commodity price shocks and the potato famine to the fall in 19th century Irish agricultural labour demand (O'Rourke, 1991), the causes of the depression in Australia in the 1930s (Siriwardana, 1995), and the effects of technological improvements and factor growth during the industrial revolution on British output and trade patterns (Harley and Crafts, 1998).

5.1.1 Interactions Between Shocks

This chapter argues, however, that using the traditional general equilibrium modelling approach for historical questions yields an incomplete analysis because it ignores the interactive effects of simultaneous shocks. Typically, synergies exist between different shocks so that, for example, the net effect on GNP of a specific level of immigration coupled with a fall in world commodity prices is likely to be different from the effect on GNP of adding the individual effect of increasing immigration to the individual effect of the price change. The possibility that simultaneous shocks may be mutually enhancing or off-setting also implies that the effects of a single specific shock will depend on the presence or absence of other shocks in the economy. A full analysis of the effect of a specific shock would, therefore, include a series of conditional measures: its effect in the absence of all other shocks; its effect in the presence of each other shock individually; its effect in the presence of each pairwise combination of other shocks and so on. The final conditional measure would be the case in which the effect of the shock is measured conditional on the presence of all other shocks.

This chapter proposes that economic historians can use information about shocks and net changes in an economy over a specified interval to undertake a systematic decomposition analysis in which both the individual and the interactive effects of the various shocks are assessed. Decomposition analysis requires a simple modification to traditional applied general equilibrium modelling techniques. Instead of specifying a single base year equilibrium, as is the standard approach of policy analysts, equilibria for both an initial and a final year of interest are specified and the

main shocks to the economy in the interim identified. Calibration ensures that introducing all shocks to the initial equilibrium yields the final equilibrium as a model solution.

Decomposition analysis is undertaken by sequentially introducing individual shocks or subsets of shocks to the modelled system, solving the model and comparing variables from the ensuing solution equilibria with their values in one of the two known equilibria. By comparing the effects of individual shocks with the effects of combined shocks, the modeller can assess the synergies among different combinations of shocks. The various model solutions arising from decomposition analysis also allow the modeller to report conditional measures for the contribution to change of a particular shock of interest.

The decomposition analysis is illustrated in this chapter using a simple model of railroads in the US in the interval 1870-1890. The focus of the modelling exercise is the relationship between technological change in the railway sector and growth, as measured by GNP. Four types of shocks are considered: technological change in the railroad sector, factor endowment growth, changes in consumer preferences, and other technological change.

The overriding result of the simulations is that the shocks to the US economy in the late 19th century were highly interactive: the additional increase in GNP from technological change in the railroad sector in the presence of the remaining shocks is more than six times its value in their absence. The implication of this result is that a properly specified question about the effects of a particular historic change must be

made conditional on the status of the remaining shocks.[43] Hence, asking what the contribution of technical change in the railway sector between 1870 and 1890 was to the growth of US GNP is an ill-posed question. A more appropriate question would ask, for example, what the contribution of technical change in the railway sector to the growth of national product would have been had consumer preferences and factor endowments remained at their 1870 levels, and technology in other sectors of the economy been allowed to adjust to 1890s levels.

This chapter is organized as follows. Section 5.2 provides the framework for decomposition analysis. It gives a formal description of the traditional approach to applied general equilibrium modelling, characterises the modelling strategies of economic historians with respect to that approach, and describes the modifications to the traditional approach required for decomposition analysis. Section 5.3 introduces the illustrative model and provides a detailed discussion of its calibration. Section 5.4 gives the results of the decomposition analysis using the model, and Section 5.5 concludes.

5.2 The Modelling Framework for Decomposition Analysis

Although many modellers consider the joint effect of more than one shock, the possibility that simultaneous shocks to an economy may interact with one another has not been explored using applied general equilibrium models. The results of all modelling exercises are implicitly conditional on the presence or absence of such

[43] Because contemporary policy analysts always consider the effects of potential future policy changes relative to the current equilibrium, the implicit conditionality in the issues they address is that other potentially shockable parameters retain their base case values.

interactions. This section characterizes existing modelling activities as 'forward' simulations in which the absence of interactions among shocks is the implicit condition or 'backward' simulations, which implicitly include the interaction of all shocks. These forward and backward simulations represent the two endpoints for the spectrum of decomposition analysis introduced at the end of this section.

5.2.1 Forward Simulation

In recent models, economic historians have employed the traditional 'counterfactual' applied general equilibrium modelling methodology given in Shoven and Whalley (1992) and standardized in a multitude of trade and tax applications. This procedure, which has been explained in detail in Section 2.3, requires the modeller to specify an initial equilibrium for the economy. Calibration in this traditional modelling approach finds values for the model parameters that determine the static specification of the model in the initial period. These parameters include the values of the initial period policy parameters. Once the model is calibrated, the modeller specifies counterfactual values for the model's policy parameters. The counterfactual values are the shocks or policy changes, such as a change in tax rates or a different quota regime, which are the focus of the modelling exercise. The modeller then solves the model using the counterfactual policy parameters, and infers the effects of the policy change by

comparing the values of variables in the new, 'counterfactual' solution with their values in the initial, 'benchmark' equilibrium.[44]

This approach, under which a benchmark equilibrium is modelled and a future counterfactual is specified, is termed here 'forward simulation.' Forward simulations are *ceteris paribus* exercises - the effects of a specific policy change or a set of policy changes are gleaned by assuming that they will be the sole shocks to the economy. Any predictions made by the model using forward simulations are, of course, conditional on this *ceteris paribus* assumption.

Economic historians have adopted forward simulations to examine economic history questions. Their simulation procedures are analogous to those of contemporary policy analysts, except that the simulated changes are based on actual rather than potential shocks to the initial system. Siriwardana (1995), for example, uses 1934/1935 data (as a proxy to 1920s data) to calibrate a model of the Australian economy, introduces a sequence of exogenous shocks that are representative of the actual shocks faced by the Australian economy in the 1930s, including a decrease in real wages, a decrease in aggregate investment, an increase in tariffs, and a fall in export prices, and observes the effects of each on macroeconomic and sectoral variables to conclude that the largest contributions to the onset of the depression came from decreases in investment and falling export prices.

---

[44] Such models are static and the adjustment occurs instantaneously. Hence, they cannot address issues surrounding path dependency.

5.2.2 Backward Simulation

Unlike policy analysts, however, economic historians can observe the final equilibrium. Their ability to specify an *ex post* equilibrium means that they are not limited to *ceteris paribus* simulation exercises. They can undertake 'backwards simulations' in which they model the post-change economy as an equilibrium system and develop insight about the effect of a shock by simulating its removal. For example, Harley and Crafts (1998) seek to find the effects of technological improvements and factor growth during the Industrial Revolution in Britain by calibrating a model that reproduces post-revolution 1841 data as the initial equilibrium, and then removing the effects of technological change and factor growth. In the case of Harley and Crafts, the simulation experiment is undertaken by introducing values for technology parameters and factor endowments for a pre-Industrial Revolution year, 1770, and solving the model. The effects of technological change and factor growth are inferred by comparing the actual 1841 values of sectoral outputs and trade volumes with what they would have been in 1841 had the factor endowments and technology remained as they were in 1770.

Backward simulations, which are a unique feature of historical analysis, are *mutatis mutandis* exercises; the *ex post* equilibrium includes all economic changes up to the year for which it is derived, and the simulations are undertaken assuming that the economy is subjected to all shocks except the shocks of interest.

The relationship between forward and backward simulations depends on whether the modelled shocks are an exhaustive representation of the actual shocks affecting an economy. In addition to knowing the *ex post* equilibrium, economic

historians can, in principle, know what all of the main shocks to an economy were over a given time frame. If the values of all of the pre- and post-change policy parameters are known, then a backwards simulation in which the post-change equilibrium is the initial equilibrium and all of the values for the policy (or shock) parameters from the pre-change equilibrium are imposed as counterfactual parameters, will yield the pre-change equilibrium as the model solution, so that full forward and backward simulations, in which all of the policy changes and shocks are either introduced or removed, become equivalent procedures.[45]

5.2.3 Decomposition Analysis

The equivalence between full forward and backward simulations is not very interesting, since if the net effect of the simultaneous shocks to an economy is the focus of research, a modeller could simply create *ex ante* and *ex post* equilibria from data and would have no need to model the transition from one to the other explicitly. For a set of historical events whose net outcome is known, the more interesting question is the relative contribution made by each of the component events, that is, decomposing the net change into its constituent causes.

Although backwards and forwards simulations are equivalent for finding the net effects of simultaneous shocks, the contribution of an individual shock to change is, in general, conditional on the direction of simulation. This asymmetry arises because

[45] Because only relative prices matter, modellers use one price or a price index as a numeraire. This equivalence between forward and backward simulations will only hold if the choice of numeraire is constant.

forward and backward simulations differ in the status of the other shocks to the system: in a forward simulation no other shocks are present, whereas in a backward simulation all other shocks are present. Furthermore, the individual shocks to an economy are unlikely to have additive effects, so that the status of the remaining shocks in the system will affect the marginal contribution to change made by the shock of interest.

The shocks may be mutually enhancing so that for example, a specific technological innovation may make a greater contribution to growth in an economy with large labour force than in an economy with a small labour force. Conversely, they may dampen each other's effects. Thus, an exercise, such as that undertaken in Siriwardana (1995), in which the cumulative effect of all the shocks to the economy is inferred from summing the individual effects, ignores the potential for interactions or synergies among the shocks.

Some historians have implicitly recognized the distinction between analyses conducted in the presence and absence of other shocks. O'Rourke (1991), for example, models the Irish economy using data from the period 1840-1845 and introduces the agricultural price shocks which occurred between 1845 and 1876 alone to observe whether they can account for the observed decrease in agricultural labour demand. Because they cannot, he concludes that the famine was a significant factor in the Irish population decline. The simulation is repeated using the price shocks in conjunction with observed capital accumulation and both sets of results are reported. In both cases, the inability of the simulations to reproduce actual observations leads O'Rourke to conclude that the Famine was important. Thus, in this modelling exercise, the importance of the Famine is assumed by residual.

The assumptions underlying O'Rourke's simulations bear closer scrutiny. First, the implicit assumption is that the Irish economy faced three main shocks following the period 1840-1845: price shocks, capital accumulation, and the Famine. If additional shocks were present, then the conclusion that the Famine mattered could not be made; the source of the mismatch between the simulation outcome and observations could not be identified. Second, although no numerical claims are made for the impact of the Famine, the qualitative conclusion that it was important is conditional on the simultaneous presence of price shocks and capital accumulation.[46]

5.2.4 Calibration for Decomposition Analysis

Decomposition analysis provides a systematic way of explicitly separating the individual and the interactive effects of simultaneous shocks. Unlike O'Rourke's simulations, it requires all the significant mechanisms of change to be modelled and to be consistent with both an initial and a final equilibrium. The information available to economic historians about the net result of simultaneous changes is used to find values for unknown parameters. Typically not all of the values for the initial and final shock/policy parameters are known. The absence of information about the full set of policy shocks and the new equilibrium makes forward simulation the only option available to policy modellers. For economic historians, however, knowledge of the *ex post* equilibrium can be used to find unknown values for some of the shock or policy

[46] If capital accumulation was a factor of change, the conclusions from simulations with price shocks only should be that the effect of the combination of Famine and capital accumulation was important to the Irish economy.

parameters. While some of the values for these parameters, such as tax rates or quota allotments, will be known since they will be available from published data, others, such as technological change parameters, are seldom available from data and must be calibrated.[47] Calibration in this context uses information about the second equilibrium, to find values for the unknown policy or shock parameters.

Calibration for decomposition analysis becomes a two stage process. In the first stage, the initial equilibrium is used to find values for those parameters that remain constant in the initial and final equilibria - the time invariant parameters. Let $\mathbf{Y}^*$ denote the initial equilibrium, and let $\mathbf{p}^*$ be the known, initial period values of the policy and shock parameters. If F is the system of equations that characterises the model, the first stage in the calibration is the standard procedure given in Chapter 2. It finds values for the vector of time-invariant parameters, $\boldsymbol{\alpha}$, such that

$$F(\boldsymbol{\alpha}, \mathbf{p}^*) = \mathbf{Y}^*. \tag{5.1}$$

The second stage finds values for the unknown second period policy or shock parameters. Let $\mathbf{Y}^{**}$ denote the second period equilibrium values, let $\boldsymbol{\theta}^{**}$ be the vector of exogenously specified second period policy parameter values which are known from data, and let $\boldsymbol{\lambda}^{**}$ be the unknown second period policy or shock parameters. The second stage of calibration finds values for $\boldsymbol{\lambda}^{**}$ such that when the model is solved using those

[47] Harley and Crafts (1998), however, do not calibrate technology parameters from the 1770 benchmark data, but instead use exogenously specified values.

calibrated values and the values for θ^{**}, the second equilibrium is returned as the model solution:

$$F(\alpha, \lambda^{**}, \theta^{**}) = \mathbf{Y}^{**}. \tag{5.2}$$

As with the calibration of the static α parameters, the calibration of the unknown transition parameters assumes that modelled economy represents a deterministic system: once the elements of θ^{**} have been exogenously specified, the values of the calibrated transitional parameters, λ^{**}, absorb all of the residual changes in the economy. Such an assumption would be reasonable for a historical problem in which the main determinants of change are known.[48]

This use of the second observation for the calibration of applied general equilibrium models is rare. One recent example, which does not fall under the traditional domain of economic history, is Hill (1995) who isolates the injury to the Canadian economy, measured by employment changes, caused by changes in the world price of imports. In the period 1972 -1980, the Canadian economy was subjected to changing tax rates, factor endowments, preferences and technology as well as world price shocks. Hill specifies changes to tax rates, factor endowments and world prices from statistical publications, but calibrates preference and technical change parameters

[48] An alternative approach would be to specify *a priori* the number of transitional parameters in the model and then to calibrate the model to both equilibria under some least squares minimization criterion. Implicit in such an approach is the idea that the two equilibria are not true equilibrium representations of the economy.

so that when all the shocks are introduced to the 1972 economy, the 1980 benchmark data are reproduced as a solution to the system. In the counterfactual simulation, Hill allows all the changes to take place, but fixes the world price of industry output so that the share of domestic goods in domestic consumption remains at 1972 levels. The impact of the trade shocks is obtained from comparing the actual trade shock inclusive data to the counterfactual, trade shock free solution.

Although economic historians have not used a second observation to infer values for the parameters of their models, they have made use of information about the actual economic changes to test the performance of their model. This testing is undertaken by introducing all the known shocks to the model, solving, and comparing the model solution to the post-change observation. Typically, no formal testing criterion is applied. The modeller subjectively compares the magnitudes and directions of change arising from the model and asserts whether or not they are consistent with observed changes.

Hence, O'Rourke (1994), who explores the effects of repealing the Corn Laws on Irish agriculture, calibrates a model to data for the period 1856-1860, and in an initial exercise, exogenously introduces the actual wage and price shocks for the period 1956-1860 to 1874-76, together with a shock for the decline in the demand for potatoes. The simulation outcomes are compared to observed changes, and upon inspection, are deemed sufficiently close that the model can be used as an accurate representation of the economic process. The counterfactual is then undertaken by simulating the wage and potato demand shocks as before, but without the Repeal-based price shocks.

O'Rourke concludes from these simulations that the Repeal led to a large reduction in agricultural labour demand in Ireland.

Thus, O'Rourke and others who employ applied general equilibrium models to explore problems in economic history, have exploited their information about *ex post* equilibria to test their models, but they have not made full use of this information in their model specification. Decomposition analysis is a methodology that does make use of this information. It offers the possibility for a richer systematic understanding of the causes of economic change, especially since the ways in which simultaneous shocks interact with one another has remained unexplored in economics. Knowledge of a second equilibrium opens this topic to investigation by modellers.

5.3 Implementing Decomposition Analysis: US Railroads in the 19th Century

The use of applied general equilibrium models for decomposition analysis is illustrated here using a simple closed economy model of railroads in the United States during the penultimate decades of the 19th century. The contribution of the railways to US growth has long been the subject of debate in the economic history literature (see, for example, Fogel, 1995; and Williamson, 1974). The objective here is neither to dispute nor to embellish upon the results of this earlier work, but instead, to introduce model-based decomposition as a numerical tool for historians.

5.3.1 Model Structure

The illustrative model used to analyze the impact of technological change in the railroad sector in the US has a simple structure, which is summarized in Table 5.1. A single, representative agent with Cobb-Douglas preferences consumes three classes of goods: passenger rail services, agricultural goods and an aggregate residual good which is comprised of all other goods, services and savings in the economy. This agent's income is derived from the endowments of capital, labour and land. Production occurs in three sectors; railway transport, agriculture and the aggregate sector. Output from the railway sector is assumed to be produced from capital and labour with a constant elasticity of substitution (CES) production technology.

The railway sector supplies passenger services and freight services using a constant elasticity of transformation technology with a high elasticity of transformation (2.0) to reflect the relative ease of shifting production between the two types of services. Freight services are intermediate inputs into the production of the agricultural and the aggregate goods, while passenger services are produced for final consumption.

The agricultural sector is assumed to employ land, labour and freight services using CES technology with a low elasticity of substitution (0.15). The choice of fixed coefficients technology in agriculture was rejected on the basis that farmers likely had some control over the destination of their outputs and the origin of inputs and thus could vary the proportion of transportation per unit output, and that a low degree of substitution between quantity of land and labour intensity in agricultural production seems reasonable.

Table 5.1

Summary of the Model Structure Used to Illustrate Decomposition Analysis

Production Structure

- Three production sectors - the railroad sector, the agricultural sector, and an aggregate residual sector - produce final consumption goods.

- The production of agricultural goods and the aggregate good use railroad services (freight) as an intermediate input.

- <u>The railroad sector</u> produces railroad services using CES technology and capital and labour inputs, allowing for the possibility of sector and factor specific technological change:
  $Q^r = \zeta^r[\eta^r(\tau_K{}^r K^r)^{(\xi-1)/\xi} + (1-\eta^r)(\tau_L{}^r L^r)^{(\xi-1)/\xi}]^{\xi/(\xi-1)}$ where Q^r is the quantity of railroad services produced, K^r and L^r are the capital and labour inputs to the railroad sector, $\tau_K{}^r$ and $\tau_L{}^r$ are the technological change parameters for capital and labour inputs for the sector, ζ^r is a scale parameter, η^r is the CES share parameter for capital in the railroad good sector and ξ is the constant elasticity of substitution

- <u>Output from the railroad sector</u> is allocated between passenger and freight services in a CET transformation function, again allowing for sector and factor specific technical change:
  $Q^r = \gamma^r[\alpha^r(t_F{}^r F)^{(\chi-1)/\chi} + (1-\alpha^r)(t_P{}^r P)^{(\chi-1)/\chi}]^{\chi/(\chi-1)}$ where Q^r is the quantity of total railroad services produced, F and P are the quantities of freight and passenger services, $t_F{}^r$ and $t_P{}^r$ are the technological change parameters for freight and passenger outputs for the sector, γ^r is a scale parameter, α^r is the share parameter for freight services, and χ is the constant elasticity of transformation.

- <u>The aggregate good sector</u> produces output using a CES technology with freight services and value added as inputs, again allowing for sector and factor specific technological change:
  $Q^a = \gamma^a(\alpha^a (t_F{}^a F^a)^{(\psi-1)/\psi} + (1-\alpha^a)(t_V{}^a V^a)^{(\psi-1)/\psi})^{\psi/(\psi-1)}$, where Q^a is the output of the aggregate good sector, F^a and V^a are the freight and value added inputs, $t_F{}^a$ and $t_V{}^a$ are the technological change parameters for freight and value added inputs for the sector, γ^a is a scale parameter, α^a is the share parameter for the freight input, and ψ is the constant elasticity of substitution.

- <u>Value added in the aggregate good sector</u> is produced using CES technology from capital and labour, and allowing for the possibility of sector and factor specific technological change:
  $V^a = \zeta^a [\eta^a(\tau_K{}^a K^a)^{(\phi-1)/\phi} + (1-\eta^a)(\tau_L{}^a L^a)^{(\phi-1)/\phi}]^{\phi/(\phi-1)}$ where V^a is the quantity of value added produced, K^a and L^a are the capital and labour inputs to the aggregate good sector, $\tau_K{}^a$ and $\tau_L{}^a$ are the technological change parameters for capital and labour inputs for the sector, ζ^a is a scale parameter and η^a is the CES share parameter for capital in the aggregate good sector. The parameter ϕ is the constant elasticity of substitution.

- <u>The agriculture sector</u> produces agricultural output using CES technology and land, labour, and freight inputs, allowing for the possibility of sector and factor specific technological change:
  $Q^f = \zeta^f[\lambda_G{}^f(\tau_G{}^f G^f)^{(\sigma-1)/\sigma} + \lambda_L{}^f(\tau_L{}^f L^f)^{(\sigma-1)/\sigma} + \lambda_F{}^f (\tau_F{}^f F^f)^{(\sigma-1)/\sigma}]^{\sigma/(\sigma-1)}$ where Q^f is the quantity of agriculture produced, G^f, L^f and F^f are the land, labour and freight inputs to the agriculture sector, $\tau_G{}^f$, $\tau_L{}^f$ and $\tau_F{}^f$ are the technological change parameters for land, labour and freight inputs to the sector, ζ^f is a scale parameter, $\lambda_G{}^f$ is the CES share parameter for land in the agriculture sector, $\lambda_L{}^f$ is the CES share parameter for labour in the agriculture sector, $\lambda_F{}^f$ is the CES share parameter for freight in the agriculture sector, and σ is the constant elasticity of substitution.

Consumption Structure

- The single consumer has Cobb-Douglas preferences over agricultural goods, the aggregate good, and passenger rail services and maximizes the utility function:
  $U = (Q^f)^{\beta}(Q^a)^{\delta}P^{(1-\beta-\delta)}$ where U is the level of utility, β is the consumer's expenditure share on the agricultural good, δ is the consumer's expenditure on the aggregate good, Q^f, Q^a, and P are quantities consumed of the agricultural good, the aggregate good and passenger services.

- The consumer is endowed with capital (κ), labour (Γ), and land (G).

Production of the aggregate good is undertaken from freight services and value added, again assuming CES technology with a low elasticity of substitution (0.25). As in agriculture, the producer is assumed to have some choice over the destination of output and, hence, over the quantity of freight services required per unit output. A large increase in fares, however, would be required for him or her to substitute away from transportation in production. Since the aggregate good is a residual good that encompasses all non-agricultural, non-railway goods and services in the economy, more substitutability between freight and value added input in production is assumed than in agricultural production. Value added in the production of the aggregate good is undertaken using capital and labour inputs with CES technology and an elasticity of 1.1. In equilibrium, the markets for all goods and factors identified in the model clear, each production sector makes zero profits and the consumer's budget is balanced.

The model uses a very strong closed economy assumption. Given the importance of trade (around 15 percent of GDP) and the increase in the trade between the US and Europe in the latter decades of the 19th century, this specification is clearly ahistorical, but it was chosen for transparency in presenting the decomposition analysis methodology.

5.3.2 Calibration

The two equilibria in the model are derived from 1870 and 1890 data. The choice of years is somewhat arbitrary since the majority of railway construction in the US

occurred in the 1850s and 1860s. The period of major railway construction, however, was before the American Civil War, while its long run benefits were felt more towards the turn of the century. The presence of the war during in this interval represents a large structural shift which applied general equilibrium models are ill-equipped to handle. By 1870, however, much of the economic disruption caused by the war would have ended and hence, it is chosen as the base year for the simulations. The choice of 1890 for the second data set represents the passage of sufficient time for the modelled changes to occur, but not so long that a large number of unmodelled structural changes are likely to have occurred.

One of the pragmatic features of general equilibrium modelling is its relatively few data requirements. The model of the US economy specified in Table 5.1 can be calibrated using the raw data for 1870 and 1890 given in Section I of Table 5.2. These data, taken from four sources, represent the minimum data requirement for the model calibration.

Calibration for decomposition analysis is a more lengthy process than the traditional calibration employed for counterfactual analysis, and thus the process is explained here in detail. It entails four steps: the derivation of microconsistent benchmark data sets; the decomposition of the final period data set into equilibrium price and quantity observations; calibration from the initial benchmark data set of the time invariant parameters - those that are constant in 1870 and 1890; and calibration of the shock parameters, in this case the final period technology and preferences.

Table 5.2

Raw Data Used in Constructing the Benchmark Data Sets

I. Raw Data. The benchmark equilibrium data sets for the US economy in 1870 and 1890 were constructed from the following raw data.

	1870	1890	source	notes
GNP in 1929 prices	$10334M	$23143M	Kuznets. p. 118	Values are annual averages for 1869-78 and 1889-98.
Proportion of National Income in Agriculture	20.1%	17.1%	US Bureau of Census p.238	
Passenger Miles by Rail	4.1B	12.1B	Fishlow p. 585.	
Ton-Miles by Rail	11.7B	80.0	Fishlow p. 585.	
Passenger Rate in current prices	$0.00280	$0.00220	Fishlow p. 585.	
Freight Rate	$0.00218	$0.00092	Fishlow p. 585.	
Real Net Capital Stock in Railroad Sector Index 1909=100	16.6	61.9	Fishlow p. 606.	Index is for 1869 and 1889 respectively.
Total Labour Force	12930T	23320T	Lebergott p. 118.	
Labour Employed in Railway Transport	160T	750T	Lebergott p. 118.	
Labour Employed in Agriculture	6790T	9960T	Lebergott p. 118.	
Annual Earning of Non-farm Labour in 1914 prices	$375	$519	US Bureau of Census p. 165.	
Monthly Earnings of Farm Labourers in current prices	$16.57	$13.93	US Bureau of Census p. 163.	Values are inclusive of board.
Land in Farms (acres)	407735T	623219T	US Bureau of Census p. 457.	
Consumer Price Index 1958=100	38	27	US Bureau of Census p. 211.	Index for 1929; 51.3; Index for 1914;30.1.
Farm Output Index 1947-49 =100	23	43	US Bureau of Census p. 429.	

II. Constant Value Calculations. The following calculations were undertaken to transform the data into constant 1914 dollars.

	1870	1890	calculation
GNP in 1914 prices	$6063M	$13579M	GNP in 1929 prices x (0.301/0.513).
Value of Passenger Rail Services in current prices	$114.8M	$266.2M	Passenger miles x passenger rate.
Value of Rail Freight in current prices	$255.06M	$736M	Ton miles x freight rate.
Value of Passenger Rail Services in 1914 prices	$91M	$297M	1870: current value x (0.301/0.38); 1890: current value x (0.301/0.27).
Value of Rail Freight in 1914 prices	$202M	$821M	1870: current value x (0.301/0.38); 1890: current value x (0.301/0.27).
Annual earning of Non-farm Labour in 1914 prices	$375	$519	none.
Monthly Earnings of Farm Labourers in constant prices	$13.13	$15.53	1870: current value x (0.301/0.38); 1890: current value x (0.301/0.27).

The number of workers in the production of the aggregate consumer good is derived as the residual of the total labour force after agricultural labour and railway transport labour. In 1870 this value is 5980T and in 1890 it is 12610T.

Step 1. Derivation of the microconsistent data sets.

Microconsistent data sets are necessary to derive the pre-change (1870) and the post-change (1890) equilibria for the US economy, but the raw data in Table 5.2 are not microconsistent. The first step in the calibration process is, therefore, to use the data in Table 5.2 to derive microconsistent data sets for the initial and final equilibria in constant units. The benchmark data sets must both be specified in value terms and thus constant 1914 dollars are chosen as the common units. The transformation of the raw data for 1870 and 1890 into constant 1914 dollars is undertaken in the straightforward steps outlined in Section II of Table 5.2.

These constant unit series values are then used to derive the two microconsistent data sets in the manner described in Table 5.3. Some of the values in Table 5.3 are generated by applying the equilibrium assumptions of the model to the data. For example, the closed economy specification of the model implies that all agricultural production is consumed domestically - a clearly ahistorical simplification. Similarly, without domestic savings, the budget balance conditions of the model allow the value of output from the aggregate good sector to be constructed as a residual from GNP once the values of output from agriculture and railway services have been determined. The assumption that capital and labour are the sole inputs to the railway sector, coupled with the equilibrium condition that each sector make zero profits in equilibrium means that the value of the capital input to the railway sector is derived as the residual once the value of labour has been determined. Capital, as defined in the model, thus includes physical capital, investment in the sector and operating profits.

Table 5.3

Construction of the Benchmark Data Sets for Decomposition Analysis

The BED data values in constant 1914 dollars were derived as follows.

	1870	1890	calculation
1. Final Consumption			
Final Consumption of Rail Services	$91M	$297M	Passengers are assumed to be final consumers of rail transport.
Final Consumption of Agriculture	$1243M	$2322M	Proportion of Agriculture in Income multiplied by GNP in 1914 dollars.
Final Consumption of Aggregate Good	$4729M	$10960M	Residual of GNP after Railway and Agricultural Consumption.
2. Railway Sector			
Value of Total Railway Transport	$293M	$1118M	Sum of passenger and freight values.
Value of Labour in Railway Sector	$60M	$389M	Labour employed in Railway sector multiplied by average annual earnings.
Value of Capital in Railway Sector	$233M	$729M	Residual of Output Value after labour costs.
3. Agriculture Sector			
Value of Labour in Agriculture	$1070M	$1856M	Labour Employed in Agriculture multiplied by the monthly wage multiplied by 12.
Value of Freight Input to Agricultural Sector	$42M	$144M	Freight Transport assumed equal to Proportion of Agriculture in total non-rail output.
Value of Agricultural Land Input	$131M	$322M	Residual of Agricultural Output after labour and freight costs.
4. Aggregate Good Sector			
Value of Labour in Aggregate Good Production	$2243M	$5558M	Number of workers multiplied by average annual earnings.
Value of Freight Input to Aggregate Good Sector	$160M	$677M	Freight Transport assumed equal to Proportion of Aggregate Good in total non-rail output.
Value of Capital Input to Aggregate Good Sector	$2326M	$4725M	Residual of Aggregate Good Output after Labour and Freight costs.
5. Endowments			
Labour	$3373M	$7803M	Sum of labour inputs to production sectors.
Capital	$2559M	$5454M	Sum of capital inputs to production sectors.
Land	$131M	$322M	Value of land input to agriculture.

Step 2. Derivation of prices and quantities for the final equilibrium.

By adopting the Harberger (1962) convention, the quantities transacted in the initial equilibrium are defined in units costing one 1914 dollar. As a result, the 1870 benchmark data set can represent both the value of transactions and the model-defined quantities transacted in 1870. The microconsistency of the 1870 data set also implies that the market clearing equilibrium conditions for goods and factors hold.

However, no such parity between quantities and transactions values exists in the 1890 benchmark data set. Although the values of transactions must be microconsistent to meet the zero profit and budget balance conditions of the model, market clearing is not evident from the final data set because prices are no longer unity; shocks to the economy had occurred in the interim which changed the relative prices of the goods and factors in the economy. To ensure that market clearing holds, the transactions values in the final equilibrium must be decomposed into quantity and price observations.

The process is further complicated by the requirement that the quantities in the 1890 benchmark equilibrium be measured in model-consistent units; so that published data measured in tons, numbers of individuals, acres, ton-miles, number of locomotives, etc., cannot be used. Instead, the published data are used to create indices of changes in the quantities produced or transacted, and these indices are then applied to the model-consistent base quantities.

The quantities in the model are found, therefore, by using the ratio in non-model consistent units of a quantity produced in 1890 to the quantity produced in 1870 and multiplying the 1870 model-consistent quantity by that ratio. For example, the raw data in Table 5.2 indicate that the quantity of land in farms was 623 million acres in 1890

and 408 million acres in 1870. From Table 5.3, the quantity of land in agriculture was 131 million model-consistent units in 1870, so that the model-consistent quantity of land in agriculture in 1890 is given by 131 million multiplied by 623/408. The price of land is then found by dividing the value of land in agriculture by its quantity. Similarly, the quantity of the labour endowment and the price of labour are derived from data about the relative size of the labour force in the two years of interest. The quantities of labour inputs to the various sectors are inferred by dividing the value of the labour input in each sector by the price of labour.

Table 5.4 gives a detailed description of how the decomposition of the 1890 benchmark transactions values was undertaken, while Table 5.5 summarises the ensuing prices and quantities. The weakest elements in this derivation are likely to be the price and quantities for capital. As was noted earlier, the definition of capital in the model includes all inputs except land and labour. To find the quantity of capital input to the railroad sector, indices of the real net capital stock in the railroad sector were used. This index is based on physical capital stock and is employed as a proxy to the model definition of capital stock. Using this index then requires the implicit and probably unrealistic assumption that the ratio of physical capital stock used in the index (locomotives and track) to the broader notion of capital stock in the railway sector defined in the model (locomotives and track, operating profits, investment) remained constant in the interval 1870-1890. The quantity of capital in the railway sector, thus derived, is then used to infer the price of capital. Because capital is homogenous in the model, this price of capital then forms the basis for the quantity observation of the

Table 5.4

Derivation of Quantities and Prices Used to Calibrate Technology Parameters

1890 Variable Name	Derivation in Model-Consistent Units	Calculation	Value
1. Labour			
Quantity of labour endowment	Ratio of the total labour force in 1890 to the total labour force in 1870 multiplied by the 1870 labour endowment.	(23320T/12930T) x 3373M	6083M
Price of labour	Value of labour in 1890 divided by the labour endowment quantity.	$7803M / $6083M	$1.2828
Quantity of labour input to the railway sector	Value of the labour input to the railway sector in 1890 divided by the price of labour in 1890.	$380M / $1.2828	303M
Quantity of labour input to agriculture	Value of the labour input to agriculture in 1890 divided by the price of labour in 1890.	$1856M / $1.2828	1447M
Quantity of labour input to the aggregate good sector	Value of the labour input to the aggregate good sector in 1890 divided by the price of labour in 1890.	$5558M / $1.2828	4333M
2. Capital			
Quantity of capital input to the railway sector	Ratio of the index of real net capital stock in the railway sector in 1890 to the index of real net capital stock in the railway sector in 1870 multiplied by the 1870 quantity of capital input to the railway sector.	(61.9 / 16.6) x 233M	869M
Price of capital	Value of capital input to the railway sector in 1890 divided by the quantity of capital input in 1890.	$729M / 826M	$0.8391
Quantity of capital endowment	Value of the capital endowment in 1890 divided by the price of capital in 1890.	$5454M / $0.8391	6500M
Quantity of capital input to the aggregate good sector	Value of capital input to the aggregate good sector in 1890 divided by the price of capital in 1890.	$4725M / $0.8391	5631M
3. Land			
Quantity of land endowment	Ratio of acres of farmland in 1890 to acres of farmland in 1870 multiplied by the 1870 quantity of land endowment.	(623219T / 407735T) x 131M	200M
Price of land	Value of the land endowment in 1890 divided by the quantity of the land endowment in 1890.	$322M / 200M	$1.6081
4. Final Consumption			
Quantity of passenger services consumed	Ratio of passenger miles in 1890 to passenger miles in 1870 multiplied by the 1870 quantity of passenger services.	(12.1B / 4.1B) x 91M	269M
Price of passenger services	Value of passenger services in 1890 divided by the quantity of passenger services in 1890.	$297M / 269M	$1.1059
Quantity of agriculture consumed	Ratio of the farm output index for 1890 to the farm output index for 1870 multiplied by the quantity of agricultural consumption for 1870.	(43 / 23) x 1243M	2324M
Price of agriculture	Value of agriculture in 1890 divided by the quantity of agriculture in 1890.	$2322M / 2324M	$0.9992
Price of the aggregate good	Price implied by the price index: Value of the 1870 consumption of the aggregate good at 1890 prices, divided by the quantity of 1870 aggregate good consumption.	($6063M - (1243M x $0.9992) - (91M x $1.1059)) / 4729M	$0.9982
Quantity of the aggregate good consumed	Value of the 1890 aggregate good consumption divided by the 1890 price of the aggregate good.	$10960M / $0.9982	10980M
5. Freight			
Quantity of freight produced	Ratio of ton-miles in 1890 to the ton-miles in 1870 multiplied by the 1870 output of freight.	(80 / 11.7) x 202M	1381M
Price of freight	Value of freight produced in 1890 divided by the quantity of freight produced in 1890.	$821M / 1381M	$0.5944
Quantity of freight input to agriculture	Value of the freight input to agriculture divided in 1890 by the price of freight in 1890.	$144M / $0.5944	242M
Quantity of freight input to the aggregate good sector	Value of freight input to the aggregate good sector in 1890 divided by the price of freight in 1890.	$677M / $0.5944	1139M

Table 5.5

Quantities and Prices for Calibration

	Quantities (millions of model-consistent units)		Prices (constant 1914 dollars)	
	1870	**1890**	**1870**	**1890**
1. Final Consumption				
Final Consumption of Passenger Services	91	269	1.0000	1.1059
Final Consumption of Agriculture	1243	2324	1.0000	0.9992
Final Consumption of Aggregate Good	4729	10980	1.0000	0.9982
2. Railway Sector				
Output of Passenger Services	91	269	1.0000	1.1059
Output of Freight Services	202	1381	1.0000	0.5944
Capital Input to Railway Sector	233	869	1.0000	0.8391
Labour Input to Railway Sector	60	303	1.0000	1.2828
3. Agriculture Sector				
Output of Agriculture	1243	2324	1.0000	0.9992
Labour Input to Agriculture	1070	1447	1.0000	1.2828
Land Input to Agriculture	131	200	1.0000	1.6081
Freight Input to Agriculture	42	242	1.0000	0.5944
4. Aggregate Good Sector				
Output of Aggregate Good	4729	10980	1.0000	0.9982
Labour Input to Aggregate Good Sector	2243	4333	1.0000	1.2828
Capital Input to Aggregate Good Sector	2326	5631	1.0000	0.8391
Freight Input to Aggregate Good Sector	160	1139	1.0000	0.5944
5. Endowments				
Labour	3373	6083	1.0000	1.2828
Capital	2559	6500	1.0000	0.8391
Land	131	200	1.0000	1.6081

capital input to the aggregate good sector.

The derivation of the price of the aggregate good also warrants comment. The benchmark data sets for 1870 and 1890 are constructed to be in common units of constant 1914 dollars. Under the constant dollar condition, the consumer price index for both data sets is the same. Hence, the consumer's bundle of goods in the 1870 data set can be bought for the same expenditure in 1890. This condition allows the price of the aggregate good in 1890 to be inferred from the prices of the agricultural good, and passenger services in 1890, GNP in 1870, and the quantities consumed in 1870.

Step 3 ***Calibration of the initial period parameters.***

The calibration of the initial period parameters is undertaken using the 1870 benchmark equilibrium data set only. It employs the standard procedure described in Mansur and Whalley (1984), Shoven and Whalley (1992) and summarized in Chapter 2; sufficient parameters are specified exogenously that the remainder can be exactly determined from the 1870 benchmark data set. In this case, the exogenously specified parameters are the elasticities of substitution in the CES and CET functions, the values of which are given in Table 5.6; the initial period factor endowments, which are obtained from the benchmark data set for 1870; and the specification of the base period technology parameters (τ_K^r, τ_L^r, τ_K^a, τ_L^a, τ_G^f, τ_K^f, τ_L^f, t_F^r, t_P^r, t_F^a, and t_V^a from Table 5.1) which are set to unity.

The base case calibrated parameters include the share and scale parameters in the model's CES production functions and the Cobb-Douglas consumer expenditure

Table 5.6

Production Parameter Values

Parameter	Symbol used in Table 5.1	Value
1. Exogenously Specified Parameters		
CES substitution elasticity (capital, labour) in railway services production	ξ	0.8
CET transformation elasticity (passenger, freight services) in railway output	χ	-2.0
CES substitution elasticity(land, labour, freight) in agriculture production	σ	0.15
CES substitution elasticity(capital, labour) in values added production	ϕ	1.1
CES substitution elasticity (value added, freight) in aggregate good production	ψ	0.25
2. Calibrated Parameters		
CES share of capital in the railway sector	η^r	0.845
CES share of labour in the railway sector	$1-\eta^r$	0.155
Scale parameter for railway services production	ζ^r	1.603
CET share of freight output in the railway sector	α^r	0.402
CET share of passenger services output in the railway sector	$1-\alpha^r$	0.598
Scale parameter for the output railway services	γ^r	2.080
CES share of freight in the agriculture sector	λ_f^f	<0.001
CES share of labour in the agriculture sector	λ_l^f	<0.001
CES share of land in the agriculture sector	λ_G^f	≈1.000
Scale parameter in agriculture	ζ^f	1.193
CES share of labour in value added	$1-\eta^a$	0.492
CES share of capital in value added	η^a	0.508
Scale parameter in the production of value added	ζ^a	2.000
CES share of value added in the production of the aggregate good	$1-\alpha^a$	≈1.000
CES share of freight in the production of the aggregate good	α^a	<0.001
Scale parameter in the production of aggregate good	γ^a	1.047

shares. The values of the production parameters are summarized in Table 5.6, while the preference shares in the initial equilibrium are given in Table 5.7.

Step 4 ***Calibration of the final period parameters.***

For the illustrative purposes of this exercise, four types of shocks are assumed to have affected the US economy in the interval 1870 to 1890; increases in factor endowments, changes to preferences, technological change in the railroad sector, and other technological change. Each of the four types of shock must be represented as changes to model parameters, and these changes are summarised in Table 5.7. The increases in factor endowments are taken directly from the 1890 quantity observations in Table 5.5. Similarly, the values for the Cobb-Douglas preference shares in 1890 are derived from the expenditure data in the 1890 benchmark data set in the same way that they are calculated in the base case: they are determined from the 1890 values of the consumer's income (given by the sum of the value of endowments in Table 5.3), and the quantities and prices of goods consumed in 1890 (given in Table 5.5).

The calibration of the technological change parameters, however, represents an extension to the standard, single period calibration technique. In a model without technological change, a standard CES production function is given by

$$Y = \theta(\textstyle\sum_i \omega_i X_i^{(\rho-1/\rho)})^{\rho/(\rho-1)}, \qquad \textbf{(5.3)}$$

where Y is the quantity produced, ρ is the elasticity of substitution, θ is a scale

Table 5.7

Parameter Value Changes in the Model

Parameter	1870 Value	1890 Value
1. Railroad Technology Shocks		
Technological Growth in Railroad Production	1	1.778
Technological Growth in Railroad Output	1	1.000
2. Other Technology Shocks		
Technological Growth in Agriculture	1	1.401
Technological Growth in Aggregate Good Production	1	1.010
Technological Growth in Value Added Production	1	1.086
3. Factor Endowment Shocks		
Labour endowment (in model-consistent units)	3373M	6083M
Land endowment (in model-consistent units)	131M	200M
Capital endowment (in model-consistent units)	2559M	6500M
4. Preference Shocks		
Cobb-Douglas expenditure share on the agricultural good	0.205	0.171
Cobb-Douglas expenditure share on the aggregate good	0.780	0.807
Cobb-Douglas expenditure share on passenger services	0.015	0.022

parameter, X_i is the quantity used of input i, and ω_i is the CES share of input i such that $\sum_i \omega_i = 1$. If input prices are denoted by w_i, the corresponding unit cost, C, is given by

$$C = (1/\theta)(\textstyle\sum_i \omega_i^{\rho} w_i^{(1-\rho)})^{-1/(\rho-1)}, \tag{5.4}$$

and the cost-minimizing demand for factor i, x_i, by

$$x_i = Y(\omega_i C / w_i)^{\rho}\, \theta^{(\rho-1)}. \tag{5.5}$$

Technological change is introduced to this model using the specification introduced in Hill (1985). This specification modifies the standard CES function of equation (5.3) using technological change parameters for each factor, z_i, so that the production function becomes

$$Y = \theta(\textstyle\sum_i \omega_i (z_i X_i)^{(\rho-1/\rho)})^{\rho/(\rho-1)}. \tag{5.6}$$

The CES functions of the model presented in Table 5.1 employ this specification for technology. The corresponding unit cost and factor demand functions are given by

$$C = (1/\theta)(\textstyle\sum_i \omega_i^{\rho} z_i^{(\rho-1)} w_i^{(1-\rho)})^{-1/(\rho-1)}, \tag{5.7}$$

and

$$x_i = Y(\omega_i z_i^{(\rho-1)/\rho} C / w_i)^{\rho} \theta^{(\rho-1)}. \tag{5.8}$$

Rearranging equation (5.8) allows the parameters z_i to be expressed as

$$z_i = (x_i \theta^{(1-\rho)}/Y)^{1/(\rho-1)} (w_i/\omega_i C)^{\rho/(\rho-1)}. \tag{5.9}$$

Thus modelled, technological change has a scaling effect on the quantity produced and an allocative effect on the relative intensity of the input use. These two effects can be illustrated by defining Ω,

$$\Omega = \sum_i \omega_i z^{(\rho-1/\rho)}, \tag{5.10}$$

and Z_i,

$$Z_i = \omega_i z_i^{(\rho-1/\rho)}/\Omega, \tag{5.11}$$

and rewriting equation (5.6) as

$$Y = \theta\, \Omega^{\rho/(\rho-1)} (\sum_i Z_i X_i^{(\rho-1/\rho)})^{\rho/(\rho-1)}. \tag{5.12}$$

If the change in the technology parameters has a positive effect on output, $\Omega^{\rho/(\rho-1)}$ will be greater than one. The changes in the factor intensities in production arising from the

new technology are captured by the new CES share parameters, Z_t.

In the initial equilibrium, the technology parameters are fixed at unity and the values for θ and ω_t are calibrated from the values in the base period benchmark equilibrium data set. Calibration of the second period technology parameters is then undertaken by substituting the second period price observation for C and quantity observations x_t and Y into equation (5.8) and solving. Because the values for θ and ω_t are derived for the case where the initial benchmark technology parameters are unity, the values of the final equilibrium technology parameters are measured relative to 1. The measure of technological growth in the sector is derived by substituting the technology parameters into equation (5.10). The values of the technological growth parameters in the model sectors described in Table 5.1, are given in Table 5.7. These values suggest that the greatest technological gains occur in the production of agriculture and railway services.

Although advances in agriculture and transportation technologies seem to be plausible features of late 19th century US economic development, the calibrated technology parameter values in Table 5.7 should be treated with some scepticism. Some of the criticisms that have been levelled against current practices in general equilibrium modelling are that the deterministic approach to calibration forces the model's parameters to absorb all the stochastic disturbances in the data, and that it offers no avenue for testing the choice of model specification (see McKitrick, 1995). This shortcoming of the traditional methodology persists in the strategy described here for calibrating technology parameters - once the exogenous shocks have been specified, technology absorbs the residual causes of change. For such a model to be compelling

and the decomposition analysis to be meaningful, the modeller must be confident that no significant causes of change have been omitted. In this case, however, the model has been constructed to illustrate a methodology rather than to contribute to the debate about the role of railroads in US growth. Thus, features of the US economy in the late 19th century which certainly have had a role in growth, such as trade in agriculture, have been omitted to maintain transparency. The parameter values in Table 5.7 are, therefore, derived for an ahistorical model.

5.4 Decomposition Analysis

Decomposition analysis is undertaken by solving the model repeatedly using exhaustive combinations of shocks, and comparing the ensuing solution values of a specific variable of interest. The number of shocks that form the basis of a given decomposition analysis, however, depends on the question to be addressed. In the model considered here, the focus of the analysis is the effect of technological change in the railroad sector on GNP relative to the effects of other technological change, the change in endowments, and the change in preferences. Hence, four types of shocks are identified for the decomposition analysis. Because some of the four shocks involve more than one parameter change, they could, in principle, be split into other component shocks. For example, a modeller who was interested in the relative effects of the increased labour supply and land expansion on growth might choose to categorize five shocks: shocks to each of the three factor endowments, all technological change and changes in preferences.

If the economy is categorized as being subjected to N shocks, the number of solutions required for decomposition analysis is (2^N-2).[49] Hence, the effects of the four types of shocks identified in Table 5.7 are determined from fourteen simulations. The results of these simulations are given in Table 5.8. Section 1 of Table 5.8 gives the results of simulations in which railroad technology is changed both by itself and in conjunction with other shocks to the economy. Section 2 considers the cases in which the railroad technology is fixed at its 1870 level and the remaining shocks are introduced to the system. A glance at Table 5.8 indicates that the greatest increase in GNP came from the growth in factor endowments. This result is hardly surprising, since the endowments grew by between 50 and 100 percent over the two decades.

The question of greater interest for decomposition analysis, however, is the contribution of each individual shock to growth. This contribution is presented in Table 5.9, which shows unambiguously that the effect of a specific shock depends on the presence or absence of the remaining shocks. Table 5.9 gives the marginal growth in GNP from introducing each individual shock, *conditional* on the status of the other three shocks.[50] In this case, the response to the question 'what was the contribution to GNP growth of the change in railroad technology?' has eight different answers. These

---

[49] The number of possible combinations of shocks is 2^N, but this number includes the case in which no shocks are introduced to the system (the initial benchmark) and the case in which all shocks are introduced (the final benchmark). Hence, the number of informative solutions required is 2^N-2.

[50] These values are derived from the results in Table 5.8. For example, Table 5.8 shows that introducing the change in factor endowments leads to GNP of \$12509M. Introducing the change in factor endowments *and* railroad technology leads to a GNP of \$12599M. Hence, the contribution of railroads to GNP, conditional on the presence of changes in factor endowments is (\$12599M-\$12509M) = \$90M.

Table 5.8

GNP in Millions of 1914 USD Arising from Changes in Parameter Configurations

1890 Parameters Used in Place of 1870 Parameters	GNP
None (1870 Base Case)	6063
1. Simulations with Railroad Technology Shocks	
Railroad Technology Only	6126
Railroad Technology and Preferences	6130
Railroad and Other Technology (all technology)	6614
Railroad Technology and Preferences and Other Technology	6609
Railroad Technology and Endowments	12599
Railroad Technology, Endowments and Preferences	12677
Railroad Technology and Endowments and Other Technology	13529
Railroad Technology, Preferences, Endowments and Other Technology (1890 case)	13579
2. Simulations with Other Shocks	
Preferences Only	6069
Preferences and Other Technology	6359
Preferences and Endowments	12589
Preferences, Endowments and Other Technology	13176
Other Technology Only	6362
Endowments Only	12509
Endowments and Other Technology	13129

Table 5.9

The Effects of Individual Shocks on US GNP between 1870 and 1890

(values are the increase in GNP in millions of constant 1914 USD from the 1870 value)

1. The Effect of Changes in Railway Technology

i)	Compared to the 1870 base case	63
ii)	Given other technology	252
iii)	Given preference changes	61
iv)	Given endowment changes	90
v)	Given other technology and preference changes	250
vi)	Given other technology and endowment changes	400
vii)	Given preference and endowment changes	88
viii)	Given other technology, preference and endowment changes (all other changes)	403

2. The Effect of Changes in Other Technology

i)	Compared to the 1870 base case	299
ii)	Given railroad technology	488
iii)	Given preference changes	290
iv)	Given endowment changes	620
v)	Given railroad technology and preference changes	479
vi)	Given railroad technology and endowment changes	930
vii)	Given preference and endowment changes	587
viii)	Given railroad technology, preference, endowment changes (all other changes)	902

3. The Effect of Changes in Endowments

i)	Compared to the 1870 base case	6446
ii)	Given other technology	6767
iii)	Given preference changes	6520
iv)	Given railroad technology changes	6473
v)	Given other technology and preference changes	6817
vi)	Given other technology and railroad technology changes	6915
vii)	Given preference and railroad technology changes	6547
viii)	Given other technology, preference, railroad tech. changes (all other changes)	6970

4. The Effect of Changes in Preferences

i)	Compared to the 1870 base case	6
ii)	Given other technology	-3
iii)	Given railroad technology changes	4
iv)	Given endowment changes	80
v)	Given other technology and railroad technology changes	-5
vi)	Given other technology and endowment changes	47
vii)	Given railroad technology and endowment changes	78
viii)	Given other technology, railroad technology, endowment changes (all other changes)	50

answers vary considerably. Given changes in preferences, the marginal contribution of the change in railroad technology to growth was 61 million 1914 USD, whereas its marginal contribution given all other changes in the economy was 403 million 1914 USD, (the value from a backward simulation). A forward simulation would yield an effect on GNP from changes in railroad technology of 63 million 1914 USD.

One way of simplifying the answer would be to report the average effect, but such a strategy hides some interesting results. For example, both the average effects of each of the shocks and the results presented in Table 5.8 seem to suggest that all four shocks led to growth in GNP. The conditional results in Table 5.9, however, show that in two cases, the change in preferences would have had a small but negative effect on growth.

The limitations of undertaking only a forward or a backward simulation to find the effects of a specific shock also emerge from Table 5.9. A forward simulation of this model would have suggested that only 0.8 percent of the growth in GNP between the years 1870 and 1890 was accounted for by technological change in the railroad sector, whereas a backward simulation would have concluded that this figure was 5.4 percent.

The results in Table 5.9 indicate that although the growth in GNP between 1870 and 1890 is $7516M, the sum of the individual effects of each shock relative to the base case is only $6814M. This discrepancy suggests that synergies exist between the different kinds of shocks. The four shocks clearly have a net mutually enhancing effect which accounts for a growth in GNP of $702M. The question then arises of how the shocks interact.

5.4.1 Interactions Between Shocks

The ways in which shocks are mutually enhancing or offsetting can be derived from decomposition simulations. Consider the case where the modeller is interested in the effects of introducing two shocks, A and B to an initial equilibrium. The effect of A alone is found by model simulations using just A, while the effect of B alone is found from simulations with just B. The synergy of A and B is given as the difference between the net combined effect of A and B and the sum of their individual effects. Thus, if R(A) denotes the change in the model result from solving the model using shock A, if R(B) denotes value arising from solving the model using shock B, and if R(A, B) is the value from solving the model using both shocks A and B, then the interaction of A and B, denoted here by I(A, B), can be expressed as

$$I(A, B) = R(A, B) - R(A) - R(B). \tag{5.13}$$

This index is derived for a forward simulation. A more general expression can be found for the synergies between two simultaneous shocks. Let A and B be indexed by 0 in the initial equilibrium and by 1 in the final equilibrium, and let R_0 be the result of a forward simulation. Equation (5.13) which gives the interaction index for the forward simulation, I_0, becomes

$$I_0(A, B) = R_0(A_1, B_1) - R_0(A_1|B_0) - R_0(B_1|A_0). \tag{5.14}$$

Alternatively, the interactivity index could be measured relative to the final equilibrium. If such an index is denoted by I_1, and if the results of backward simulations are denoted by R_1, its value is given by

$$I_1(A, B) = R_1(A_0, B_0) - R_1(A_0|B_1) - R_1(B_0|A_1). \qquad \textbf{(5.15)}$$

The derivation of pairwise interactivity indices becomes somewhat more complicated when a third shock is present. As with the individual effect of a shock, the value of the interactivity expression will be conditional on the value of the remaining shock. Consider a third shock, C, which is also indexed by 0 and 1 to denote its values in the initial equilibrium and in the final equilibrium. The forward interaction between A and B must then be expressed as conditional on the status of C, that is either

$$I_0(A, B|C_0) = R_0(A_1, B_1| C_0) - R_0(A_1|B_0, C_0) - R_0(B_1|A_0, C_0), \qquad \textbf{(5.16)}$$

or

$$I_0(A, B|C_1) = R_0(A_1, B_1| C_1) - R_0(A_1| B_0, C_1) - R_0(B_1|A_0, C_1). \qquad \textbf{(5.17)}$$

These two conditional indices will not necessarily be equivalent - the pairwise synergy between A and B depends on the presence or absence of C. If the modeller is interested in reporting a general measure of the interactivity between two specific shocks, an

unconditional interactivity index could be constructed by taking the average of the two conditional indices.

The existence of a third shock means that the interaction between A and B is only one of three possible pairwise synergies. Decomposition analysis also allows the modeller to report the conditional synergies for A and C, and for B and C. Furthermore, a third shock introduces the possibility of a three-way interaction. This interaction is derived as the residual after all of the individual and pairwise effects of the shocks relative to the initial equilibrium have been considered. Hence, if $R_0(A_1, B_1, C_1)$ is the result of introducing all shocks in a forward simulation, (which by definition is the difference between the initial and the final benchmark equilibria), then the three-way interactivity index would be given by

$$\begin{aligned} I_0(A, B, C) = \; & R_0(A_1, B_1, C_1) - R_0(A_1 | B_0, C_0) - R_0(B_1 | A_0, C_0) - R_0(C_1 | A_0, B_0) \\ & - I_0(A, B | C_0) - I_0(A, C | B_0) - I_0(B, C | A_0). \end{aligned} \quad \textbf{(5.18)}$$

Rearranging (5.18) gives

$$\begin{aligned} R_0(A_1, B_1, C_1) = \; & R_0(A_1 | B_0, C_0) + R_0(B_1 | A_0, C_0) + R_0(C_1 | A_0, B_0) \\ & + I_0(A, B | C_0) + I_0(A, C | B_0) + I_0(B, C | A_0) + I_0(A, B, C), \end{aligned} \quad \textbf{(5.19)}$$

which represents the full forward decomposition analysis for the case of three shocks to an economy. Thus, the difference between the two equilibria can be presented as the

sum of the individual contributions to change of the three shocks, their pairwise and three-way synergies.

Decomposition analysis can be generalized beyond the case of three shocks to the case in which a modelled economy faces a set of shocks $\{S\}$ with N elements. An exhaustive decomposition of the overall change between two time periods would include the effects of the 1-way, 2-way, ..., N-way interactions, where the 1-way interaction gives the individual effects of each of the N shocks, and the remaining interactions capture the spectrum of possible inter-shock synergies.

Let the set $\{S\}$ be partitioned into a non-empty set, $\{Y\}$, with y elements which are the shocks that form a y-way interaction, and $\{Z\}$, with $(N\text{-}y)$ elements which are the remaining shocks in $\{S\}$. The number of y-way interactions in the decomposition analysis is given by $N![(N\text{-}y)!y!]^{-1}$, so that the four-shock illustrative model considered here includes four individual effects, six pairwise synergies, four three-way synergies, and one four-way interaction. The total number of terms in the decomposition analysis, then, is

$$\sum_{y=1}^{N} N![(N\text{-}y)!y!]^{-1} = 2^{N}\text{-}1. \qquad \textbf{(5.20)}$$

Furthermore, each of the y-way synergies will be conditioned on the values of the elements in $\{Z\}$, and the number of possible conditional values for a given y-way interaction will be $2^{N\text{-}y}$. Hence, in the illustrative model, each single shock has the eight conditional effects on GNP given in Table 5.9. The six pairwise interactions each have

four conditional values, the three-way synergy has two, and the four-way interaction, one.

5.4.2 The Interactions Among Shocks in the Illustrative Model

The derivation of the four conditional synergies for each of the six pairs of shocks in the model is given in Table 5.10. The conditional joint effects of each pair of shocks, found in the first column of Table 5.10, are calculated from the results in Table 5.8. Hence, the net change in GNP attributed to the introduction of 1890 railroad technology and preferences, in the presence of the shock to other technology but the absence of the endowment shock (E=0, T=1) is given by the difference in GNP from introducing railroad technology, other technology and preferences (6609 million 1914 USD) and the effect of introducing other technology alone (6362 million 1914 USD). The second and third columns give the conditional individual effects of shocks and are taken from Table 5.9.

The pairwise synergies in Table 5.10 suggest that the endowment, and wider technology changes in the penultimate decades of the 19th century US economy had significant mutually enhancing effects, accounting for between 4 and 6 percent of the total change in GNP. Hence, any analysis of the contribution of either of these shocks to growth in the absence of the other would understate its importance. This effect is particularly true of the other technology shock where in some cases, the interactive effects with the endowment changes are greater than its individual effect. In contrast, the change in preferences over the interval offsets the effects of changes in the other

Table 5.10

Derivation of Pairwise Interactions for the Four Model Shocks

Railroad Technology (R) and Preference (P) Interaction: I(R, P)

		ΔGNP(R, P)	ΔGNP(R)	ΔGNP(P)	GNP Change due to R and P interaction
E=0	T=0	67	63	6	-2
E=1	T=0	168	90	80	-2
E=0	T=1	247	252	-3	-2
E=1	T=1	450	400	47	3
				mean =	-1

Railroad Technology (R) and Endowment (E) Interaction: I(R, E)

		ΔGNP(R, E)	ΔGNP(R)	ΔGNP(E)	GNP Change due to R and E interaction
P=0	T=0	6536	63	6446	27
P=1	T=0	6608	61	6520	27
P=0	T=1	7167	252	6767	148
P=1	T=1	7220	250	6817	153
				mean =	64

Railroad Technology (R) and Other Technology (T) Interaction: I(R, T)

		ΔGNP(R, T)	ΔGNP(R)	ΔGNP(T)	GNP Change due to R and T interaction
P=0	E=0	551	63	299	189
P=1	E=0	540	61	290	189
P=0	E=1	1020	90	620	310
P=1	E=1	990	88	587	315
				mean =	251

Preference (P) and Other Technology (T) Interaction: I(P, T)

		ΔGNP(P, T)	ΔGNP(P)	ΔGNP(T)	GNP Change due to P and T interaction
R=0	E=0	296	6	299	-9
R=1	E=0	483	4	488	-9
R=0	E=1	667	80	620	-33
R=1	E=1	980	78	930	-28
				mean =	-20

Preference (P) and Endowment (E) Interaction: I(P, E)

		ΔGNP(P, E)	ΔGNP(P)	ΔGNP(E)	GNP Change due to P and E interaction
R=0	T=0	6526	6	6446	74
R=1	T=0	6551	4	6473	74
R=0	T=1	6814	-3	6767	50
R=1	T=1	6965	-5	6915	55
				mean =	63

Endowment (E) and Other Technology (T) Interaction: I(E, T)

		ΔGNP(E, T)	ΔGNP(E)	ΔGNP(T)	GNP Change due to E and T interaction
R=0	P=0	7066	6446	299	321
R=1	P=0	7403	6473	488	442
R=0	P=1	7107	6520	290	297
R=1	P=1	7449	6547	479	423
				mean =	371

technology. This negative synergy, measured relative to the initial equilibrium, is greater in magnitude than the individual positive effect on GNP of changed preferences, resulting in a net negative effect.

The differences in conclusions about the synergies between shocks arising from forward and backward simulations also emerges from Table 5.10, particularly with respect to the changes in railroad technology. For each pair of shocks, forward simulations would generate the first of the four interactions presented, and backward simulations, the last. A forward simulation would lead the modeller to believe that railroad technology and preferences exhibited a slight negative synergy, while a backward simulation would suggest a slight positive synergy. Although the signs of the interaction for railroad technology and other technology shocks would be positive for both forward and backward simulations, a backward simulation would suggest that the magnitude of the synergy is more than five times its value from a forward simulation.

Table 5.11 presents the conditional three-way interaction values, and Table 5.12, the four-way value. These higher order indices are derived to complete the decomposition analysis, but are of little intrinsic interest, mainly because they are very small. About 91 percent of the change in GNP over the modelled interval can be attributed to individual shocks, while about 8 percent is captured by pairwise synergies. The combined effect of the three-way and four-way interactions is just over 1 percent. Thus, the magnitude of the pairwise synergies from this model, particularly those between endowments and technology may motivate an exploration of their interaction in other economies. However, imagining how the higher order interactions might have general implications is difficult.

Table 5.11

Derivation of Three-Way Interaction Values[1]

Railroad Technology, Other Technology and Preferences Index: I(R, T, P)

	ΔGNP(R, T, P)	ΔGNP(R, T)	ΔGNP(R, P)	ΔGNP(T, P)	ΔGNP(R)	ΔGNP(T)	ΔGNP(P)	**GNP Change due to (R,T, P) Interaction**
E=0	546	189	-2	-9	63	299	6	0
E=1	1070	310	-2	-33	90	620	80	5
							mean =	3

Railroad Technology, Other Technology and Endowment Index: I(R, T, E)

	ΔGNP(R, T, E)	ΔGNP(R, T)	ΔGNP(R, E)	ΔGNP(T, E)	ΔGNP(R)	ΔGNP(T)	ΔGNP(E)	**GNP Change due to (R,T,E) Interaction**
P=0	7466	189	27	321	63	299	6446	121
P=1	7510	189	27	297	61	290	6520	126
							mean =	124

Railroad Technology, Endowments and Preferences Index: I(R, E, P)

	ΔGNP(R, E, P)	ΔGNP(R, E)	ΔGNP(R, P)	ΔGNP(E, P)	ΔGNP(R)	ΔGNP(P)	ΔGNP(E)	**GNP Change due to (R,P,E) Interaction**
T=0	6614	27	-2	74	63	6	6446	0
T=1	7217	148	-2	50	252	-3	6767	5
							mean =	3

Other Technology, Endowments and Preferences Index: I(T, E, P)

	ΔGNP(T, E, P)	ΔGNP(E, T)	ΔGNP(E, P)	ΔGNP(T, P)	ΔGNP(E)	ΔGNP(T)	ΔGNP(P)	**GNP Change due to (P,T,E) Interaction**
R=0	7113	321	74	-9	6446	299	6	-24
R=1	7453	442	74	-9	6473	488	4	-19
							mean =	-22

note 1: The shocks are denoted as follows. R: railroad technology, P: preferences, T: other technology, E: endowments. The index 0 denotes parameter values in 1870; 1 in 1890.

Table 5.12

Derivation of the Four-Way Interaction Value

Total Change (calculated from Table 5.8)

ΔGNP(R, T, P, E)	7516

Effects of Individual Factors Relative to 1870 (from Table 5.9)

ΔGNP(P)	6
ΔGNP(E)	6446
ΔGNP(R)	63
ΔGNP(T)	299
sum	6814

Effects of Pairwise Interactions Relative to 1870 (from Table 5.10)

I(P, E)	74
I(P, T)	-9
I(T, E)	321
I(R, T)	189
I(R, P)	-2
I(R, E)	27
sum	600

Effects of Three-way Interactions Relative to 1870 (from Table 5.11)

I(R, T, E)	121
I(R, T, P)	0
I(R, P, E)	0
I(P, T, E)	-24
sum	97

Four-Way Interaction Relative to 1870

I(R, T, P, E) = total change - three-way effects - pairwise effects - individual effects
= 7516 - 6814 -600 -97 = **5**

5.5 Conclusion

The simple application of traditional applied general equilibrium modelling techniques to questions in economic history does not make full use of information about the net outcome of changes which might be available to historians but not to contemporary policy analysts. This chapter has proposed an adaptation of the traditional methodology which incorporates this information and allows historians to undertake a systematic decomposition analysis. Such a decomposition analysis seeks to quantify the individual and interactive contributions of simultaneous shocks to historical change. The interactive effects of shocks can be significant: in a simple model of the US, they account for about 9 percent of the change in GNP over the period 1870-1890. Furthermore, the existence of synergies among shocks has implications for the way modellers present their results. The effect of a shock will depend on the presence or absence of other shocks in the simulations, and modellers must explicitly report their simulation results conditional on the status of the remaining shocks in the model.

Finally, although this chapter focusses on the narrow use of static applied general equilibrium models in economic history, decomposition analysis is not limited to such models. Instead, it can be applied to any historical model for which an initial and a final equilibrium can be specified, and the exhaustive set of shocks which cause the transition from the initial to the final equilibrium can be identified. Dynamic stochastic general equilibrium models represent one such extension to the framework proposed here. As with static models, the parameters of such dynamic models could be calibrated so that the set of shocks to the initial period steady state would yield the final

period steady state as an equilibrium. Decomposition analysis could be undertaken by repeatedly solving the model using subsets of the shocks. Insight about the conditional contributions to change of individual shocks and the interactive effects of subsets of those shocks would arise from comparing the ensuing counterfactual steady states. Such a decomposition analysis is technically more challenging than the static case and it provides a direction for future research in this area.

Chapter 6

Conclusions

This thesis has explored extensions to the traditional calibration techniques in applied general equilibrium modelling: Chapter 3 has proposed a methodology for undertaking sensitivity analysis with respect to uncertainty in the data used to calibrate a model and hence, to uncertainty in the values of the model's calibrated parameters; Chapter 4 has explored the role of data adjustments in applied general equilibrium models and illustrated that under statistical criteria applied to the model results, the choice of data adjustment algorithm can be an important element of the modelling process; and Chapter 5 has shown that by modifying the traditional calibration methodology to include information about the combined effects of several shocks or policy changes, applied general equilibrium models can be used to decompose a known historical change into its component causes.

One of the main implications of the research in Chapters 3 and 4 is that important components of the applied general equilibrium modelling process occur in the early stages of data accumulation and adjustment, before the formation of the BED. The calibration element of the modelling process contains several stages: raw data are balanced to form a BED that meets the equilibrium conditions of the model; and together with the specified elasticity parameters, the BED determines the joint values of the calibrated parameters.

The sensitivity analyses developed in Chapter 3 require modellers to focus on the early steps in calibration. They need to know the values of the unadjusted, raw data;

they must have information about the relative reliability of those data; and they have to be able to reproduce the adjustment process. In current calibration practice, however, modellers often discard some of this information. If they report their data at all, they typically report the BED and the values of the elasticities used in calibration. Modellers tend to concentrate their energy and enthusiasm on specifying the model structure and are somewhat cavalier about the original data and how they are adjusted. Adjustments are often undertaken as a series of disparate activities, many of which are undocumented. Hence, one of the normative implications of the sensitivity analysis methodologies proposed in Chapter 3 is that modellers record both the values of the unadjusted data and the data adjustment process.

The conclusions with respect to the importance of the adjustment process in modelling are much stronger in Chapter 4 than in Chapter 3. Chapter 4 argues that not only should modellers be careful to record the data adjustment process, but that they should evaluate possible adjustment procedures before choosing one, and that the evaluation criteria should be the effect of the adjustments on the statistical properties of the model results. The experiments in Chapter 4 illustrate that the choice of adjustment algorithm can significantly affect the statistical properties of the model results, and that, therefore, this choice is a potentially important component of the adjustment process.

For a given policy experiment and model, one of the factors that affects the performance of the adjustment algorithms tested in Chapter 4 is the relative magnitudes of the elements of the unadjusted data. A second set of experiments in Chapter 4 shows that as these relative magnitudes change, the best choice of adjustment algorithm also

changes. These experiments suggest, therefore, that no universal best choice of algorithm exists, but instead that the best choice is specific to the model, the policy experiment in question, and the characteristics of the data.

Chapter 5 shows that by including information on the combined outcome of several simultaneous changes which might be available to historians but not to contemporary policy analysts, economic historians can undertake a systematic decomposition analysis. Such a decomposition analysis seeks to quantify the individual and interactive contributions of simultaneous shocks to historical change. These interactive effects of shocks are not insignificant: in the simple model of the US considered in Chapter 5, they account for about 9 percent of the change in GNP over the period 1870-1890. Furthermore, the existence of synergies among shocks has wider implications for the way modellers present their results. The effect of a specific shock will depend on the presence or absence of the other shocks in the simulations, and modellers must, therefore, explicitly report their simulation results conditional on the status of the remaining shocks in the model.

6.1 Directions for Future Research

Finally, the research in this thesis suggests two directions for further research on applied general equilibrium model calibration. The first is to develop a systematic data adjustment algorithm evaluation technique, so that the algorithm choice can be integrated into the calibration process. Chapter 4 shows that this choice matters for the

model results and, therefore, that developing such a technique is important, but it leaves the development of such a procedure to future research.

The second direction for further research deals with applying the general approach of decomposition analysis to dynamic models. Chapter 5 focusses on decomposition analysis using static applied general equilibrium models in economic history, but the general approach for decomposition analysis can be applied to any historical model for which an initial and a final equilibrium can be specified, and the exhaustive set of shocks which cause the transition from the initial to the final equilibrium can be identified. Dynamic stochastic general equilibrium models fit into this framework. Their parameters could be calibrated so that imposing the set of shocks on the initial steady state causes the economy to move to the final period steady state. As in the static case, decomposition analysis could be undertaken by repeatedly solving the model using subsets of the shocks. A comparison of the initial and counterfactual steady states would provide insight about both the individual and the interactive conditional contributions to change of the various shocks. Although such a decomposition analysis may be technically challenging, it provides a subject for future research.

References

Adams, P., and P. Higgs, "Calibration of Applied General Equilibrium Models from Synthetic Benchmark Equilibrium Data Sets," *Economic Record* 66 (1990), 110-126.

Bacharach, M., *Biproportional Matrices and Input-Output Change* (Cambridge: Cambridge University Press, 1970).

Brixen, Peter, "ECUSAM -80, A Social Accounting Matrix for Ecuador, 1980," an extended version of Chapter 5 of the PhD. Thesis, *The Financial Sector in Applied General Equilibrium Models - The Case of Ecuador*, (1995).

Brooke, A., D. Kendrick, and A. Meeraus, *GAMS: A User's Guide* (Redwood City, California: The Scientific Press, 1988).

Browning, Martin, Lars Peter Hansen, and James J. Heckman, "Micro Data and General Equilibrium Models," in John Taylor and Michael Woodford (eds.), *Handbook of Macroeconomics* (Amsterdam: North-Holland Publishing Company, forthcoming 1999).

Byron, R.P., "Diagnostic Testing and Sensitivity Analysis in the Construction of Social Accounting Matrices," *Journal of the Royal Statistical Society* A 159 (1996), 133-148.

_____, "The Estimation of Large Social Accounting Matrices," *Journal of the Royal Statistical Society* A 141 (1978), 359-367.

Chia, Ngee-Choon, Sadek Wahba, and John Whalley, "A General Equilibrium-Based Social Policy Model for Côte d'Ivoire," Poverty and Social Policy Series Paper 2, The World Bank (1992).

Clements, K. W., "A General Equilibrium Econometric Model of an Open Economy," *International Economic Review* 21 (1980), 469-488.

Crossman, P., "Balancing the Australian National Accounts," *Economic Record* 64 (1988), 39-46.

de Melo, J., "SAM-Based Models: An Introduction," *Journal of Policy Modeling* 10 (1988), 321-325.

Dervis, K., J. de Melo, and S. Robinson, *General Equilibrium Models for Development Policy* (Cambridge: Cambridge University Press, 1982).

DeVuyst, Eric A., and Paul V. Preckel, "Sensitivity Analysis Revisited: A Quadrature-Based Approach," *Journal of Policy Modeling* 19 (1997), 175-185.

Dixon, Peter B., B. R. Parmenter, John Sutton and D. P. Vincent, *ORANI: A Multi sectoral Model of the Australian Economy* (Amsterdam: North-Holland Publishing Company, 1982).

Fishlow, Albert, "Productivity and Technological Change in the Railroad Sector 1840-1910," in National Bureau of Economic Research Conference on Research in Income and Wealth, Studies in Income and Wealth Volume Thirty, *Output, Employment, and Productivity in the United States After 1800* (New York: Columbia University Press, 1966).

Fogel, Robert, "Notes on the Social Saving Controversy," Chapter 11 in Robert Whaples and Dianne C. Betts (eds.), *Historical Perspectives on the American Economy* (Cambridge: Cambridge University Press, 1995).

Golan, Amos, George Judge, and Sherman Robinson, "Recovering Information from Incomplete or Partial Multisectoral Economic Data," *Review of Economics and Statistics* 76 (1994), 541-49.

Günlük-Şenesen, G., and J. M. Bates, "Some Experiments with Methods of Adjusting Unbalanced Data Matrices," *Journal of the Royal Statistical Society* A 151 (1988), 473-480.

Harberger, A. C., "The Incidence of the Corporation Income Tax," *Journal of Political Economy* 70 (1962), 215-240.

Harley, C. Knick and N. F. R. Crafts, "Productivity Growth during the First Industrial Revolution: Inferences from the Pattern of British External Trade," London School of Economics and Political Science Working Papers in Economic History 42/98 (1998).

Harrigan, F. J., J. W. McGilvray, and I. J. McNicoll, "Simulating the Structure of a Regional Economy," *Environment and Planning A* 12 (1980), 927-36.

Harrigan, Frank, and Iain Buchanan, "A Quadratic Programming Approach to Input-Output Estimation and Simulation," *Journal of Regional Science* 24 (1984), 339-58.

Harrison, Glenn W., Richard C. Jones, Larry J. Kimbell, and Randall M. Wigle, "How Robust is Applied General Equilibrium Analysis?," *Journal of Policy Modeling* 15 (1992), 99-115.

Harrison, Glenn W., and H. D. Vinod, "The Sensitivity Analysis of Applied General Equilibrium Models: Completely Randomized Factorial Sample Designs," *The Review of Economics and Statistics* 74 (1992), 357-362.

Hill, Roderick, "Trade Shocks and Employment Change in Canadian Manufacturing Industries: An Applied General Equilibrium Approach," *International Economic Journal* 9 (1995), 73-88.

Hoover, K. D. (1995), "Facts and Artifacts: Calibration and the Empirical Assessment of Real Business Cycle Models," *Oxford Economic Papers* 45, 24-45.

International Bank for Reconstruction and Development, The World Bank, Socio-economic Time-Series Access and Retrieval System, CDROM Data Base Version 3.0 (1993).

James, John, "The Use of General Equilibrium Analysis in Economic History," *Explorations in Economic History* 15 (1984), 231-253.

Jorgenson, D., "Econometric Methods for Applied General Equilibrium Analysis," Chapter 4 in H. E. Scarf, and J. B. Shoven (eds.), *Applied General Equilibrium Analysis* (Chicago: University of Chicago Press, 1984).

Jorgenson, Dale W., Daniel T. Slesnick, and Peter J. Wilcoxen, "Carbon Taxes and Economic Welfare," *Brookings Papers on Economic Activity* Microeconomics 0 (1992), 393-341.

Kehoe, Timothy J. (1985), "Multiplicity of Equilibria and Comparative Statics," *Quarterly Journal of Economics* 10(1), 119-47.

Kehoe, Timothy J., and John Whalley, "Uniqueness of Equilibrium in Large-Scale Numerical General Equilibrium Models," *Journal of Public Economics* 28 (1985), 247-54.

Khan, Abdul Qayyum, "Comparisons of Naive and RAS Methods of Updating Input-Output Tables: The Case of Pakistan," *Economic Systems Research* 5 (1993), 55-61.

Kocherlakota, Narayana R., "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature* 34 (1996), 42-71.

Kuznets, Simon, *National Product Since 1869* (New York: National Bureau of Economic Research, 1946).

Kydland, F. E., and E. C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica* 50 (1982), 1345-70.

Lebergott, Stanley, "Labor Force and Employment, 1800-1960," in National Bureau of Economic Research Conference on Research in Income and Wealth, Studies in Income and Wealth Volume Thirty, *Output, Employment, and Productivity in the United States After 1800* (New York: Columbia University Press, 1966).

Lecomber, J. R. C., "A Critique of Methods if Adjusting, Updating and Projecting Matrices," Chapter 1 in R. I. G. Allen and W. F. Gossling (eds.), *Estimating and Projecting Input-Output Coefficients* (London: Input-Output Publishing Company, 1975).

Lynch, R. G., "An Assessment of the RAS Method of Updating Input-Output Tables," in Ira Sohn (ed.), *Readings in Input-Output Analysis, Theory and Application* (New York: Oxford University Press 1986).

Malizia, Emil, and Daniel L. Bond, "Empirical Tests of the RAS Method of Interindustry Coefficient Adjustment," *Journal of Regional Science* 14 (1974), 355-65.

Mansur, A., and J. Whalley, "Numerical Specification of Applied General Equilibrium Models: Estimation, Calibration and Data," Chapter 3 in H. E. Scarf, and J. B. Shoven (eds.), *Applied General Equilibrium Analysis* (Chicago: University of Chicago Press, 1984).

McKitrick, Ross R., "The Econometric Critique of Computable General Equilibrium Modeling: The Role of Functional Forms," *Economic Modelling* 15 (1998), 543-573.

_____,"The Econometric Critique of Applied General Equilibrium Modeling: The Role of Parameter Estimation," University of British Columbia, Department of Economics Discussion Paper 95/27 (1995).

Mehra, Rajnish, and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15 (1985), 145-61.

Miller, Allen, and Thomas Rice, "Discrete Approximations of Probability Distributions," *Management Science* 29 (1983), 352-362.

O'Rourke, Kevin, "The Repeal of the Corn Laws and Irish Emigration,"*Explorations in Economic History* 31 (1994), 120-38.

_____,"Did the Great Irish Famine Matter?," *The Journal of Economic History* 51 (1991), 1-22.

Pagan, A., and J. Shannon, "Sensitivity Analysis for Linearized Computable General Equilibrium Models," in J. Piggott and J. Whalley (eds.), *New Developments in Applied General Equilibrium Analysis* (Cambridge: Cambridge University Press, 1985).

_____, "How Reliable are ORANI Conclusions?," *Economic Record* 63 (1987), 33-45.

Parikh, Ashok, "Forecasts of Input-Output Matrices Using the R.A.S. Method," *The Review of Economics and Statistics* 61 (1979), 477-81.

Piggott, J., and J. Whalley, *U.K. Tax Policy and Applied General Equilibrium Analysis* (Cambridge: Cambridge University Press, 1985).

Perroni, Carlo, and Thomas F. Rutherford, "A Comparison of the Performance of Flexible Functional Forms for Use in Applied General Equilibrium Modelling," *Computational Economics* 11 (1998), 245-63.

Prebisch, R., "The Economic Development of Latin America and its Principal Problems," *Economic Bulletin for Latin America* 7 (1962), 1-22. (First published as an independent booklet by UN ECLA, 1950).

Preckel, Paul V., and Eric DeVuyst, "Efficient Handling of Probability Information for Decision Analysis under Risk," *American Journal of Agricultural Economics* (1992), 655-662.

Roberts, Barbara M., "Calibration Procedure and the Robustness of CGE Models: Simulations with a Model for Poland," *Economics of Planning* 27 (1984), 189-210.

Rutherford, Thomas F., "General Equilibrium Modelling with MPS/GE," unpublished manuscript, Department of Economics, University of Western Ontario, Apr. 1989.

Rutström, E. E., *The Political Economy of Protectionism in Indonesia: A Computable General Equilibrium Analysis* (Stockholm: Economic Research Institute, The Stockholm School of Economics, 1991).

Scarf, H. E. (with T. Hansen), *The Computation of Economic Equilibria* (New Haven: Yale University Press, 1973).

Schneider, Michael H., and Stavros A. Zenios, "A Comparative Study of Algorithms for Matrix Balancing," *Operations Research* 38 (1990), 439-455.

Shoven, John B., and John Whalley, *Applying General Equilibrium* (Cambridge: Cambridge University Press, 1992).

_____, "Applied General Equilibrium Models of Taxation and International Trade: An Introduction and Survey," *Journal of Economic Literature* 22 (1984), 1007-1051.

_____, "A General Equilibrium Calculation of the Effects of Differential Taxation of Income from Capital in the U.S.," *Journal of Public Economics* 1 (1972), 281-322.

Singer H. W., "The Distribution of Gains Between Investing and Borrowing Countries," *American Economic Review, Papers and Proceedings* 40 (1950), 473-85.

Siriwardana, Mahinda, "The Causes of the Depression in Australia in the 1930s: A General Equilibrium Evaluation," *Explorations in Economic History* 32 (1995), 51-81.

St.-Hilaire, F., and J. Whalley, "A Microconsistent Equilibrium Data Set for Canada for Use in Tax Policy Analysis, " *Review of Income and Wealth* 29 (1983), 175-204.

Stone, R., "The Development of Economic Data Systems," Foreword to G. Pyatt and J. Round, *Social Accounting for Development Planning* (New York: Cambridge University Press, 1978).

Temin, Peter, "General-Equilibrium Models in Economic History," *Journal of Economic History* 31 (1971), 58-75.

Thomas, Mark, "General Equilibrium Models and Research in Economic History," Chapter 4 in A. Field (ed.) *The Future of Economic History* (Boston: Kluwer-Nijhoff, 1987).

United States, Bureau of the Census, *Historical Statistics of the United States: Colonial times to 1970* (Washington: United States Bureau of the Census, 1976).

Watson, M. W. (1993), "Measures of Fit for Calibrated Models," *Journal of Political Economy* 101(6), 1011-41.

Weale, Martin, "Testing Linear Hypotheses on National Account Data," *The Review of Economics and Statistics* 67 (1985), 685-89.

Whalley, J., *Trade Liberalization Among Major World Trading Areas* (Cambridge MA: MIT Press, 1985).

Wiese, Arthur M., "On the Construction of the Total Accounts From the U.S. National Income and Product Accounts: How Sensitive Are Applied General Equilibrium Results to Initial Conditions," *Journal of Policy Modeling* 17 (1995), 139-162.

Wigle, R. M., "The Pagan-Shannon Approximation: Unconditional Systematic Sensitivity in Minutes," in J. Piggott and J. Whalley (eds.), *Applied General Equilibrium* (Heidelberg: Physica-Verlag, 1991).

Williamson, Jeffrey G., *Late Nineteenth-Century American Development* (New York: Cambridge University Press, 1974).