

**ESTIMATING CONSUMER BEHAVIOUR IN AN
ENERGY-ECONOMY POLICY MODEL**

by

Dale Beugin
B.A.Sc., University of British Columbia, 2003

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF RESOURCE MANAGEMENT

In the
School of Resource and Environmental Management

© Dale Beugin, 2007

SIMON FRASER UNIVERSITY

Fall, 2007

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Dale Beugin
Degree: Master of Resource Management
Title of Thesis: Estimating Consumer Behaviour in an Energy-Economy Policy Model
Project No.: 443
Examining Committee:
Chair: Vicky Weekes
Master of Resource Management Candidate

Mark Jaccard
Senior Supervisor
Professor
School of Resource and Environmental Management
Simon Fraser University

John Nyboer
Supervisor
Adjunct Professor
School of Resource and Environmental Management
Simon Fraser University

Sean Cox
Supervisor
Assistant Professor
School of Resource and Environmental Management
Simon Fraser University

Date Defended/Approved:

Dec 3, 2007



SIMON FRASER UNIVERSITY
LIBRARY

Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website <www.lib.sfu.ca> at: <<http://ir.lib.sfu.ca/handle/1892/112>>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

ABSTRACT

CIMS is a technologically explicit hybrid energy-economy model that forecasts effects of policy alternatives on technological change and greenhouse gas emissions. It strives towards realistically representing consumer behaviour to better forecast effects of policies. This study attempts to improve the behavioural realism of the model by calibrating the parameters representing consumer behaviour using historical data from 1990 to 2004. A statistical simulation called Markov Chain Monte Carlo generates a probability distribution over the behavioural parameters for three technology competitions. The calibrated model is then applied to a policy analysis forecasting the effects of a carbon tax on residential furnace emissions to 2050. Despite insufficient variation in energy prices over the historical period, uncertainty in model structure, and an absence of revealed preference data for emerging technologies, historical calibration can improve model credibility and thus usefulness for policy-makers, particularly when used in combination with other, stated preference parameter estimation research.

Keywords: Energy-economy models; Calibration; Consumer behaviour; Uncertainty; Energy policy

Subject Terms: Energy policy – mathematical models; Bayesian statistical inference

ACKNOWLEDGEMENTS

Thanks to everyone who has helped me, supported me, and generally kept me sane during this project:

- Mark Jaccard for challenging me and providing the inspiration for the project
- John Nyboer for many productive and stimulating brainstorming and problem solving sessions
- Sean Cox for statistical and simulation wisdom / black magic
- Nic Rivers, Bill Tubbs, Chris Bataille, Aaron Springford, Jotham Peters, Rose Murphy, Jillian Mallory and Pritam Ranjan for help and good ideas throughout this process
- Steve Groves for moral support, alternative perspectives, and editing
- Vicky Weekes for editing, balance, and all around back-up
- Many other REMmers and EMRGers for their enthusiasm, academic inspiration, and friendship

Thanks also to the National Science and Engineering Research Council, the Canadian Institute of Energy, the Energy and Materials Research Group, and Simon Fraser University for providing the funding for this project.

TABLE OF CONTENTS

Approval	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	viii
List of Tables	x
List of Acronyms	xi
Chapter 1: Project Rationale	1
1.1 Introduction.....	1
1.2 Background.....	2
1.2.1 Energy economy modelling	2
1.2.2 Consumer behaviour in energy-economy models.....	5
1.2.3 CIMS and behaviour	8
1.2.4 Empirical basis for behaviour	12
1.2.5 Sources of model uncertainty.....	15
1.3 Objectives	16
1.4 Report structure.....	17
Chapter 2: Methodology	18
2.1 Overview.....	18
2.2 Principles of calibration and backcasting	18
2.3 Review of complex model calibration approaches	20
2.3.1 Subjective parameter estimation	20
2.3.2 Partial calibration	21
2.3.3 Deterministic optimization or model fitting	21
2.3.4 Parameter estimation with uncertainty.....	23
2.3.5 Bayesian Monte Carlo updating.....	24
2.3.6 Calibration under uncertainty	24
2.3.7 Sequential experimental design	25
2.4 Bayesian Markov Chain Monte Carlo (MCMC) theory	25
2.5 Implementing Metroplis-Hastings for CIMS.....	29
2.5.1 Scoping the analysis.....	30
2.5.2 Matching historical and model frameworks	31
2.5.3 Selecting likelihood and posterior functions.....	33
2.5.4 Constraining parameter space	35
2.5.5 Informative prior probability distributions	35

2.6	Methodological summary	36
Chapter 3: Analysis and Discussion		37
3.1	Overview.....	37
3.2	Refrigerator node	37
3.2.1	Dynamic historical refrigerator market share trends.....	37
3.2.2	Declining capital costs and autonomous energy efficiency increases	40
3.2.3	Changing services	41
3.2.4	Effects of policies	42
3.2.5	Dynamic intangible costs	43
3.2.6	Refrigeration node summary.....	45
3.3	Furnace node	46
3.3.1	Posterior probability distribution for residential furnace parameters	47
3.3.2	Justification for convergence	49
3.3.3	Importance of prior probability distributions.....	51
3.3.4	Correlation between parameters	53
3.3.5	Modeling historical policies.....	55
3.4	Vehicle node	55
3.4.1	Dynamic historical vehicle preferences	57
3.4.2	Modeling dynamic preferences for vehicles	59
3.4.3	Vehicle parameters with dynamic intangible costs.....	62
3.4.4	Alternative model structures	65
3.5	Summary	67
Chapter 4: Policy analysis under uncertainty		69
4.1	Introduction.....	69
4.2	Energy policy instruments	69
4.3	Stochastic policy analysis	71
4.4	Incorporating uncertainty into a CIMS policy analysis	72
4.5	Policy simulation results	74
4.5.1	Impact of calibrated parameters.....	75
4.5.2	Impact of explicitly incorporating uncertainty.....	77
4.6	Summary	79
Chapter 5: Conclusions.....		80
5.1	Modelling conclusions	80
5.1.1	Estimating behavioural parameters	80
5.1.2	Incorporating uncertainty	80
5.1.3	Assessing model structure.....	81
5.2	Policy conclusions	82
5.3	Limitations of the analysis and potential future research	82
5.3.1	Accounting for the possible multiple parameter values.....	83
5.3.2	Data limitations and revealed preferences	83
5.3.3	Structural uncertainty and issues in extrapolation	85
5.3.4	No autocorrelation in parametric uncertainty	87
5.4	Assessment of historical calibration for technology explicit models	88

Appendices	91
Appendix A: The Bayesian framework in limited mathematical detail	92
Bayes' Rule.....	92
Prior probability distributions	93
Likelihood functions	94
Appendix B: Model results of forecast analysis	95
Reference List	96

LIST OF FIGURES

Figure 1: Framework for assessing the usefulness of energy-economy models.....	3
Figure 2: Effect of the value of heterogeneity parameter, ν on consumer preferences for a simplified technology competition between two technologies, A and B with life cycle costs LCC A and LCC B.....	10
Figure 3: Conceptual model of a single CIMS technology competition (residential refrigeration).....	11
Figure 4: Process flow diagram for conventional Metropolis-Hastings Markov Chain Monte Carlo parameter estimation simulation.....	29
Figure 5: Process flow diagram for modified Metropolis-Hastings Markov Chain Monte Carlo parameter estimation simulation	35
Figure 6: Marginal posterior probability distributions for refrigerator parameters, heterogeneity, ν , and intangible costs i_1 through i_4 , for low, medium, high and super efficient refrigerators respectively.	38
Figure 7: Historical (points) and forecasted (lines) new market shares (forecasted trajectories are generated by running the model on the most probable parameter combination as estimated by calibration).....	39
Figure 8: Calculated dynamic intangible costs for refrigerator archetypes at $\nu = 10$ and $r = 0.6$	44
Figure 9: Unscaled marginal posterior probability distributions output from MCMC calibration routine with 40000 iterations; Ratio of accepted points to total number of candidate points = 0.228; standard deviation of jump distribution = (0.01, 0, 5, 0, 5, 5, 5, 0.0).	47
Figure 10: Trace Plots for MCMC run with 40,000 iterations (burn-in iterations removed); Ratio of accepted points to total number of candidate points = 0.228.....	50
Figure 11: Comparison of forecasts with modes of MCMC-output marginal posterior probability distributions ($\theta = 6.35, 670, 15, -585, -5695, 0.3$) (lines) vs historical trends (points)	51
Figure 12: Correlation between medium efficiency natural gas furnace intangible cost, and high efficiency natural gas furnace intangible cost in posterior sample points.....	54
Figure 13: Structure of gasoline vehicle node for backcasting	56

Figure 14: Comparison of historical data (points) with forecasted trends using calibrated, best-fit parameters (lines)	58
Figure 15: Unscaled marginal posterior probability distributions for behavioural parameter values. Ratio of accepted points in calibration = 0.278; number of iterations = 50,000 (additional 50,000 removed as burn-in).....	62
Figure 16: Comparison of historical and forecasted (using behavioural parameters equal to the mode of the estimated posterior distribution) market shares from 1990 - 2003 for gasoline vehicles.....	63
Figure 17: Historical (points; prior to 2005) and forecasted greenhouse gas trajectories for Ontario residential furnace technologies under three alternative modelling scenarios under no policy, or “business as usual” (solid) and under an escalating carbon tax (dashed). Dotted lines in panel (c) represent 90% uncertainty bounds. 300 iterations were used in the Monte Carlo stochastic run.....	75

LIST OF TABLES

Table 1: Overview of Consumer Reports refrigerator cost data (freezer-on-top models only)	40
Table 2: Correlation between calibrated parameters in estimated joint posterior probability distribution	53
Table 3: Results from other CIMS parameter estimation studies for values of behavioural parameters	65
Table 4: Summary of estimated parameters through calibration	67
Table 5: Escalating carbon tax simulated in policy analysis, paralleling Bataille <i>et al.</i> (2007)	73
Table 6: Behavioural parameters used in forecasting for each of three modelling scenarios	73
Table 7: Previous parameter values used in CIMS	93
Table 8: Forecasted technology new market shares in 2005 under each of three modelling scenarios	95
Table 9: Forecasted technology new market shares in 2050 under each of three modelling scenarios	95

LIST OF ACRONYMS

AEEI	Autonomous Energy Efficiency Improvements
BAU	Business as Usual
BIC	Bayesian Information Criterion
CAFE	Corporate Average Fuel Economy
CFCs	Chlorofluorocarbons
CGE	Computable General Equilibrium
CIMS	Formerly, the CIMS model was known as the “Canadian Integrated Modelling Sytem.” Since it has applied to other countries, “CIMS” is no longer an acronym, merely the name of the model
GDP	Gross Domestic Product
HFCs	Hydrofluorocarbons
LCC	Life cycle cost
MCMC	Markov Chain Monte Carlo
SUV	Sport Utility Vehicle

CHAPTER 1: PROJECT RATIONALE

1.1 Introduction

Human-induced climate change is a critical environmental policy problem for Canada and the rest of the world. Society's reliance on a wide range of energy-using technologies, a major source of greenhouse gas emissions, makes the problem complex. Assessing policy options for such a complex problem is challenging. Energy-economy models are therefore essential tools for policy makers attempting to manage complex environmental issues such as climate change. One such model is CIMS, which forecasts the effects of policies on technological change. Analysts can use these forecasts to determine the relative effectiveness of alternative policy packages for reducing greenhouse gas emissions. These results can provide practical information for policy decisions.

Yet how much confidence should policy-makers place in model forecasts? After all, no model is a perfect representation of reality. One way in which models such as CIMS try to provide better forecasts of the impacts of policies is with parameters intended to realistically simulate the decisions and behaviour of firms, households, and individual consumers.¹ However, these parameters are highly uncertain, particularly in terms of representing future consumer behaviour.

A possible approach to addressing this uncertainty involves projecting future behaviour based on past behaviour. More specifically, behavioural parameters can be adjusted so that model outputs from a simulation over a historical period match actual historical outcomes as closely as possible. This process is called *calibration*. Since calibration is based on empirical (historical) data, it can improve the credibility of forecasts. Further,

¹ Henceforth, I will refer to all purchasers of energy-using equipment (firms, households, and individuals) simply as "consumers." Residential energy consumers are typically individuals or households, while Industrial and Commercial consumers are firms.

using a statistical simulation approach, calibration can generate probability distributions for behavioural parameter values. Distributions express the probability that parameter values reflect historical behaviour. Generating these distributions and integrating the results into the forecasting model in order to improve its credibility is the overarching goal of this study.

In the remainder of this chapter, I provide background information and context within the energy-economy modelling literature to frame the goals of the study. In section 1.2, I first review key issues in modelling consumer behaviour in energy-economy models and describe important elements of the CIMS model. I also assess the advantages and disadvantages of alternative approaches to empirically estimating behavioural parameters. I then discuss issues of uncertainty in energy-economy models. In section 1.3, I build on this foundation of background research to identify three specific objectives for this study. Finally, in section 1.4, I conclude the chapter with a road map to the structure of this report in full.

1.2 Background

The evolution of energy-economy models in recent years highlights the challenges of realistically representing consumer behaviour in models. To help frame the objectives of this calibration study, in this section I review issues in modelling consumer behaviour and quantifying uncertainty in models.

1.2.1 Energy economy modelling

Jaccard (2005a) identifies three criteria to assess the “usefulness” of a model in terms of its ability to evaluate the combined effects of alternative policies: technological explicitness, behavioural realism, and the incorporation of *macro-economic equilibrium feedbacks* between energy-technology decisions and the larger economy.² Hourcade *et al.* (2007) redefine these criteria as technological explicitness, microeconomic realism, and

² “Macro-economic feedbacks” refers to interactions between the energy system and other sectors of the economy as a whole. This criterion thus assesses how well an energy model ensures consistency with other economic factors such as capital and labour.

macro-economic explicitness, emphasizing that “behaviour” in an energy-economy modelling context refers to the purchasing preferences of agents (firms or individual consumers). Since this study focuses on estimating *behavioural* parameters, I continue to use the “behavioural realism” label. In the past, two main energy-economy modelling paradigms, “bottom-up” and “top-down” models, have competed as potential analytical tools for policy makers. Figure 1 illustrates a conceptual framework for assessing energy-economy models and shows how neither paradigm meets all three criteria perfectly.

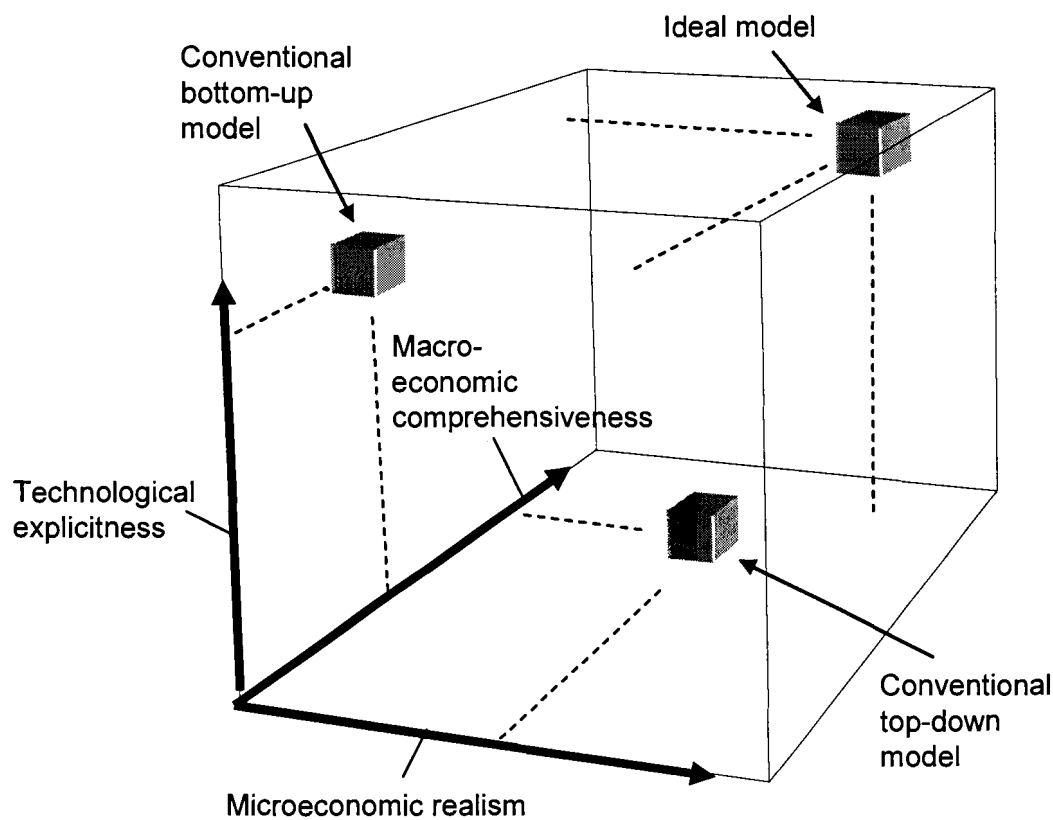


Figure 1: Framework for assessing the usefulness of energy-economy models

Source: used with permission from Hourcade et al., 2007

Conventional “bottom-up” models are technologically explicit. These models compare the *lifetime cost* of a (typically large) number of technologies that compete to provide a

service.³ They then assume consumers have perfect foresight and unerringly choose the optimal, lowest cost alternative. Since the lifecycle costs of newer technologies with lower greenhouse gas emissions tend to be slightly lower over the full life of the technology, bottom-up models tend to show potential for relatively inexpensive reductions of emissions through the diffusion of these technologies (Jaccard, 2005a). In reality, however, operating and capital costs are not the only factors on which consumers base decisions. Similarly, not all consumers have identical preferences and do not make identical choices. For these reasons, typical bottom-up models are not behaviourally realistic. Further, conventional bottom-up models do not accurately represent the macro-economic feedback effects of different energy pathways and policies that might affect economic growth (Hourcade et al., 2007). They thus fare less well on the second two criteria, as illustrated in Figure 1.

Conventional top-down models, on the other hand, are not technologically explicit, but instead represent energy-economy interactions on a macro-economic scale. These models depict production systems and technologies on an aggregate scale, though some allow for varying degrees of detail within sectors (Carraro and Hourcade, 1998). Computable General Equilibrium (CGE) models are one example of top-down models. Computable general equilibrium models are more behaviourally realistic than conventional bottom-up models if models use historical data to estimate aggregate relationships between energy costs, market shares and other inputs to the economy. These relationships are linked to total economic output on a broader macro-economic equilibrium framework fulfilling the third criteria for model usefulness (Jaccard et al., 2003). Top-down models, however, are useful only for modelling top-level policy instruments such as carbon taxes or tradable permits: the aggregate nature of top-down models precludes any ability to model technology-specific policies or to determine industries most affected by a given policy (Jaccard et al., 2003).

³ A life cycle cost calculates the net present value of the costs associated with a technology over its lifespan, including annual operating and maintenance, fuel, and capital costs. Alternatively, a “levelized life cycle cost” calculates all costs associated with the technology amortized over its lifespan.

By combining elements of both bottom-up and top-down models, hybrid models such as CIMS attempt to meet all three criteria to as great an extent as possible, again represented conceptually in Figure 1. Technologies that demand and supply energy are represented explicitly in CIMS, and recent work has improved the model's representation of macro-economic feedbacks (Bataille et al., 2007). Finally, CIMS also attempts to realistically represent consumer technology-choice behaviour through empirically estimated parameters (Jaccard, 2005a). My focus in this study is the behavioural realism criterion. In it, I continue efforts to improve the "behavioural realism" of CIMS by providing better empirical support for model parameters.

1.2.2 Consumer behaviour in energy-economy models

Assumptions about consumer behaviour drive energy-economy models. Consumers' preferences when purchasing a new car, a light bulb, an industrial steam boiler, or any energy-consuming equipment, dictate how an energy system changes over time. Different models, however, take different approaches to representing consumer behaviour. Nyboer and Bataille (2007) provide an overview of the evolution of behavioural modelling literature in the context of energy-economy models in particular. This section extracts a few key elements from their analysis. Since the goal of this study as a whole is calibration, not the total reinvention of CIMS' behavioural framework, I do not attempt to provide an in-depth review of all paradigms of behavioural modelling. Rather, I provide a brief overview of key ideas in the field to provide context for CIMS' representation of consumer behaviour.

Bottom-up models typically assume that consumers purchase a product that has the lowest life cycle costs. This principle of cost minimization is consistent with the "rational actor" construct and principles of utility maximization, and thus relies on greatly simplified micro-economic theory more than empirical realism. Empirical studies, however, indicate a disjoint between real-world behaviour and theoretical cost minimization; they suggest that the revealed discount rate, reflecting consumers' time preferences, is substantially higher than the productivity of capital (Train, 1985). This discrepancy is known as the "energy efficiency gap" (Jaffe, Newell, Stavins, 1999).

Essentially, the gap indicates that consumers were less prone to purchasing energy-efficient goods that would cost more in the short term, but save money in the longer term. New energy efficient technologies have diffused through the market much more slowly than predicted by bottom-up optimization models. In order to simulate actual behaviour, modelling choices through pure cost minimization theory is thus clearly insufficient, and further analysis of possible mechanisms for actual consumer behaviour and the energy efficiency gap is required.

Consumer behaviour literature presents different perspectives on the energy efficiency gap. One perspective, paralleling bottom-up models, frames the lack of investment in energy-efficient technology as the result of *market failures*, which result in consumers making socially sub-optimal choices. More specifically, Ruderman *et al.* (1987) point to a lack of information for consumers, constraints on available capital, small expected savings, and manufacturers choosing which products to market as possible causes of market failure. High transactions costs, including the high costs of searching for product information, as well as negotiating, monitoring, and enforcing contracts may add to the market failure. Moxnes' (2004) empirical study supports these hypotheses, suggesting that efficiency standards can improve consumers' welfare given that they make sub-optimal decisions as a result of a lack of information. As a bottom line, this perspective suggests that policies that overcome the market failures and the lower operating costs of efficient new technologies can lead to substantial penetration of these technologies.

A second perspective, associated with economists, argues that market *barriers* representing real costs, not market *failures*, are responsible for the lack of penetration of efficient technologies (Jaffe, Newell and Stavins, 1999). Given uncertainty, rational consumer decisions may transcend simple cost minimization. Trade-offs exist between economic efficiency and energy efficiency (Jaffe, Newell and Stavins (1999)).⁴ Further, the substantial heterogeneity in the needs of consumers, and thus in their expected value for a given technology, further decreases the penetration of an energy-efficient technology. Adopting energy-efficient products can have real, though varying, costs for

⁴ If a choice is not economically efficient, the real costs of a technology are greater than the real benefits, as experienced by the consumer.

consumers (albeit costs not included in a conventional bottom-up model). Including these costs in an energy economy model recognizes that reducing the greenhouse gas intensity of an energy system can be more expensive than the pure bottom-up approach assumes.

Under uncertainty, consumers' choices of technology might be more complicated. Sutherland (1991) argues that energy efficient investments can be very risky due to high transaction costs, irreversibility, and the fact that for small households, risks cannot be diversified by choosing a combination of high and low risk products. Households have real incentives to make cautious, risk averse choices. If a household with limited income invests in an expensive vehicle with new, efficient technology, the consequences are substantial if the vehicle ends up proving unsatisfactory; the large cost of a replacement car might make the decision irreversible. Similarly, unlike a firm, a household cannot diversify its risk by populating a corporate fleet of vehicles with a range of newer and older technology vehicles. Hassett and Metcalfe (1993) suggest that conservation investments require large sunk or irreversible costs; as such, waiting for more information before investing has *option value*. Again, the consumer has an incentive not to purchase a new technology before gathering additional information, which in itself might be costly. Paralleling the market failure explanation, Johnson (1994) argues that the "state of information" of the decision-maker plays a role and that the cost of collecting information must be included in comparing product options. Finally, uncertain future energy prices increase the value of delaying a decision. A high discount rate might therefore represent optimal investment strategic behaviour under uncertainty.⁵ I discuss how these market barriers might be incorporated into a model simulating consumer behaviour in the next section.

In an issue quite separate from the energy efficiency gap debate, energy-economy models also have varied in their approaches to addressing how consumer behaviour has changed over time. Nyboer and Bataille (2007, p. 10) suggest that many economists have

⁵ Other explanations for revealed consumer behaviour and the energy efficiency gap try to move away from relying solely on an economic framework. Jager *et al.* (2000) incorporate elements of psychology and implicitly describe other sources of behavioural inertia, suggesting that in addition to having sufficient information, a consumer must be dissatisfied with the status quo before altering their behaviour.

generally tended toward modelling consumer preferences as “an exogenous, unchanging parameter derived from historical data” so as to avoid the pitfalls of trying to endogenously model complex behaviour. On the other hand, analysts such as Norton *et al.* (1998) argue that consumers’ changing preferences are too important to ignore. For example, empirical research by Mau (2002) identifies a “neighbour effect,” which suggests that consumers may be more inclined to invest in a new technology as its market shares increases and uncertainty surrounding its performance decreases.

1.2.3 CIMS and behaviour

To reconcile the discrepancies between revealed consumer behaviour and theories of pure cost-minimization, CIMS builds on insights from the behavioural literature in modelling technology-purchase decisions. Similar to bottom-up models, CIMS calculates the life-cycle costs of different competing technologies to predict the market share of future purchases of these technologies.⁶ In addition to capital costs and operating costs, however, behavioural parameters play an important role in forecasting technology market shares. Drawing lessons from the economist perspective on consumers’ technology preferences, CIMS attempts to realistically forecast the adoption of energy efficient technologies, given the existence of market barriers.⁷ These behavioural parameters add a degree of behavioural realism not present in most bottom-up models and address issues developed in the behavioural literature.

CIMS applies a market share function to forecast new market shares for alternative technologies. The function uses the life-cycle costs of different competing technologies to calculate new market shares, $MS_{k,t}$, for each technology option, k , at each time interval, t , as shown in Equation 1.

⁶ CIMS also tracks stock turnover, modelling retirement of existing stock and demand for new stock. Further, the macro-economic component of the model, inspired by top-down models, ensures equilibrium between the energy sector and the rest of the economy by adjusting energy prices if required. This component helps CIMS partially satisfy the third criterion for model usefulness, “incorporation of macro-economic feedback effects.” Again however, the focus of this study is improving the model’s performance in terms of “behavioural realism.”

⁷ In building on the second, economist perspective on the energy efficiency gap issue, as presented in Section 1.2.2, CIMS can also calculate the real costs resulting from imposing a given policy option. Costs of policies, however, are outside the scope of this analysis. My focus here is describing the mechanisms driving consumer preferences.

$$MS_{k,t} = \frac{\left[(CC_k + i_k) \times \left(\frac{r}{1 - (1+r)^{-N}} \right) + MC_k + EC_{k,t} \right]^{-v}}{\sum_{j=1}^K \left[\left((CC_j + i_j) \times \left(\frac{r}{1 - (1+r)^{-N}} \right) + MC_j + EC_{j,t} \right) \right]^{-v}} \quad \text{Equation 1}$$

Where:

- CC_k = the capital costs for technology option k
- v = heterogeneity factor
- i_k = intangible costs of technology option k
- r = private discount rate
- N = lifespan of technology
- MC_k = annual maintenance costs of technology option k
- $EC_{k,t}$ = annual energy operating costs of technology option k at time; $EC_{k,t}$ is calculated as the product of annual energy requirements of technology k times the cost of fuel at time t .
- K = total number of representative technology options, or *archetypes*, competing for the service

As evident in Equation 1, three main behavioural parameters are used in the CIMS market share calculation to represent consumer behaviour: v , i , and r respectively represent market heterogeneity, intangible costs, and a discount rate. While the discount rate, r , still represents consumers' time preferences, the v and i parameters attempt to

reflect that actual choices cannot be described through conventional discounted life-cycle cost minimization.

Market heterogeneity, ν , represents the extent to which different consumers in the same market choose different technologies. Heterogeneity can also be regarded as a measure of consumers' sensitivity to cost: at high values of ν , consumer choices show little variation, and most choose the least cost option. At low values of ν , large variation in technology choice suggests consumers are insensitive to the cost of a technology and make their choices based on other criteria. Figure 2 illustrates the calculated market share splits between two technologies, A and B. As illustrated in the figure, when $\nu = 1$, even if one technology has a lifecycle cost (LCC) twice that of a competing technology, a consumer choosing the expensive technology has a probability of almost 40%. At the other end of the spectrum, when $\nu = 15$, the model trends toward a 100% probability of the less expensive alternative being purchased.

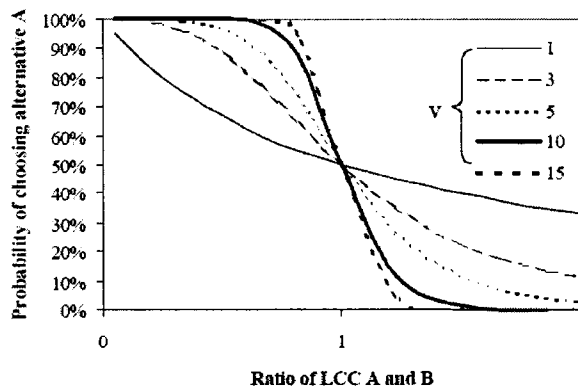


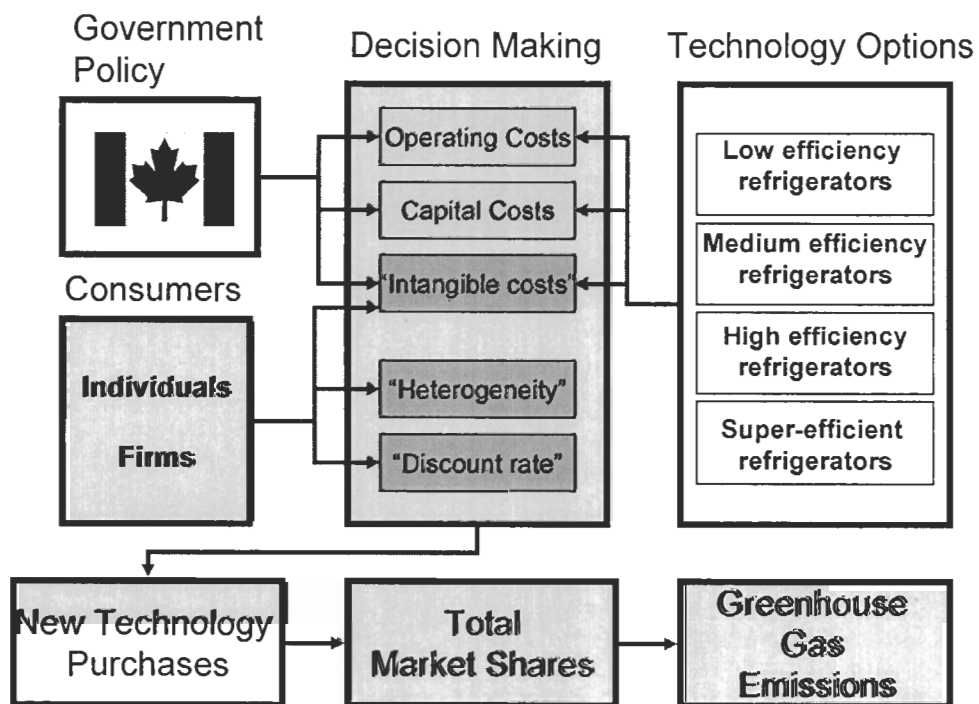
Figure 2: Effect of the value of heterogeneity parameter, ν on consumer preferences for a simplified technology competition between two technologies, A and B with life cycle costs LCC A and LCC B.

Source: Rivers, 2003

Intangible costs, i , represent apparently non-financial costs associated with a given technology. They include factors such as the risk of failure of a new technology, option value in delaying adoption of new technologies, or the costs of overcoming market

barriers such as lack of information regarding the technology. Each competing technology is thus associated with an intangible cost that represents these non-financial costs.

Figure 3 illustrates a conceptual model for CIMS' representation of technology purchase decisions for one technology competition, or "node," (residential refrigeration in the case of the figure) based on these behavioural parameters and technology costs. The figure also suggests how the model represents the effects of policies on decision-making: subsidies might reduce costs of a given technology; information programs might attempt to reduce consumers' perceptions of risk or overcome information market barriers and thus reduce intangible costs.



5

Figure 3: Conceptual model of a single CIMS technology competition (residential refrigeration)

1.2.4 Empirical basis for behaviour

By including the behavioural parameters v , i , and r , the CIMS algorithm can potentially represent the purchasing dynamics described in Section 1.2.2 and improve on conventional, cost-minimizing, bottom-up models. This improvement can only be realized, however, if the values for behavioural parameters are supported by empirical evidence.

Prior to 2000, behavioural parameters for CIMS were estimated using a “combination of literature review, judgment, and meta-analysis” (Jaccard, 2005a, p.14). Two main approaches exist as options for improving the empirical justification for parameter values, discrete choice modelling and calibration.

Recently, analysts using CIMS have quantified the behavioural parameters using discrete choice methods (Horne, 2003; Rivers, 2003; Eyzaguirre, 2004). Discrete choice models, drawing on the well-established field of choice theory, can be useful in representing choices between discrete options. Such models quantify the trade-offs made by consumers choosing a technology. Choice models can be informed either through stated or revealed preferences; each approach has strengths and weaknesses.

Discrete choice models based on revealed preference are based on actual choices made by real consumers. These models therefore fully account for all aspects of consumers’ decisions, such as income level and access to given technologies (Axsen, 2006). The revealed preference approach results in a reliable, credible representation of consumer behaviour. However, revealed preference data is limited to current or historical conditions; extrapolating to inform a model forecasting future decisions between technology options that are not yet available, or choices under dramatically different market conditions, is problematic. The revealed preference approach is also statistically more complicated, because factors effecting choice may not have varied substantially over the historical period, or multiple effects may have changed in the same way (colinearity) (Axsen, 2006).

In stated preference studies on the other hand, surveys estimating consumer preferences from respondents' hypothetical choices under a range of conditions inform the model. Horne (2003) surveyed consumers' stated vehicle technology preferences. Rivers (2003) uses surveys to estimate industrial consumers preferences regarding steam generation technologies. Surveys are flexible and can poll respondents choices under a variety of conditions. Further, a stated preference approach can assess preferences for new and emerging technologies for which no revealed preference data yet exists. Eyzaguirre (2004), for example, uses discrete choice surveys to estimate preferences for hydrogen powered vehicles. Similarly, Mau (2005) quantifies how consumer preferences for hybrid vehicles might change with an increasing hybrid market share. However, stated preference approaches are vulnerable to a hypothetical bias: respondents' responses may differ significantly from their actual behaviour. As summarized by Jaccard (2005a), these discrepancies may exist because survey respondents do not face real-world budgetary or information constraints, may not understand the survey properly, or may answer strategically to influence results (Louviere et al., 2000; Train, 2002). Urban et al. (1996) suggest that consumers are often more inclined to select higher efficiency technologies in surveys than in reality.

Finally, discrete choice surveys can sometimes combine stated and revealed preference approaches. Axsen's (2006) study attempts to overcome the weaknesses of each approach using a discrete choice survey that combines stated preferences survey data with revealed preference based on past hybrid car purchases in California.

Calibration provides an alternative approach to estimating consumer behaviour. A calibration process involves adjusting model inputs (parameters) to produce the best fit between model outputs and real world data. In the case of CIMS, the model can be run over a historical period, and forecasted market shares of selected technologies compared to historical market shares over that period. Values for behavioural parameter that result in model outputs that best match empirical data can then be determined. This approach would represent a *dynamic* calibration in that it could capture dynamics in model

forecasts by calibrating to time series data, unlike a simpler *static* calibration that matches model outputs to real world data for only a single time period.

Behaviour estimated through calibration represents revealed preferences, given it incorporates actual, historical consumer choices. Calibration thus explicitly connects the model with the real world. Cooley (1997, p. 56) writes, “Calibration is a strategy for finding numerical values of parameters of artificial economic worlds... In the calibration approach, measurements are used to give content to theory.” Calibration provides a framework for comparing the theoretical model representation with reality by running the model over a historical period and comparing it to real world data. This framework can thus also allow us to improve the connection between theory and empiricism, model and reality, and attempt to reduce the uncertainty associated with representing real consumer behaviour with a theoretical model.

Nevertheless, like a discrete choice modelling approach that relies solely on revealed preferences, calibration may be constrained in how much it can inform a model about *future* consumer preferences. Parameters calibrated from historical data could be fully valid over a future time period only if market conditions such as energy prices and the availability of technology options are consistent with historical conditions. Given that the model is used to simulate the effects of policies that would result in substantially different market conditions (a carbon tax would effectively change the price of carbon-intensive energy, for example), this limitation may be problematic. Still, using calibration to estimate behavioural parameter can supplement previous stated and revealed preference work and further improve the credibility of models like CIMS.

While static calibration is often used in energy-economy models to establish consistency in technological parameters in a base year, estimation of consumer behaviour by dynamic calibration has not yet been attempted for technologically rich models. One exception may be a study by Boonekamp (2005), which runs a bottom-up model for the Netherlands over a historical period to determine the effect of historical policies and estimate the historical impact of energy price on energy consumption. The study estimates heterogeneity parameters similar to the ν parameter in CIMS. This estimation is

not the main focus of the study, however, and no statistical estimation approach is applied. Calibration approaches have been applied more extensively and more rigorously, however, to models in other disciplines. In Chapter 2, I assess a range of specific calibration methodologies, and review their applications in natural, physical, and social science models as well as energy-economy models.

1.2.5 Sources of model uncertainty

While uncertainty associated with the behavioural parameters is central to this calibration study, other sources of uncertainty are relevant. These additional sources can pose challenges to calibrating a model using historical data and are worth a brief review

Tschang and Dowlatabadi (1995) recognize two distinct types of uncertainty in models: *parametric uncertainty* regarding the value of input parameters and *model structure uncertainty* pertaining to how the model maps inputs to outputs. Smith (2003) suggests that model structure uncertainty, also called model inadequacy, can be a more significant problem than parametric uncertainty.

Peterson (2006) and Kann and Weyant (2000) add to the list of model uncertainties with the idea of *stochasticity*, or uncertainty due to variability in the modeled system itself. In the case of energy-economy models, stochastic uncertainty could represent random, non-deterministic events such as the energy price spike of the 1970s.

Observational uncertainty can also be an important component of model uncertainty (Kennedy and O'Hagan, 2001). Observational uncertainty pertains to the uncertainty associated with the actual data on which a model is based. This is an important issue for calibration, a process dependent on empirical data.

Finally, Kennedy and O'Hagan (2001) also describe *code uncertainty* as the uncertainty associated with modelling complex systems using computer code. They suggest that if the model is complex, it cannot be tested with every possible configuration of inputs and that uncertainty about outputs due to unanticipated effects of undetected bugs must therefore be acknowledged.

All of these types of uncertainty exist in CIMS. The behavioural parameters v , i , and r are associated with significant parametric uncertainty. While other parameters, such as technology costs and fuel prices are also associated with uncertainty, the behavioural parameters are more difficult to estimate because they are not observable as unique components; only the aggregate effects can be observed in historical consumer choices. Model structure uncertainty is also an issue. Although CIMS models macro-economic feedbacks, applies a high level of technological detail, and aims for behavioural realism in modeling choices of consumers and firms (Jaccard, 2005a), it is not a perfect representation of the energy-economy system; the system is too complex for this to be possible. Stochastic uncertainty is also unavoidable, as CIMS will never be able to predict stochastic events and shouldn't try to, given the large range of potentially relevant random effects. Similarly, code uncertainty is an issue for CIMS as it becomes more complex. Testing the full range of possible inputs becomes impossible.

1.3 Objectives

The goal of this calibration study is to improve the credibility and usefulness of forecasts by improving the empirical foundation of the model's behavioural parameters. This calibration exercise is an attempt to inform the model's theoretical framework with a historical empirical framework. Given the issues of uncertainty and modelling behaviour described in this chapter, I define three specific objectives to frame this overarching goal:

1. **Estimate probability distributions for behavioural parameters through calibration to historical data.** To better model future behaviour, in this study I attempt to gain insight into historical consumer behaviour by estimating values for the CIMS behavioural parameters. Since historical data will be used, estimated parameters will be based on revealed preferences. Forecasts using these calibrated parameters should better simulate technology choices. Further, quantifying parametric uncertainty in the behavioural parameter by estimating probability distributions will allow me to better assess the parameters currently used in CIMS and compare results to previous stated and revealed preference work designed to estimate these parameters.

2. Explicitly incorporate uncertainty analysis into CIMS forecasting.

Uncertainty in parameter values can be explicitly incorporated into forecasts by statistically sampling from an estimated probability distribution. This approach could quantify uncertainty in forecast outputs, given the uncertainty in behavioural parameter inputs (as quantified through empirical calibration). Using explicit representations of uncertainty would help clearly communicate the potential range of impacts of policies and improve the credibility of model forecasts.

3. Understand key issues resulting in differences between historical time trends and model forecasts. Comparing model forecasts over a historical time-period to empirical historical data over this period can also lead to insights into dynamics in the energy-economy system not represented in CIMS. Analysing differences between historical and forecasted trends can allow me to qualitatively assess structural uncertainty in the model to complement the quantitative calibration of parametric uncertainty. I can then evaluate the importance of these effects and make recommendations to improve the model if they are necessary and practical.

1.4 Report structure

In Chapter 2, I review alternative methodologies used to calibrate energy economy models as well as other models of complex systems. I also present a statistical simulation methodology used to estimate behavioural parameters in CIMS; in Chapter 3, I present the results of this analysis. In Chapter 4, I apply the results of the calibration to policy analysis using the improved model to forecast the effectiveness of policy options while explicitly incorporating uncertainty analysis. Finally, I conclude in Chapter 5 by discussing the significance of the results of this study, recommendations for further improvements to CIMS, and a description of potential areas for future research.

CHAPTER 2: METHODOLOGY

2.1 Overview

Incorporating revealed historical consumer preferences into the CIMS model can improve the usefulness and credibility of the model as a forecasting tool. In this chapter, I develop a specific methodology designed to meet both this overarching goal as well as the specific objectives established in Chapter 1. In Section 2.2, I present the ideas of calibration and backcasting in general. Building on this big-picture foundation, in Section 2.3 I survey specific calibration methodologies used for other energy-economy models as well as for complex models from other disciplines. In Section 2.4, I present in detail the methodology used in this study, Markov Chain Monte Carlo (MCMC) parameter estimation. In Section 2.5, I conclude the chapter by discussing issues in implementing the MCMC calibration of behavioural parameters in the CIMS model.

2.2 Principles of calibration and backcasting

A CIMS simulation produces time series data consisting of technology market shares for each time period in the simulation. To calibrate the model over such a time period, the model must be run from an initial (historical) date, and then its outputs compared to real world historical data.

To provide a valid comparison point between historical data and CIMS outputs, a process known as *backcasting* allows modellers to forecast over a historical period. The model's initial conditions are reset to be consistent with a historical starting year, and the model is simulated forward through time. Rather than *forecasting* future trends, the model's outputs instead *backcast* trends over the historical period. Calibration then involves modelling different combinations of inputs and assessing how well the resulting model outputs correspond to real world data (Tschang and Dowlatabadi, 1995). Parameter values that result in model outputs that best match empirical data can then be determined. In the case of CIMS, I run the model over a period from 1990 to the present and compare

predicted market shares of selected technologies with historical market shares over that period.

The problem can also be reframed as what is known as the “inverse problem” (Tarantola, 2005). Essentially, model outputs (historical “observations”) are known, while some of the model inputs required to generate the given outputs are unknown and must be estimated. The inverse problem in this case, however, may be somewhat challenging given that multiple parameter combinations may result in the required outputs. Multiple solutions exist due to the parameterization of the CIMS market share function, as described in Equation 1. The behavioural parameters overlap in their representation of behaviour. Heterogeneity, v , partly refers to heterogeneity in how different consumers perceive non-financial costs represented by intangible costs, i . Risk associated with new technologies is also associated with temporal preferences, represented by the discount rate, r . A further issue arises because the market share function, as described in Section 1.2.3, calculates market shares for competing technology archetypes based on their lifecycle costs. Conceivably, multiple “good fit” combinations could exist for intangible costs: if for example, in a simplified competition between two technologies, A and B, if $i_A = \$1000$ and $i_B = \$4000$, the calculated market share might be similar as to the case in which $i_A = \$2000$ and $i_B = \$5000$, in which the relative difference is comparable.^{8,9}

Calibration through backcasting should thus provide values for the uncertain parameters that cause the model to behave in accordance with observed data. That is, these parameters allow the model to represent historical purchasing behaviour of firms and consumers over the historical period. Policy analysts can then use these improved

⁸ This example is artificial and clearly dependent on the capital and operating costs for each technology. Nonetheless, because the market share function depends on the *relative* lifecycle costs of competing technologies, multiple combinations of intangible costs could clearly replicate a given market share split.

⁹ It could be that multiple solutions might exist to the calibration inverse problem not only due to the parameterization of the market share function, but also because the multiple years of data available in a historical time trend provide essentially only one observation: each year of data is correlated with the previous. This limited number of observations suggests that this inverse problem is under-constrained (there are more parameters than informative data points), and that multiple (perhaps an infinite) number of solutions to the problem will exist. Applying calibration approaches has not shown this to be the case; supplemental analysis indicated the historical data matrix was indeed informative for parameter values. I nevertheless include this hypothesis here for completeness.

parameters in new forecasts. Future behaviour may differ from past behaviour. However, in support of previous estimates from stated preference discrete choice survey research, the calibration process should improve CIMS' representation of behaviour and improve CIMS forecasts because it relies on an empirical, revealed preference approach.

2.3 Review of complex model calibration approaches

Due to the breadth and complexity of the CIMS model, calibrating behavioural parameters from historical data is not a simple task. Specific approaches used to calibrate parameters in other complex models can provide insight into the critical issues involved. In this section, I survey calibration methodologies from a variety of disciplines (including statistics, economics, and the physical, natural, and social sciences) though I refer to energy-economy models wherever possible.

Calibration is essentially an attempt to quantify uncertainty using empirical (historical) data. While the focus in this analysis is primarily on *parametric* uncertainty in the behavioural parameters, the existence of multiple sources of uncertainty in a model such as CIMS, particularly *model structure uncertainty* has some important implications for estimation of behavioural parameters. A survey of calibration methodologies suggests that, in general, the simplest approaches are deterministic, more complex approaches attempt to explicitly account for parametric uncertainty, and the most complex approaches also try to manage other sources of model uncertainty. I review alternative approaches here moving from the simplest to the most complex.

2.3.1 Subjective parameter estimation

The simplest method of calibration is to use judgement or expert opinion to set parameter values and to check the validity of the values by comparing model outputs to empirical data. Even when a more rigorous calibration approach is applied, the calibration problem can be under-determined (there are more unknown parameters than observations) and can only be solved if some parameter values are exogenously determined. When econometric estimates are not available, these calibrations are often based on assumptions and judgement (Wing, 2006).

Subjective parameterization has the advantages of being simple to implement and easily incorporating the intuition of experts. The approach does not, however, explicitly incorporate empirical data into the model or account for uncertainty. Nevertheless, these “semi-quantitative” assessments of uncertainty can be useful in improving a model’s credibility (Peterson, 2006).

2.3.2 Partial calibration

Simple or “partial” calibration techniques are often used to estimate supply and demand parameters in computable general equilibrium (CGE) models (Bergman and Henrekson, 2003; IPCC, 2004; Bohringer, 1998). In these cases, smaller components of the larger model are calibrated to exogenous data. Similarly, in CIMS, *technological* parameters are statically calibrated using historical data from one time period, the base year of the simulation. Disaggregated data for technological parameters (e.g., costs, energy efficiency) are not always available at a high level of detail; modellers can, however, estimate annual energy consumptions of specific technologies by calibrating the simulated energy consumption to aggregate real world data for a given base year (Jaccard, 2005a). While this process is helpful in verifying technology cost parameters in the model, it is unrelated to improving parameters that represent consumers’ technology acquisition behaviour.

2.3.3 Deterministic optimization or model fitting

As an improvement over subjective or partial approaches, uncertain parameters can be more formally optimized to a single set of values such that model outputs best match empirical data. Liu et al. (2004), for example, calculate optimum values for elasticities of substitution in a CGE model. By comparing model outputs over a backcasting period with historical data, they generate a likelihood function and choose parameter values with the maximum likelihood.

Though it does not quantify or explicitly account for uncertainty, this approach is common due to its relative simplicity. Peterson (2006), in her survey of uncertainty in

energy policy models, suggests that few complex models quantify uncertainty, and even then, usually do so through ad-hoc, or guesstimated approaches.

Sometimes, uncertainty is addressed, though not quantified, through calibration.

Freedman et al. (2005), for example, calibrate non-linear parameters in a groundwater hydrology model using an optimization approach. To limit the scope of their analysis, they first use sensitivity analysis to determine which parameters have the greatest effect on outputs. They then replace the least sensitive parameters with constants and optimize the remaining parameters to best fit historical data. Uncertainty is handled separately from the calibration by comparing model predictions to a separate set of experimental field data and applying confidence intervals.

To calibrate a CGE model, Roberts (2004) also takes an ad-hoc approach to uncertainty. In this study, some parameters are optimized in a calibration process while others are calibrated subjectively. Model outputs are statically calibrated to a single year, which is assumed to be an equilibrium state. The analysts then implement a kind of sensitivity analysis; the process is repeated for alternate benchmark years to test sensitivity and robustness of calibrated parameter values. Again, however, this approach does not dynamically calibrate to the entire time series.

Warr and Ayres (2006) calibrate the Resource Exergy Services (REXS) model using American historical data spanning a century. Parameters are calibrated with a standard optimization algorithm; best fitting parameters minimize the root mean square of error between prediction and data. Interestingly, the authors also take an innovative approach to modelling temporal dynamics in consumer preferences: to correctly reproduce empirical time series, parameter values were allowed to shift over the backcasting period. Simulating a simple shift in consumer preferences, optimization methods were again used to identify the years in which parameter values should change.

2.3.4 Parameter estimation with uncertainty

Explicitly quantifying uncertainty in parameter values adds complexity to a calibration approach but credibility to model forecasts. Monte Carlo approaches can sample from probability distributions over parameter input values in order to quantify uncertainty in model forecast outputs. Parameter estimation involves calculating probability distributions or confidence intervals over parameter values, thus quantifying parametric uncertainty if model structure uncertainty can be ignored. If the model structure is assumed to be accurate, or close to accurate, all discrepancy between modeled outputs and empirical (historical) data is assumed to be due to parametric uncertainty (Balakrishnan et al., 2003; Smith 2003).

Balakrishnan et al. (2003) use a Bayesian statistical framework to estimate posterior probability distributions for parameter values in a ground-water hydrology model. In Bayesian statistics (an alternative to more conventional, frequentist statistics) posterior probability distributions expresses uncertainty in parameter values given both empirical data and prior knowledge regarding appropriate values.¹⁰ They apply a Markov Chain Monte Carlo (MCMC) approach and use a Metropolis-Hastings simulation to approximate the posterior without having to reduce the analytical complexity of their model. This is the approach I apply to the CIMS model in this study (see Section 2.4 for more details)

Walters and Ludwig (1994), Kuhner (1995), and Patz and Junker (1999) provide other examples of studies that explicitly quantify uncertainty in calibrated parameters. Walters and Ludwig (1994) explore the suitability of a range of parameter values using a simple grid-search to estimate posterior probability distributions for key fish population parameters. Kuhner (1995) calibrate a genetics model using a maximum likelihood framework while Patz and Junker (1999) take a Bayesian approach in calibrating a psychology model that quantifies attitudes from survey response data.

¹⁰ For a brief summary of the Bayesian statistical framework, see Appendix A.

2.3.5 Bayesian Monte Carlo updating

Applying an alternative Bayesian approach, Tschang and Dowlatabadi (1995) present a Bayesian updating technique for reducing uncertainty in the Edmonds-Reilly global energy model (a partial equilibrium model). Similarly, Dowlatabadi and Oravetz (2006) apply this approach to estimating parameters representing technological change in a top-down energy-economy model. Their approach uses observations of model outputs to filter out combinations of parameter values that do not match outputs. The approach is Bayesian because it takes into account initial uncertainty distributions for inputs (or “prior probability distributions”), as generated by expert opinion or other subjective approaches. Backcasting then improves the prediction quality of the model by combining these distributions with the results of a comparison with historical data to generate posterior probability distributions.

Tschang and Dowlatabadi (1995) also suggest that their calibration approach can account for both model structure and parametric uncertainty. They suggest that comparing outputs of competing model parameterizations, or functional structures, can be used to update the posteriors to include model structure uncertainty. Casman et al. (1999) further analyse the implications of combined model structure and parametric uncertainty in a Bayesian framework. As I will discuss in Chapter 3, interactions between structural and parametric uncertainty is an important issue for calibration.

2.3.6 Calibration under uncertainty

Some of the latest statistical literature on calibration builds on the ‘calibration under uncertainty’ approach pioneered by Kennedy and O’Hagan (2001).¹¹ This approach explicitly recognizes that the model to be calibrated is imperfect. A stochastic process (essentially a randomized error term), called a “discrepancy function”, is used to represent uncertainty in model structure in the comparison of observed and forecasted data. The statistical model thus quantifies both parametric and structural uncertainty as distinct elements. Calibration then uses multiple data sets to calibrate both parametric uncertainty distributions and the discrepancy function. Higdon *et al.* (2004) build on

¹¹ The classification “calibration under uncertainty” was assigned by Trucano *et al.* (2006).

Kennedy and O'Hagan's (2001) approach by applying a Markov Chain Monte Carlo (MCMC) simulation to explore the posterior distributions for parameters in both engineering and nuclear physics models.

Challenor (2002) lays out the framework for an initial attempt at applying this Bayesian calibration under uncertainty approach to an integrated assessment energy policy model. This effort appears to be the first research aiming toward applying Kennedy and O'Hagan's (2001) statistical approach to calibrating an energy-economy model.

2.3.7 Sequential experimental design

A sequential experimental design is a calibration approach drawn from new, methodological research being developed in the Simon Fraser University statistics department. This methodology specifically targets the challenge of the existence of multiple combinations of behavioural parameters that would allow the model to reproduce historical data trends.

The sequential experimental design methodology (Ranjan *et al.*, 2007) deals with this challenge explicitly by outputting a *contour*. The contour through parameter space specifies *all* behavioural parameter combinations that result in a best-fit with historical data, explicitly recognizing the potential for multiple solutions. Uncertainty bounds around the contour can then quantify uncertainty for the infinite number of combinations represented by the contour. This approach is more sophisticated than simple parameter estimation which quantifies uncertainty around only a single best-fit parameter combination. The methodology achieves this goal through minimizing the number of trial parameters combinations that must be simulated by optimizing the expected value of a candidate point before running the forecast for this combination of parameter values. The approach therefore generates the contour with a reasonable number of model runs.

2.4 Bayesian Markov Chain Monte Carlo (MCMC) theory

Given the variety of possible approaches available, what then is the best methodology for calibrating CIMS through estimation of behavioural parameters? In developing a

methodology, a balance must be achieved between outputting informative results and moderating the complexity required for implementation. The research objectives outlined in Section 1.3 emphasize the importance of explicitly accounting for uncertainty, so the simplest approaches described in Section 2.3 provide insufficient detail to meet my objectives. On the other end of the spectrum, implementing a new methodology such as Sequential Experimental Design is outside the scope of this analysis. Similarly, the limited data available (particularly due to the time correlation between successive historical data points) inhibits the application of Kennedy and O'Hagan's (2001) approach to accounting for model structure uncertainty. As a compromise between simple implementation and informative results, this study will therefore apply a Bayesian Markov-Chain Monte Carlo (MCMC) parameter estimation approach.

MCMC has been evaluated in detail in statistical literature and in biological, physical, and social science applications. Energy-economy models have not yet applied the MCMC approach to parameter estimation. Metropolis-Hastings is one, quite versatile, example of an algorithm implementing MCMC. MCMC, and Metropolis-Hastings in particular, has several advantages: it is relatively simple to implement, it is capable of handling complex models, it allows the analyst to explore complex posterior distributions that might have multiple, high probability, local maxima (Gelman *et al*, 2004). Stochasticity in the algorithm's "pseudo-random walk" provides the algorithm with robustness against getting "stuck" at local maxima or saddle points (Denison *et al*, 2002). Further, the Bayesian nature of the approach makes it possible to incorporate the results of past studies or of expert opinion into the analysis in the form of prior probability distributions. For all of these reasons, this study applies the Metropolis-Hastings algorithm to the calibration of CIMS.

I will first describe a generalized Metropolis-Hastings approach, before customizing this algorithm for implementation in calibrating CIMS in Section 2.5. Figure 4 illustrates the process flow for code that implements a Metropolis-Hastings algorithm through the following steps (adapted from Denison *et al.*, 2002; Gelman *et al.*, 2004; Tanner, 1996; and Walters and Ludwig, 1994):

1. Select initial starting parameters, $\theta_{(s=0)}$, where s is the MCMC iteration and θ is a vector of the behavioural parameters (*i.e.*, v , i_I through i_K , for K competing technologies, and r).
2. Run the CIMS model with parameters θ_0 to generate forecasted market shares over the historical period.
3. Compare historical and forecasted data; calculate likelihoods at each historical interval $L(y | \theta)$ given historical market share data, y . A distribution shape (e.g. normal, log-normal, etc.) can be selected for the likelihood function.¹²
4. Combine the likelihood with prior probability distribution, $p(\theta)$ (See Section 2.5.5 for more details regarding the prior), to calculate an unscaled posterior density $p(\theta_0 | y)$:¹³

$$p(\theta_0 | y) = p(\theta) * L(y | \theta_0)$$

Equation 2

5. For $s = 1, 2, 3 \dots S$, where $S =$ the number of MCMC iterations.
 - a. Sample a new, *proposal point* θ^* from a “jumping distribution”. For this analysis, a normal distribution was used for a jumping distribution. The jumping function represents a “pseudo-random walk” algorithm because it is stochastic, but depends on the previously sampled point: the proposal

¹² Although using a likelihood function is the typical approach in an MCMC algorithm, in this study, I take a slightly different approach, as described in Section 2.5.3. However, I describe the basic approach here first in order to better make clear the nature of the MCMC algorithm.

¹³ “Unscaled” indicated that the integral of the posterior density over the full range of values does not sum to one. Scaled posterior probability densities could be calculated by normalizing the unscaled densities by this integral.

point is a random step through parameter space away from the previous point, θ_{s-1} .

- b. Run the CIMS model to calculate the unscaled posterior density $L(\theta^* | y)p(\theta^*)$ as in steps 2-4.
- c. Assess the ratio, R , of the posterior likelihood for θ_{s-1} and θ^* normalized by the jumping function:

$$R = \frac{p(\theta^* | y) / J_s(\theta^* / \theta_{s-1})}{p(\theta_{s-1} | y) / J_s(\theta_{s-1} / \theta^*)} \quad \text{Equation 3}$$

Where:

- $p(\theta^* | y)$ = the posterior probability density of a proposal parameter combination given historical data, y
- θ^* = the proposal point; a vector $(v, i, r)^*$
- $J_s(\theta^* / \theta_{s-1})$ = the probability of the random walk (jump distribution) walking to parameter combination θ^* given the previous point in parameter space, θ_{s-1} .
- $p(\theta_{s-1} | y) / J_s(\theta_{s-1} / \theta^*)$ = the normalized posterior probability density of the previous parameter space point in the MCMC iteration.

- d. If the jump increases the posterior likelihood, set $\theta_{s+1} = \theta^*$. If the jump decreases the posterior likelihood, set $\theta_{s+1} = \theta^*$ with a probability of the ratio calculated in step (c), and otherwise, set $\theta_{s+1} = \theta_s$.
6. If the MCMC algorithm has implemented a sufficient number of iterations, s , the frequency distribution of iterations of θ_s approximates the joint posterior probability distribution over all behavioural parameters.

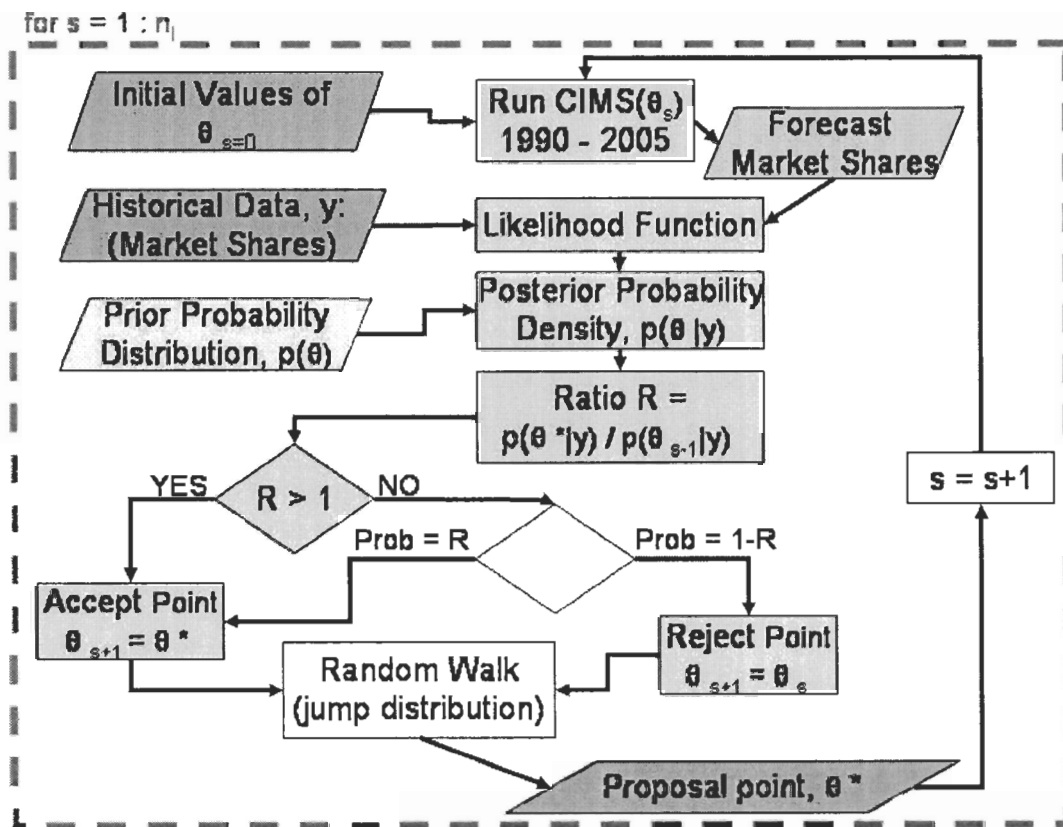


Figure 4: Process flow diagram for conventional Metropolis-Hastings Markov Chain Monte Carlo parameter estimation simulation

2.5 Implementing Metropolis-Hastings for CIMS

Applying the principles of the Metropolis-Hastings algorithm described above to the calibration of CIMS requires making some choices specific to this study. These choices

are important for customizing the implementation for CIMS and ensuring the calibration is computationally tractable.

2.5.1 Scoping the analysis

Because each competing technology in CIMS is associated with an individual intangible cost, calibrating the complete CIMS model – consisting of thousands of technologies – is clearly impractical. Rather, this study will calibrate a few key technology competitions (nodes) as indicators for other nodes in CIMS. Four criteria were applied to select representative nodes:

1. Availability of data: availability of detailed historical market share data limits which nodes can be calibrated.
2. Energy / emissions share: nodes that use larger amounts of energy and emit more greenhouse gas emissions were targeted so as to be more relevant to CIMS when it is run in its entirety.
3. Compatibility with existing CIMS structure: complex and data-intensive nodes were avoided for ease of application of the MCMC methodology and to reduce data demands.
4. Significance of behavioural parameters: calibration is particularly relevant in refining parameters specific to technology decisions for purchasing decisions that might not be explained by simple cost minimization

Given these criteria, this study calibrates the following nodes:

1. Refrigerators (residential sector)
2. Furnaces (residential sector)
3. Gasoline vehicles (transportation sector)

While it would have been useful to estimate parameters for a node in the industrial or commercial sector to compare the preferences of firms to those of individual consumers or households, obtaining sufficiently disaggregated market share from these sectors proved too difficult.

2.5.2 Matching historical and model frameworks

These three nodes were modelled separately using code that replicates the CIMS algorithm.¹⁴ These individual CIMS nodes allow for greater flexibility in slightly adjusting the configuration of the full CIMS model to better match available data, and make running MCMC statistical calibration simulations (requiring thousands of model-run iterations) possible.

Cooley's (1997, p. 58) "rules for calibration" recommend "matching the measurements to the model and matching the model to the measurements." Cooley suggests that an important part of calibration is aligning the theoretical framework (represented by the model) with a real economy (represented by data) and vice versa: a common reference point is required between model forecasts and data. For CIMS, this reference point is forecasted market shares of specific technology types. Depending on the data available, the backcast models can calculate either new market shares (matching up to sales data) or full stock turnover (matching up to stock survey data).

In terms of designing the separate backcast models, tradeoffs exist between satisfying Cooley's (1997) principles and pragmatism. The better the structure of the separate nodes corresponds with that of the full CIMS model, the more informative and applicable to the full model are the calibration results. On the other hand, if data cannot be manipulated to match the structure of the existing model, pragmatism dictates that the structure of the calibration node models must be adjusted to match the available data.

For example, in the refrigeration node, the full CIMS model calculates market shares for arbitrary refrigerator archetypes based on energy consumption (standard, 10%, 20%,

¹⁴ CIMS backcast nodes were coded in the R programming language, as was the MCMC calibration code.

30%, and 40% more efficient). Over a historical period, however, for a simulation based on these archetypes (which were designed for a base year of 2005), almost all historical refrigerator sales fall into the “standard” archetype. A calibration with these archetypes therefore does not output meaningful results. To better represent the historical preferences, a different range of arbitrary archetypes was used in the individual calibration node. While the calibrated intangible costs of these archetypes will not directly translate to the archetypes in the full model, they nevertheless do provide more insight into historical refrigeration purchase preferences. Similarly, in the furnace node, ground source heat pumps are grouped with air source heat pumps and wood furnaces are omitted due to lack of historical market share data for these technologies.

Consistent with Cooley (1997), Dowlatabadi and Oravetz’s (2006) backcasting study provides specific suggestions for establishing consistency between the model framework and historical data. First, the start year of the model is reset to the beginning of the historical period, 1990, establishing a common initial condition between model and history. Second, all known, observable parameters, such as capital and operating costs, are set to their historical values. Historical energy prices replace exogenous forecasts used in CIMS forecasts, and the macro-economic equilibrium component of CIMS is not included; no price adjustment is required for equilibrium as historical energy prices are available. Though any relevant policies implemented over the historical period should also be represented in the backcast, initially I include no policy effects over the historical period, as Canada mostly relied on ineffective subsidy or voluntary programs (Jaccard *et al.*, 2003; Jaccard *et al.*, 2006). I do, however, revisit this assumption in Chapter 3. Third, the time steps of the simulation were reduced to increase the resolution at which parameters can be calibrated. Rather than calculate market shares or stocks every five years, the individual node models calculate market shares annually, since annual historical data are available. Finally, endogenous processes are replaced with historical data. In some cases, for example, declining capital costs (as a result of learning) can be determined from historical cost data. Early analysis in the residential furnace node also indicated that the linear base-stock retirement function used to retire technology stock in the base year was not an accurate representation of historical trends. Since the retirement

function parameters are not part of the set of parameters being calibrated, however, this function was eliminated and the endogenous stock decline was replaced with the historical trend.

On a related note, Hilborn and Mangel (1997), in the ecological modelling literature, differentiate between calibrating for process uncertainty and observational uncertainty. While process uncertainty refers to the uncertainty in parameters representing processes of interest – such as the behavioural parameters in CIMS representing the process of technology choice – observational uncertainty refers to uncertainty in data observation. The distinction is important in the calibration of the residential furnace node, as described in Section 3.3, in which total stocks dependent on stocks in the previous year are calculated. This study is interested in quantifying what Hilborn and Mangel (1997) refer to as process uncertainty, a combination of the parametric and model structure uncertainties. Ideally, a process-uncertainty calibration approach would calculate stocks in year t based on the *historical* data in year $t-1$. Because CIMS requires data on the vintage of stock in year $t-1$, this approach is not possible. Instead, the model calculates stocks in year t based on the *forecasted* stocks in the previous year $t-1$. This approach is more in line with calibrating for observational uncertainty. Note, however, that by replacing endogenous trends such as the base-stock decline with historical data, the model and data frameworks are aligned as closely as possible, thus mitigating some propagation of error.

2.5.3 Selecting likelihood and posterior functions

Having matched data and model frameworks, a likelihood function (as represented by $L(y | \theta)$ in Equation 2), is then required to quantify differences between forecasts and historical market shares or stocks. However, selecting a likelihood function also introduces additional parameters (in the case of a normal distribution, the variance) that must also be estimated in a formal parameter analysis. To avoid the complexity of additional dimensionality in the parameter estimation, Walters and Ludwig (1994) replace the posterior with an expression that is independent of these “nuisance

parameters,” representing a marginal probability integrated over the variance. The posterior used in the MCMC algorithm (Equation 4) parallels this work.

$$-\log(p(\theta | y)) = 0.5 * K * n_T * \sum_{k=1}^K \left(\log \sum_{t=1}^{n_T} (P_{k,t} - O_{k,t})^2 \right) - \log(p(\theta)) \quad \text{Equation 4}$$

Where:

- $p(\theta | y)$ = the posterior probability of parameters θ , given data, y
- K = the number of technology archetypes competing
- n_T = the number of time periods simulated in a backcast
- $P_{k,t}$ = the predicted, or backcasted market share of technology k , in time period, t
- $O_{k,t}$ = the observed, or historical market share of technology k , in time period, t

The expression in Equation 4 modifies the process flow of the conventional Metropolis-Hastings algorithm described in Section 2.4. Instead of calculating the likelihood of the data using a standard deviation (that must also be estimated) and combining it with a prior distribution to estimate posterior probability, Equation 4 calculates the posterior independent of a standard deviation. This expression thus effectively replaces steps 3 and 4 of the process flow. Note that the expression calculates the negative log of the posterior. As a result, to include the prior distribution $p(\theta)$, the negative log of the prior is added to the function rather than multiplied. The modified algorithm is represented in Figure 5, as adapted from Figure 4.

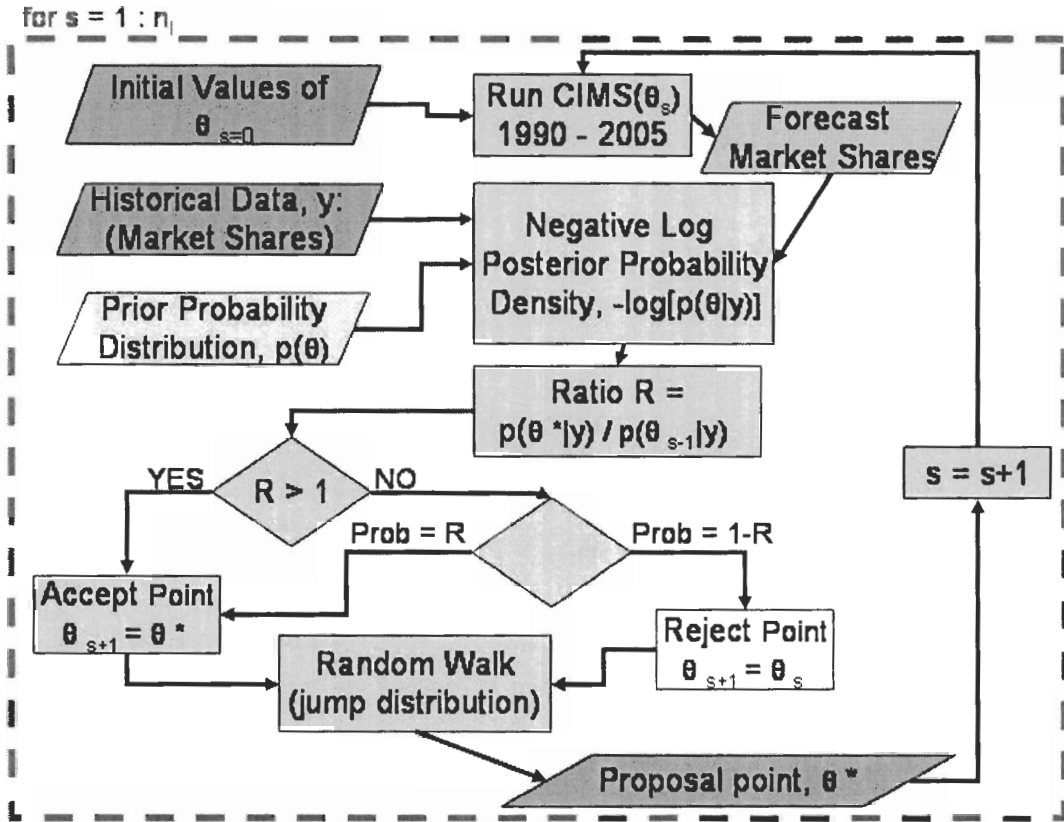


Figure 5: Process flow diagram for modified Metropolis-Hastings Markov Chain Monte Carlo parameter estimation simulation

2.5.4 Constraining parameter space

To facilitate parameter estimation, parameter values are limited to realistic values. A log transformation ensures positive values for the v parameters and a logit transformation ensures values for the r parameter between zero and one when it is included in the calibration. Negative intangible costs are permitted, so the parameter i is not transformed. Thus parameter space Θ , through which the Metropolis-Hastings algorithm steps, actually represents these transformed parameters.

2.5.5 Informative prior probability distributions

Because the MCMC algorithm I have described functions in a Bayesian statistical framework, the estimation of a posterior probability uses prior information about parameter values in the form of prior probability distributions (often referred to simply as

“priors”, and represented as $p(\theta)$ in Equation 2 and Equation 4). In this study, I use informative priors based on expert opinion as well as previous stated and revealed preference studies. Standard deviations for the distributions were set so as to result in a mildly informative prior that did not dominate the posterior, allowing the likelihood function derived from the data to remain as the predominate factor in the calculation of the posterior probability distribution. See Appendix A for additional detail.

In addition to incorporating existing prior knowledge regarding parameter values, applying informative priors is also important for the convergence of the MCMC algorithm. The calibration algorithm does not converge using non-informative priors (a uniform distribution). This behaviour may be a result of the existence of multiple parameter combinations that output backcast data that match historical trends. As I have discussed in Section 2.2, multiple combinations might exist because the market share function is over-constrained. Informative priors, however, focus the calibration algorithm on a specific region of parameter space and allow it to approximate the posterior probability distribution around a single parameter combination.

2.6 Methodological summary

A Markov Chain Monte Carlo (MCMC) approach to calibration as described above, can meet the objectives for this study. MCMC is an established and versatile method of parameter estimation and can generate probability distributions for parameters from historical data, allowing for explicit incorporation of uncertainty into policy analyses. By matching model and historical frameworks for three CIMS technology competition models, parameter distributions can incorporate historical revealed preferences, which improves the behavioural realism of the model. Finally, assessing the “reasonableness” of estimated parameter distributions can lead to insight as to the model’s successes and failures in replicating historical technological preferences.

CHAPTER 3: ANALYSIS AND DISCUSSION

3.1 Overview

To estimate behavioural parameters in CIMS, I apply the methodology to each of three CIMS nodes, or technology competitions: refrigerators, residential furnaces, and gasoline vehicles. In this chapter, I evaluate posterior probability distributions of the behavioural parameters for each technology competition, justify their validity, and assess possible drivers of historical behaviour over the calibration period that might not be captured by the model's existing structure.

3.2 Refrigerator node

The residential refrigeration node competes refrigerator technology archetypes of varying efficiencies. For this analysis, four archetypes were competed in the model: “low efficiency,” “medium efficiency,” high efficiency,” and “super efficiency” refrigerators.

¹⁵While these are not the archetypes used in the full CIMS model, this range of efficiencies does suitably illustrate the spectrum of technological changes from 1990 – 2005 as evident from historical data. The node was calibrated using sales data from British Columbia generated by Natural Resources Canada (Canada, 2005a). For calibration, sales data were compared to forecasts of new market share.

3.2.1 Dynamic historical refrigerator market share trends

The MCMC algorithm outputs what appear to be generally reasonable posterior distributions for the refrigerator behavioural parameters. Figure 6 shows approximated marginal posterior probability densities for refrigerator behavioural parameters

¹⁵ Note that “high efficiency” refers to a refrigerator that has a very good energy performance and uses a small amount of energy per volume of refrigeration. Conversely, a “low efficiency” refrigerator has relatively poor energy performance and thus uses a large amount of energy per volume of refrigeration.

(excluding r , which was fixed at 0.6).¹⁶ The y-axis on the histograms reflects the frequency of particular parameter values in the posterior sample (reflecting the frequency in which the calibration algorithm accepts a parameter value combination). Since the calibration algorithm tends to climb toward parameter values that provide the best fit with historical data, “frequency” is a measure of the unscaled probability density of a given parameter value. Calibration actually outputs a single, joint probability distribution over all parameters calibrated. Because this multi-dimensional distribution is difficult to present graphically, I instead present the marginal distributions (showing a probability of one parameter integrated over the other parameters).

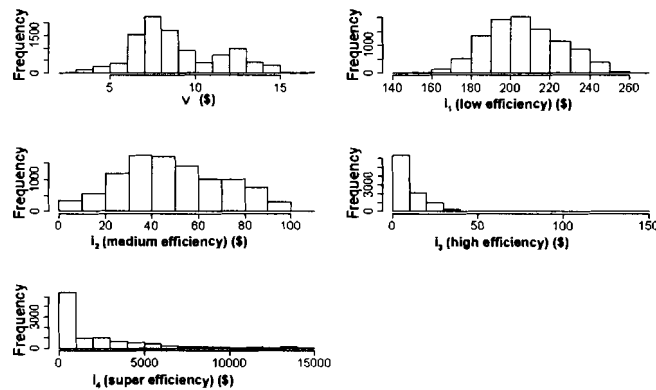


Figure 6: Marginal posterior probability distributions for refrigerator parameters, heterogeneity, v , and intangible costs i_1 through i_4 , for low, medium, high and super efficient refrigerators respectively.

However, further analysis reveals that the calibrated parameters do not result in a “good fit” between history and model outputs. Figure 7 compares historical market shares to a model backcast in which the behavioural parameters are set to the modes of the marginal distributions (the value for each parameter with the highest probability, or frequency, in the marginal distributions in Figure 6). The figure illustrates that the CIMS new market share calculation is incapable of replicating the dynamics shown in historical data. CIMS

¹⁶ This value for the discount rate, r , is consistent with Train’s (1985) survey of several studies which found consumers’ discount rates for refrigerators to be in the range of 0.39 to 1.00. 0.6 is the value currently used for refrigeration in the full CIMS model.

calculates forecasted market shares as a function of the life cycle costs of competing technology archetypes, which in turn are calculated from a technology's capital, operating, energy, and intangible costs. Because capital and operating costs were considered static, only historical changes in fuel costs could result in variation across time in new market shares. Further, because capital costs are very large relative to refrigerator operating costs, and the electricity price fluctuation is relatively small, the life cycle costs of the archetypes and the forecasted new market share are effectively static. Calibration estimates behavioural parameters that match the *average* of the historical time trend without replicating the details of these trends.

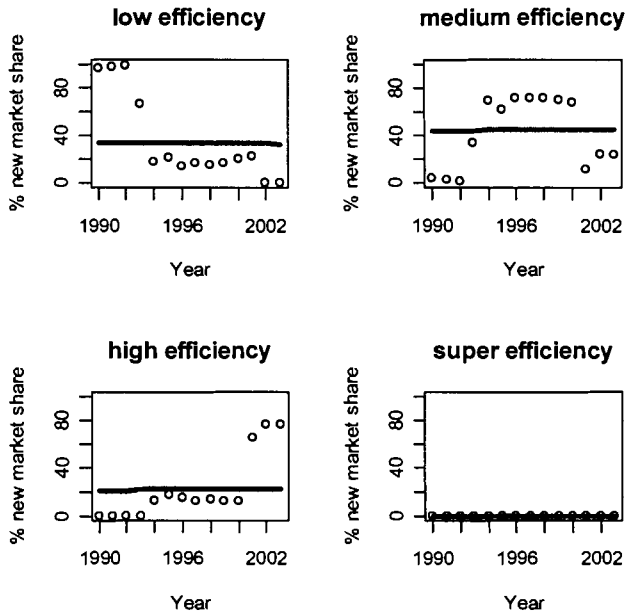


Figure 7: Historical (points) and forecasted (lines) new market shares (forecasted trajectories are generated by running the model on the most probable parameter combination as estimated by calibration)

Clearly, this issue is a problem both for calibration and for forecasting refrigerator choices into the future. CIMS can represent dynamics in the energy system through capital stock turn over: over time, even if the forecasted market share of new fridges are constant, the retirement of older, inefficient fridges will result in the gradual penetration of higher efficiency models. The historical data, however, suggests dynamics in the new

market shares, not just total stock. The inability of the model to replicate this dynamic trend, no matter what behavioural parameter values are used, suggests that CIMS is missing some important element in forecasting refrigeration market shares over the 1990 – 2004 period. Though this issue makes calibration problematic, assessing the factors driving this discrepancy is valuable in determining the effectiveness of CIMS’ forecasting approach.

3.2.2 Declining capital costs and autonomous energy efficiency increases

One possible factor causing the historical change in refrigeration preferences could be declining capital costs of higher efficiency refrigerators. Due to “learning by doing” (Arrow, 1962; Rivers and Jaccard, 2006) the costs of manufacturing a new technology tends to decrease as more units are produced and cheaper manufacturing techniques are developed. For the sake of simplicity, declining capital costs were not included in the calibration analysis, but this dynamic effect could help explain the discrepancies between new market shares and historical trends.

Consumer Reports (1992; 1994, 1995; 1997; 1999; 2002; 2004) provides useful insight into historical costs of residential refrigerators. Table 1 overviews costs of “freezer-on-top” refrigerator models from most major models sold in the United States. Costs have been adjusted to 1992 US Dollars through historical consumer price indices.

Table 1: Overview of Consumer Reports refrigerator cost data (freezer-on-top models only)

		1992	1994	1997	2002
Lowest energy consumption available	Annual Energy (kWh)	865	723	638	411
	Capital Cost (\$)	610	770	800	750
	Capacity (cu. Ft)	13.8	-	22	18.8
	Capital Cost in 1992 \$	<i>610</i>	<i>729</i>	<i>699</i>	<i>585</i>
Highest energy consumption available	Annual Energy (kWh)	1225	1072	963	523
	Capital Cost (\$)	500	550	1000	650
	Capacity (cu. Ft)	14.8	-	25	21.7
	Capital Cost in 1992 \$	<i>500</i>	<i>521</i>	<i>874</i>	<i>507</i>
Typical energy consumption model	Annual Energy (kWh)	1000	800	650	450
	Capital Cost (\$)	550	800	800	600
	Capacity (cu. Ft)	14.5	-	21- 22	18 - 22
	Capital Cost in 1992 \$	<i>550</i>	<i>757</i>	<i>699</i>	<i>468</i>

This cost data demonstrates that incorporating dynamic capital costs into the calibration is not practical. Though capital costs *did* change with time, as the *Highest* and *Lowest* capital cost trends illustrate, the rapid and substantial shift toward lower efficiency fridges over the time-period cannot be easily represented in the model. Indeed, the fridge with the lowest energy consumption available in 1997 uses more energy than the fridge with highest consumption in 2002. In this shift toward more efficient refrigerators over time, the data thus seem to illustrate what a top-down model would call *autonomous energy efficiency increases* (AEEI). In a technologically explicit model such as CIMS, however, the drivers behind such increases should be endogenous to the model.

The Consumer Reports data thus show clear changes in the range of availability of refrigerators. Consumers may not be choosing more efficient fridges over less efficient ones; the choice may instead be being made by the manufacturer. Alternatively, this shift could be a response from manufacturers to changing consumer preferences; determining definitively whether the effect is manufacturer or consumer driven is difficult.

Nevertheless, part of the reason CIMS fails to show appropriate dynamics is the rapid shift in the range of commercially available technologies. This effect cannot be easily represented in the model. Simply phasing out the least efficient models every year through an exogenous retirement function, and having a larger range and number of fridge archetypes that become available or unavailable as time passes seems arbitrary and neither particularly useful nor practical to implement.

3.2.3 Changing services

Table 1 only displays data for the standard “freezer-on-top” refrigerator model, as it is the most common fridge and the most consistent model through the historical time period. Nevertheless, the top-freezer model data does not clearly illustrate the other changes in the *service* provided by fridges. Each subsequent Consumer Report publication (1992; 1994, 1995; 1997; 2002) documents the proliferation of additional features available with refrigerators, including water dispensers, ice makers, meat drawer temperature controls, freezer lights, butter softeners, and digital displays. The 2002 report even comments on “speed chilling” features and faster ice-making capabilities. Similarly, other fridge types,

such as the less efficient but more convenient side-by-side and freezer-on-bottom models, have also significantly increased their market shares (Canada, 2005a). Finally, as evident from Table 1, and supported by Natural Resources Canada (Canada, 2005a), average refrigerator size increases over the period as well, partially offsetting gains in efficiency.

All of these trends seem to indicate that efficiency and operating costs do not play a major role in consumer choices of refrigerators. Manufacturers seem to be competing for sales by providing additional *services* alongside food refrigeration. This shift may reflect a kind of rebound effect; savings from energy efficiency have been coupled with additional energy expenditures on additional services. These effects will not be captured by CIMS and thus could contribute to the failure of the calibration of the refrigerator node.

3.2.4 Effects of policies

Another possible explanation for the observed AEEI, and thus for the dynamics in historical new market shares, is that utility and government policies between 1990 and 2004 might indeed have had a significant effect on consumers' choices. The following effects of several government and utility programs may influence the trends in the historical cost data in Table 1:¹⁷

- **In 1993, the US Department of Energy mandated a minimum increase in refrigerator efficiency of 30%.** The effect of this regulation is apparent in the changes in available fridges from 1992 to 1994 (Consumer Reports, 1992; 1994): energy consumption was reduced considerably, but capital costs actually went up as manufacturers were forced to meet the standard. The efficiency increase was achieved through several key technological changes: high efficiency compressors, thicker insulation, and better door seals, which were all implemented at this time (Consumer Reports, 1994).

¹⁷ Note that this analysis focuses on American regulations and programs as they are likely driving Canadian manufacturing decision-making given the size of the American market:

- **Fridges using CFCs (chlorofluorocarbons) were banned in 1995 (1994 fridges were the last generation to use Freon).** HFCs (hydro-fluorocarbons) were effectively a perfect substitute for CFC refrigerants; the new fridges provided similar efficiencies to their predecessors and almost indistinguishable service (Consumer Reports, 1997). Capital costs, however, increased for refrigerators as manufacturers altered designs to meet the regulation. The regulation thus may explain the jump in capital costs evident from 1994 to 1997 in the representative data in Table 1.
- **A group of major utilities sponsored the “Golden Carrot” program in 1993: they provided a \$30 million prize for a fridge that was CFC-free and 25% more efficient.** The winning design by Whirlpool became commercially available in 1995. Other manufacturers soon followed suit (Consumer Reports, 1995). This program may be largely responsible for the efficiency changes from 1994 – 1997. It perhaps also sparked a continuing competition for more efficient fridges between manufacturers responsible for the continuing (and dramatic) efficiency improvements.

These policies may provide partial explanations for the growth in efficient refrigerators, and the lack of availability of inefficient models between 1990 and 2004. I have not evaluated, however, the evolution of refrigerators in previous time-periods. Without determining whether an autonomous energy efficiency increase also existed in previous periods (in which no policies existed), I cannot confirm that the policies described above were the cause of the increase of efficient refrigerator market shares.

3.2.5 Dynamic intangible costs

Finally, my assumptions regarding technological change and dynamic parameters should also be re-evaluated. Changing preferences, modelled as dynamic intangible costs could also explain the dynamism in the historical trends. The declining intangible cost function developed by Mau (2005), for example, could be relevant if consumers felt that high-efficiency refrigerators became more desirable as more high-efficiency refrigerators

penetrated the market. However, since dynamic intangibles are more difficult to estimate through calibration, given the additional parameters involved, a “quick and dirty” approach can give some sense of how dynamic intangibles might allow the model to reproduce historical trends. By running an optimization routine for *each* year of the historical run (with ν and r fixed at 10 and 0.6 respectively), the dynamic intangible costs required for the model forecast to duplicate the historical trend can be calculated. Essentially, the freely changing intangible cost can be treated as a “fitting” variable for each year of the simulation. Figure 8 shows these calculated dynamic intangible costs.

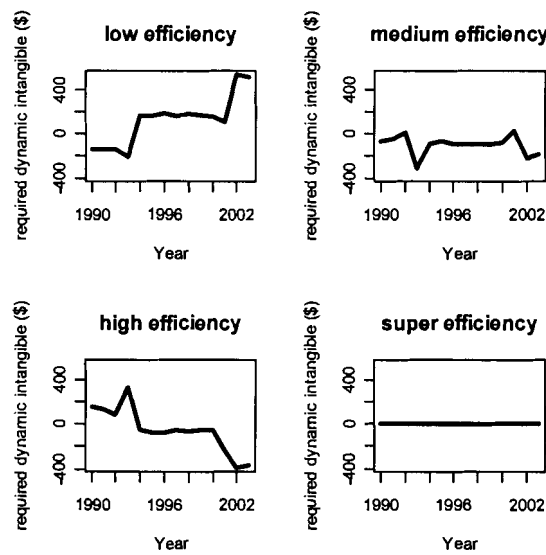


Figure 8: Calculated dynamic intangible costs for refrigerator archetypes at $\nu = 10$ and $r = 0.6$

As Figure 8 indicates, the historical shift toward efficient refrigerators could be replicated in a model backcast if the intangible cost of low efficiency refrigerators increased and the intangible cost of high efficiency refrigerators decreased. This analysis thus supports the idea that a declining intangible cost such as the “neighbour effect” might be relevant historically. Conceptually, however, the neighbour effect does not seem relevant in this case: more efficient refrigerators are not a new, untested technology that consumers might initially perceive as “risky” investments until they penetrated the market. To a

consumer, a high efficiency refrigerator provides much the same service as does a low efficiency model.

Since for this analysis, capital costs were maintained at the previous, static values, the calculated “fitting” intangible could really reflect required changes in capital costs. The “fitting” intangible refers to the optimized intangible cost for each technology for each simulation time period that is required for model forecasted market shares to match historical market shares. Note especially the jump in required cost in 1993: at this time-period, historical capital costs spiked due to regulation. The dynamic curve might therefore provide support to the hypothesis that policies enacted in the 1990s affected the diffusion of high efficiency refrigerators. Still, we should be careful about how we interpret these dynamic intangible cost curves; the “optimized” curves could be influenced by noise in other model parameters such as prices and costs as well as the availability of each archetype in a given year. Nevertheless, the “calculated dynamic intangibles” do give an indication of how changing capital costs or changing non-financial preferences might also explain the discrepancy between model forecasts and historical trends.

3.2.6 Refrigeration node summary

Overall, calibrating behavioural parameters in the refrigeration node is not possible given the model’s inability to reproduce dynamic historical trends without dynamic capital costs. Dynamic capital costs cannot be incorporated into the calibration because an apparent autonomous energy efficiency improvement results in rapid changes in the availability of CIMS refrigerator archetypes over the period of 1990 – 2005.

Still, analysis of the drivers of changing capital costs and efficiency improvements provide interesting insights into the performance of CIMS. Changes in refrigerators since 1990 may not be driven by consumers choosing more efficient fridges, but rather by manufacturers supplying only lower consumption fridges. This supply-side effect may be attributable to government regulations as well as utility demand side management programs (namely the Golden Carrot program of 1993) specifically aimed at

manufacturers. The apparent efficacy of these programs is interesting and bears further study and more quantitative analysis. To truly assess whether the programs accelerated the penetration of more efficient refrigerators, previous time periods would have to be evaluated to determine if a similar autonomous energy efficiency improvement existed in periods without policies, driven only by technological change in a competitive market. Further, since efficient and inefficient fridges provide essentially the same service, consumers may be more likely to base their decision on other factors: additional service features, greater refrigeration capacity, or convenience. CIMS currently does not account for expanding or changing energy services, though new research currently in progress will attempt to account for this effect in CIMS (Groves, unpublished).

3.3 Furnace node

The residential furnace node competes technology archetypes that provide residential heating through forced air distribution systems (electric baseboard heaters, for example, are competed in a separate node). For this analysis, I considered six archetypes: low and medium efficiency oil furnaces, low, medium, and high efficiency natural gas furnaces, and heat pumps. These archetypes are similar to those used in the full CIMS model, with two differences. First, the “high efficiency oil furnace” archetype is omitted due to lack of availability of this technology over the historical period. Natural Resources Canada (Canada, 2005b) data indicates a zero market share for this technology from 1990 to 2007, and a cursory survey of oil furnace vendors suggests that no product is available with efficiencies greater than 95% to match the archetype characteristics. Second, due to limitations in available data, a single aggregate “heat pump” archetype was used in lieu of CIMS’ current disaggregation of ground source and air source heat pumps.

The node was calibrated using data for Ontario, as generated by Natural Resources Canada (Canada, 2005b). For calibration, *total stock* data were compared to forecasted stocks, and thus (unlike for the refrigerator calibration) I implemented the CIMS stock turnover model to backcast historical stocks rather than new market shares.

3.3.1 Posterior probability distribution for residential furnace parameters

The calibration algorithm approximated reasonable probability distributions for uncertain behavioural parameters in the residential furnace node, as presented in Figure 9. Distributions were not estimated for intangible costs of the low efficiency oil furnace and the low efficiency natural gas furnace. These parameters are excluded from the calibration because these two technology archetypes were modelled as base-stock with no new stock purchased over the period (see Section 3.3.5), and behavioural parameters do not affect the retirement of base stock. Similarly, the discount rate, r , was not estimated because including this parameter added complexity to the calibration and made the convergence of the MCMC algorithm problematic. Further, economists are somewhat confident about the value of the private discount rate. Train (1985) surveys several sources to suggest a revealed discount rate range of 0.044 to 0.36 for space heating. For this analysis, r was fixed at 0.3, to match the value currently used in CIMS.

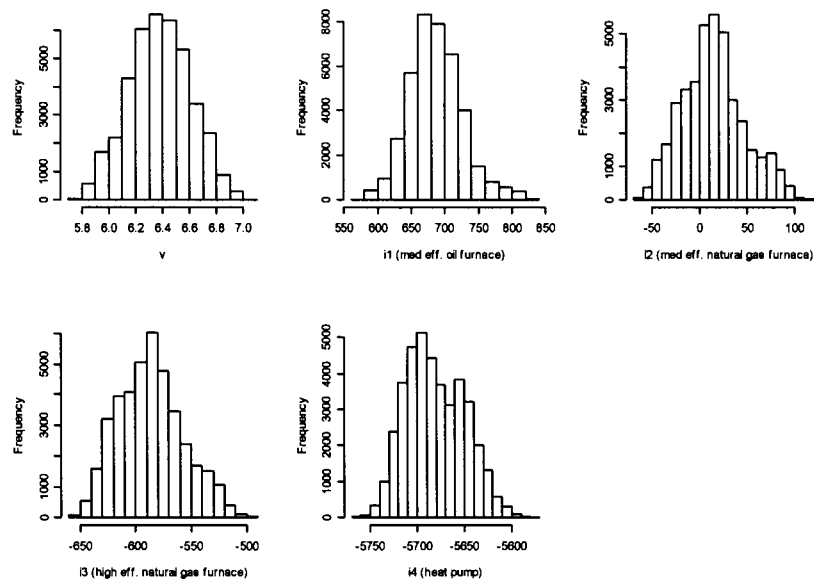


Figure 9: Unscaled marginal posterior probability distributions output from MCMC calibration routine with 40000 iterations; Ratio of accepted points to total number of candidate points = 0.228; standard deviation of jump distribution = (0.01, 0, 5, 0, 5, 5, 5, 0.0).

These marginal distributions each show the probability of one parameter integrated over the other calibrated parameters. The histograms indicate the unscaled probability density (frequency of accepted points in calibration) for given values of each behavioural parameter. The distributions were generated through the MCMC calibration algorithm using mildly informative normal prior probability distributions (priors) centred over the values currently used in CIMS: heterogeneity $\nu = 10$ and all intangibles $i_k = 0$.

As illustrated by these marginal posterior distributions, the calibration algorithm estimates reasonable values for the furnace behavioural parameters. The marginal posterior for the heterogeneity factor, ν , is centred around a value of about 6.4, suggesting a substantial degree of heterogeneity in residential furnace preferences. This value suggests market shares will be split more evenly between technologies than has been modelled in the past with CIMS (which uses a value of $\nu = 10$), despite differences in life cycle costs. The value was estimated, however, over a historical period in which the price of energy did not vary dramatically. Heterogeneity may be much larger for a forecast scenario in which a policy such as a carbon tax could result in much larger energy prices than existed during the calibration period.

The estimated intangible costs are also generally reasonable. The intangible cost of the oil furnace is the largest and suggests that owning an oil powered furnace is less desirable. This value may reflect the fact that oil heaters require a storage tank and that the tank is refilled by truck delivery; consumers may perceive the inconvenience of this system relative to a natural gas pipeline or electricity distribution lines as additional, non-financial costs. An intangible cost on oil furnaces is consistent with parameters currently used in CIMS. The negative intangible cost on the high efficiency natural gas furnace may reflect perceived non-financial benefits such as insulation from risk of future increases in the price of natural gas. The non-zero natural gas intangibles also might, however, reflect actual capital costs being different than the values used in the model.

The largest difference between calibrated and un-calibrated parameters is the intangible cost on the heat pump archetype. The large negative intangible cost estimated in calibration (with -\$5700 being the most probable value) would suggest consumers find

substantial, apparently non-financial, value in heat pumps relative to the other available technologies. This calibrated value seems counter-intuitive given how consumers can have difficulty understanding how heat pumps work, and risks associated with longer pay-offs (especially for ground source heat pumps, in which heat exchange loops are installed underground, which might be perceived as unfamiliar, high risk technology). Rather than representing non-financial benefits, however, this negative intangible cost may instead represent an incorrect capital cost for the heat pump archetype. Ground and air-source heat pumps were aggregated into a single archetype even though the actual capital costs of these two technologies are quite different. The estimated negative intangible might then suggest that more (inexpensive) air source heat pumps were sold than (expensive) ground source heat pumps, and that the aggregate capital cost of the archetype should thus be lower. While this explanation may be plausible, the large negative intangible for heat pumps has significant repercussions when extrapolated to a forecast. I return to this issue in Chapter 4.

The shape of the distribution also seems to indicate successful calibration. Each marginal distribution is relatively normal, with no major bi-modality or even substantial skewness.

3.3.2 Justification for convergence

Further analysis suggests that MCMC calibration included enough iterations to approximate the posterior probability distribution. Trace plots in Figure 10 show the trajectory of the Metropolis-Hastings algorithm's random walk through parameter space. The oscillation of the random walk results from the stochasticity in the calibration algorithm; to ensure the walk doesn't get "stuck" at a local maxima in the distribution, the algorithm sometimes accepts a point even if it doesn't improve the fit between model and data. The apparent oscillation around consistent values in Figure 10 shows good mixing through parameter space, indicating that algorithm has converged around the

mode of the distribution. The plots do not show any autocorrelation that might skew the posteriors, suggesting that a sufficient number of “burn-in” iterations were removed.¹⁸

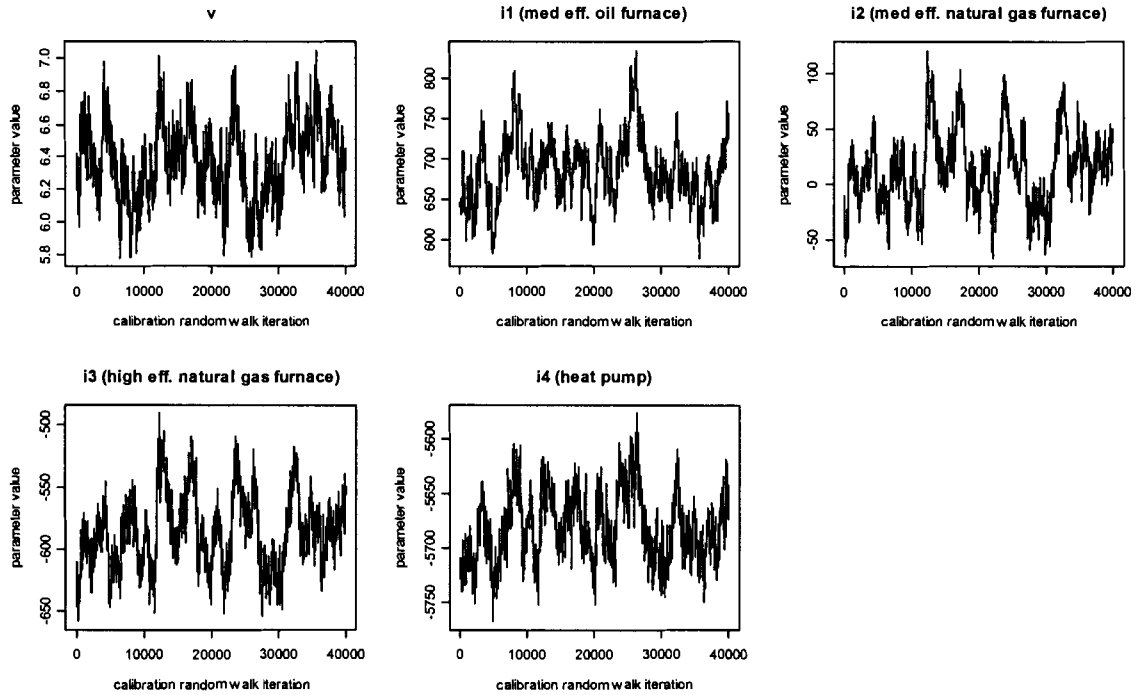


Figure 10: Trace Plots for MCMC run with 40,000 iterations (burn-in iterations removed); Ratio of accepted points to total number of candidate points = 0.228

Figure 11 provides even stronger evidence that the MCMC sample points approximate the posterior. The figure compares historical stock trends for each technology with backcasted trends, calculated using CIMS with the behavioural parameters set to the modes of the distributions output by calibration. As the figure illustrates, for each technology, the historical data correspond very well to the backcasted trends. Note that the perfect matches for the low efficiency oil and low efficiency gas furnace archetypes have been artificially imposed on the backcast so as to not distort the model’s calculation of demand. As discussed in Chapter 3, these are older base stock technologies that

¹⁸ Removing “burn-in” iterations refers to the practice of not including the initial set of MCMC iterations in parameter estimations in order to remove effects of auto-correlation (time correlation between subsequent points in the random walk) as it “climbs” toward the mode of the posterior distribution. For the furnace calibration, 10,000 “burn-in” iterations were removed.

decline with time. They are independent of behavioural parameters since no new stock is purchased; the new market share for these technologies is always zero.

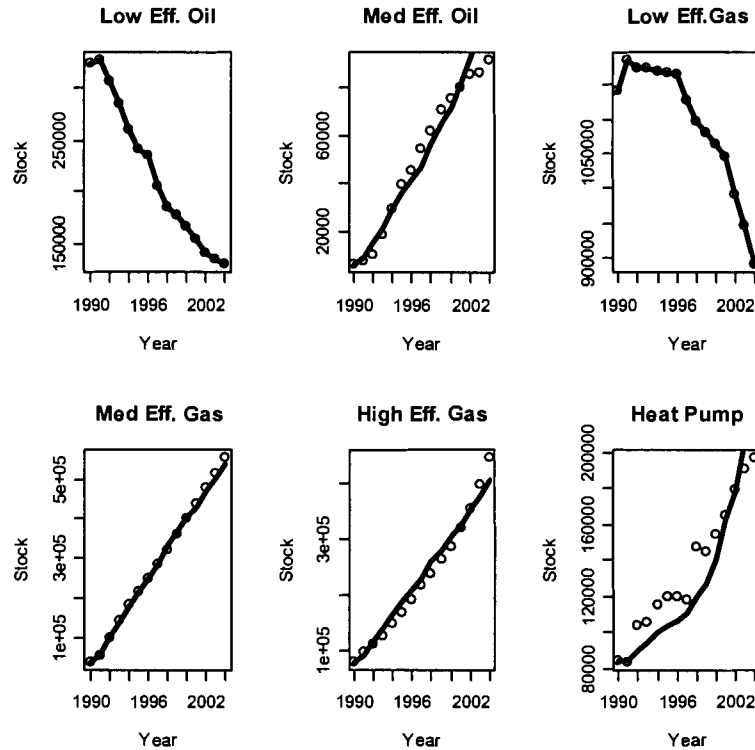


Figure 11: Comparison of forecasts with modes of MCMC-output marginal posterior probability distributions ($\theta = 6.35, 670, 15, -585, -5695, 0.3$) (lines) vs historical trends (points)

3.3.3 Importance of prior probability distributions

An advantage of a Bayesian approach to calibration is that previous research and expert opinion can supplement historical data in describing values for parameters through prior probability distributions.¹⁹ In the case of the furnace calibration, priors became critical. Without incorporating informative priors into the calibration, the MCMC algorithm did not converge. Using priors centred around the deterministic values previously used in the model (see Appendix A) causes the calibration to converge to the distribution in Figure 11.

¹⁹ Prior probability distributions, or “priors,” represent previous knowledge of parameter values. Because large standard deviations were applied to the prior probability distributions, the priors do not dominate the posterior. Historical data have the strongest effect on calibration.

Non-convergence with non-informative priors seems to result partly from the fact that the posterior function compares historical and forecast market shares for multiple technologies. A given combination might, therefore, result in a good fit with one technology, but a poorer fit with another. Since no combination results in an *exact* fit, the random walk in the calibration algorithm tends to oscillate somewhat between different parameter combinations. I confirmed that this effect had some impact on calibration by calculating the sum of squares error between forecast and historic data for each technology archetype separately. As the random walk moved to new points in parameter space the sum of squares decreased for some technologies but increased for others.

The lack of convergence might also result because relative market shares are determined by relative costs of the technology archetypes. As discussed in Chapter 2, if all intangible costs for all archetypes are increased by the same factor, relative lifecycle costs might remain approximately constant depending on capital and operating costs. Given this issue in parameterization, multiple combinations of intangible parameters could result in a good fit with data.

Informative priors, however, focus the calibration on a specific region in parameter space, and result in more distinct parameter estimation, as illustrated in the normally shaped, uni-modal joint distribution in Figure 11. By setting the means of the priors to zero, the calibration finds the mode in the joint posterior with the lowest values of intangible costs, rather than oscillating between alternative modes. A log-normal prior distribution was used for ν , and normal distributions were used for all i -parameters. Standard deviations of these distributions were set such that the magnitude of the prior distribution was small relative to the posterior; the priors are therefore only mildly informative and do not dominate the data-driven likelihood function.

But how much do the prior distributions affect the estimated posterior distributions? While carrying out a full sensitivity analysis on the value of the mean of the prior distribution is impractical given that a single calibration is computationally intensive, testing an alternative parameter “scenario” can be informative. Using an alternative prior probability distribution (with a mean for ν of 12 rather than 10), I generated a second

posterior probability distribution, centred on a different mode ($v=7.7$, $i_1 = 1720$, $i_2 = 1070$, $i_3 = 475$, $i_4 = -4600$, $r = 0.3$). This alternative combination of behavioural parameters *also* resulted in a good match with historical data. The alternative estimated intangible costs are each shifted upwards by a similar magnitude (about \$1000) from the first calibrated result to adjust for a higher value of v . The *relative* magnitude of lifecycle costs of the archetypes are similar to the first calibration.

The posterior thus does appear to be somewhat sensitive to the prior. Still, while the existence of multiple “good fit” parameter combinations makes the use of informative priors important for calibration, applying informative priors is not unreasonable. Previous parameter values have been estimated from a combination of expert opinion and previous revealed and stated preference surveys, as described in Section 1.2.4. This past research is relevant, and should not be ignored in the calibration.

3.3.4 Correlation between parameters

The posterior sample generated from calibration of the residential furnace node indicates strong correlation between the parameters, as illustrated in Table 2.

Table 2: Correlation between calibrated parameters in estimated joint posterior probability distribution

	v	i (med. eff. oil)	i (med. eff. NG)	i (high. eff. NG)	i (heat pump)
v	1.00	-	-	-	-
i (med. eff. oil)	-0.58	1.00	-	-	-
i (med. eff. NG)	0.68	0.17	1.00	-	-
i (high. eff. NG)	0.44	0.40	0.95	1.00	-
i (heat pump)	-0.15	0.76	0.58	0.72	1.00

The strong relationship between almost all parameter values is unsurprising. As I discussed in Section 2.2, the behavioural parameters overlap in how they represent consumers’ preferences. Heterogeneity, for example, can represent variation in how

different consumers perceive intangible costs and risk. Further, since the market share function calculates market shares as a function of the relative lifecycle costs of technology archetypes, multiple combinations of intangible costs could result in similarly scaled lifecycle costs and a single market share output. The strong correlation between intangible costs, therefore, is logical given that higher intangible costs for one archetype are likely associated with higher intangibles for the others.

Figure 12, for example, shows a snapshot of two dimensions of Table 2, showing the correlation between the intangible costs of the medium and high efficiency natural gas furnace, i_2 and i_3 . The figure shows all points in the calibration's random walk in the parameter space dimensions of these two parameters, illustrating that larger values of i_2 have a high probability of being associated with a larger value of i_3 . Because the parameters are strongly correlated (by a factor of 0.95, as shown in Table 2), changes in one parameter dimension are offset by changes in another.

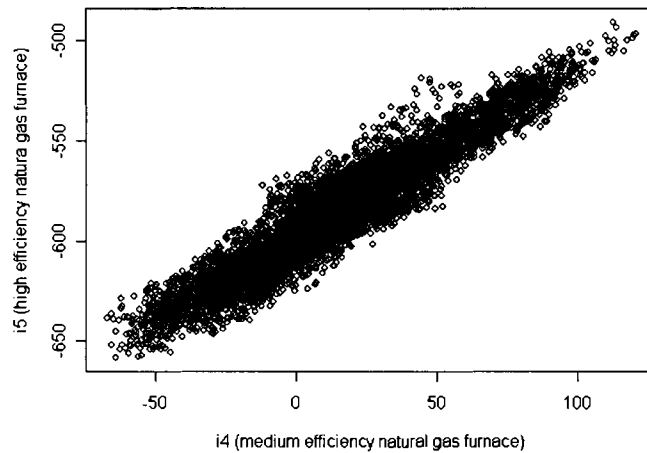


Figure 12: Correlation between medium efficiency natural gas furnace intangible cost, and high efficiency natural gas furnace intangible cost in posterior sample points.

3.3.5 Modeling historical policies

A final point of interest in the calibration of the residential furnace node concerns policies that might have affected historical furnace preferences. First, two policies likely played a role in consumer preferences over the historical period. In 1995, the Canadian federal government (Canada, 2007), required all new gas-fired furnaces to have a minimum “annual fuel utilization efficiency” (AFUE) of 78% and in 1998 required all oil-fired furnaces to have a minimum AFUE of 78%. Similarly, the Ontario provincial government had already regulated under the Energy Efficiency Act, (Ontario Ministry of Energy, 2007), that all gas furnaces manufactured after January 1, 1992 and all oil furnaces manufactured after September 1, 1994 must have minimum AFUE ratings of 78%. The Ontario regulations therefore should limit the availability of the low efficiency gas and oil-fired furnaces.

Normally, these regulations could be modelled in CIMS by not competing the low efficiency archetypes in the years in which they were not available, and retiring existing stock according to CIMS’ retirement functions. The federal and provincial regulations would only affect the retirement of base stock, not the choices of new technologies. However, as discussed in Section 3.3.2, for the analysis of the furnace node, the retirement functions were removed, and the backcasted stocks for the low-efficiency archetypes was instead set to the actual, historical stocks. Thus while the regulations might explain why the linear declining base stock function usually used in CIMS did not accurately explain historical trends, excluding an explicit representation of these policies in the backcast did not affect calibration.

3.4 Vehicle node

In the full version of CIMS, vehicle options are competed against each other to meet the demand for “vehicle-kilometres travelled”.²⁰ CIMS policy forecasts are often concerned with consumers’ choice between hybrid electric, conventional gasoline, and other emerging vehicle technology options. A backcast from 1990 to 2003, however, is

²⁰ “vehicles-kilometres travelled” is an energy service in the urban transportation node in CIMS, reflecting the demand for mobility through personal vehicles.

different in that over this historical period, hybrid vehicles had very small market shares and were not necessarily available. For the purposes of calibration in the historical backcast I model a competition between gasoline vehicles in four technology archetypes: low efficiency cars, high efficiency cars, low efficiency light trucks, and high efficiency light trucks.

Furthermore, I structure the model in two levels, as a “node-compete” competition. As illustrated in Figure 13, this structure means that the average of the two car archetypes competes against the average of the truck archetypes to determine the market split between cars and trucks, and in a nested technology competition, the low and high efficiency cars compete to determine a split within the car market share. Similarly, low and high efficiency trucks compete for the truck share of the total vehicle market. The “node-compete” nature of the vehicle node complicates model parameterization. Now three heterogeneity parameters, v , are required: one for the node competition between cars and trucks (v_{NC}), and one for each technology competition (v_{car} and v_{truck}).

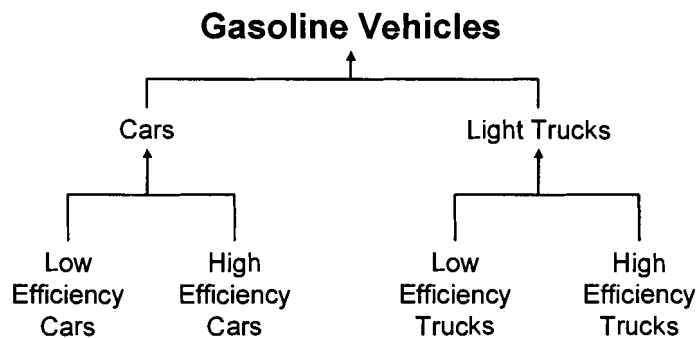


Figure 13: Structure of gasoline vehicle node for backcasting

Unlike the previous two nodes, which were calibrated to regional Canadian data, the vehicle node was calibrated with American new market share (sales) data (derived from

Automotive News (2000) data and a U.S. Department of Transportation (2004) report).²¹ American data were used because the equivalent Canadian market share data were very difficult to obtain due to confidentiality issues. An American model, however, is a good substitute. As the Canadian market is similar to the American, parameter values calibrated with the American model may be extrapolated in part to the Canadian. One caveat for this extrapolation may be that since Americans have somewhat higher incomes, their preferences on average may be more disposed towards larger, more expensive vehicles.

3.4.1 Dynamic historical vehicle preferences

Similar to the refrigerator node calibration, the historical data displays dynamic market shares through time; generally, light trucks gain new market shares at the expense of cars. Essentially, the data illustrate the growth in popularity of Sport Utility Vehicles (SUVs). However, again paralleling the refrigerator analysis, the CIMS model is incapable of replicating this trend. Fuel costs vary too little over the time-period, and consist of too small a portion of the levelized life-cycle costs for the modelled market share forecasts to vary substantially over time. Capital costs (accounting for inflation) of similar vehicle types did not vary substantially over the historical period (Consumer Reports, 1990; 2000). Again, as illustrated in Figure 14, the calibration can thus match only the average market shares over the historical period. The forecasted market shares are essentially horizontal lines, while the historical trends change through time: low efficiency cars, and high and low efficiency trucks show clear trends, while the time series for high efficiency car market shares is quite noisy and shows no clear trend from 1990 – 2003.

²¹ Historical market shares were calculated by aggregating the Automotive News (2000) nameplate data into market share splits for efficiency categories for vehicles for 2000. These splits were extrapolated over the more generalized historical time trends from 1990 to 2003 from the U.S. Department of Transportation data derived from the CAFE (corporate average fuel efficiency) program.

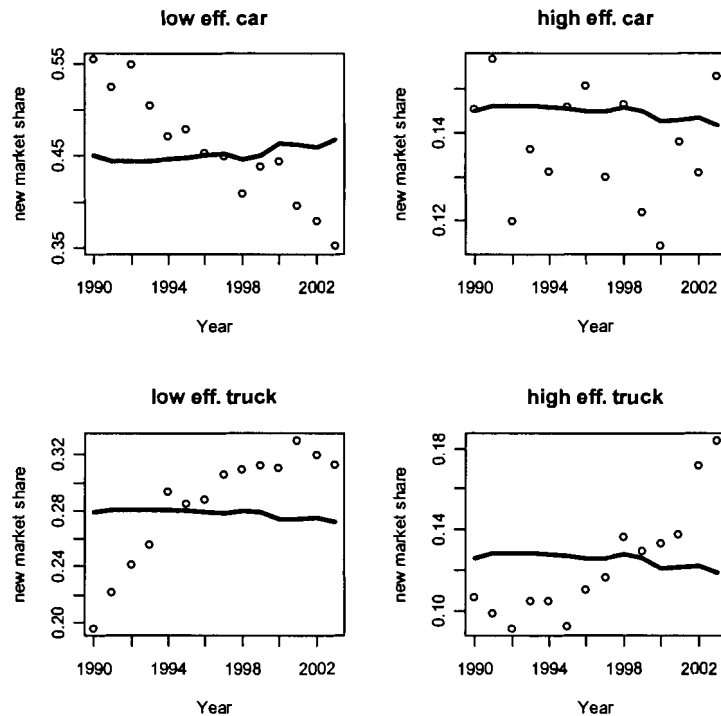


Figure 14: Comparison of historical data (points) with forecasted trends using calibrated, best-fit parameters (lines)

A combination of several effects is likely responsible for the change in consumer preferences towards light trucks and SUVs. Possible factors include:

- **Perverse effect of the CAFE (Corporate Average Fuel Efficiency) standards:** CAFE standards were introduced in the United States in 1975 and relaxed slightly in 1986. The standards required automotive manufacturers to maintain a minimum average fuel efficiency for all vehicles manufactured. Higher standards, however, were imposed on the “passenger automotive” category than on “light trucks.” Godek (1997) argues that this disparity allowed SUVs (categorized as light trucks under CAFE) to become a more economical substitution for large passenger cars.
- **Perceived safety “prisoner’s dilemma” effect:** Because SUVs are larger and heavier than cars, if an SUV collides with a car, the occupant of the SUV is more likely to avoid serious injury. As more SUVs appear on the market (and the road),

consumers have increasing incentives to *also* purchase an SUV so as to not be on the “wrong side” of a collision (Vanderheiden, 2006).

- **Larger margins for manufacturers:** Pelkmans *et al.* (2003) suggest that initially, manufacturers saw a larger price margin on the SUV than the car. This large margin caused more SUV models to be produced and promoted. This margin would have declined as more manufacturers tried to compete in the SUV market as it moved toward equilibrium.
- **Price signal lag effect:** Low energy prices in the late 1980s may have given the consumer reason to be more confident that gasoline prices would remain low, and thus could have provided incentive to purchase less fuel efficient vehicles such as SUVs.
- **Rising income effect:** Increasing Gross Domestic Product (GDP) over this period may have provided vehicle purchasers with more disposable income and more willingness to purchase more expensive vehicles that provided additional benefits (larger cargo space, better performance on backcountry roads, status) (Frank, 1999; Vlek and Steg, 2007).

3.4.2 Modeling dynamic preferences for vehicles

How then can these possible effects be accounted for in a backcast so as to match model and empirical frameworks? Perverse effects of CAFE are difficult to model, as they mostly affect supply (i.e., the range of vehicle models that manufacturers offer). Price signal lag effects are also difficult to model, given that life cycle costs are calculated with only the current price of fuel, though Peters’ (2006) work in modelling consumers expectations for the future could be relevant. Due to the difficulty associated with modelling these effects directly, however, including them in the model is outside the scope of this work.

All other effects discussed in Section 3.4.1 effectively describe dynamics in intangible costs (representing changes in perceptions of safety, provision of additional services, or status associated with owning a vehicle). The *structure* of the changing intangible costs, however, remains uncertain. The collision safety effect suggests intangible costs should be a function of the market share of light trucks already sold. Similarly, a neighbour effect (Mau, 2005) might suggest an asymptotic decline in intangible costs. An income effect suggests intangible costs as a function of GDP. Many other relationships might also be reasonable.

More complex relationships do not, however, necessarily constitute a better model. In this case, choosing a function for a dynamic intangible introduces additional model structure uncertainty since the theoretical causal mechanism (i.e., the driver of dynamic preferences) is unclear. Calibrating parameters under larger structural uncertainty makes estimating parametric uncertainty more challenging; determining which type of uncertainty is driving the differences between the historical data and the backcast is not possible. Indeed, the inability of the model to replicate changing historical preferences suggests that structural uncertainty does play a significant role. Forecasting with calibration results in this case could thus introduce additional (and unquantified) uncertainty into forecasts and CIMS policy assessments.

Nevertheless, the move toward SUVs and light trucks was clearly a significant trend in consumer preferences and cannot be ignored. For the purposes of calibration, I apply a linear dynamic intangible cost on the truck archetypes to represent the increasing appeal of light trucks over cars. Assuming that the dynamics are mostly relevant in the node-competition between trucks and cars, the slope of the linear intangible can be the same for both truck efficiency archetypes. No dynamic intangible need be applied to the car archetypes at all, as the dynamic truck intangible will drive the changes in differences in life cycle costs between cars and trucks. I therefore attempt to model the intangible costs for trucks as per Equation 5

$$i_{k,t} = i_{k,0} + m * t$$

Equation 5

Where:

- $i_{k,t}$ = the capital costs for technology option k (low efficiency or high efficiency truck) at time period, t
- $i_{k,0}$ = the initial intangible costs of technology k (low efficiency or high efficiency truck), or the y- intercept of the dynamic truck intangible
- m = a trend parameter indicating the change in the dynamic truck intangible between subsequent time periods (the same value for low efficiency or high efficiency truck archetypes)

The linear trend can serve as a approximate representation of the historical temporal dynamics. This somewhat arbitrary structure is useful for further exploration of historic preference dynamics. Only one additional parameter (the rate of intangible change, m) must be calibrated, making calibration tractable. The value of this calibrated parameter will also provide an indicator of the magnitude of the historical preference changes in general.

Nevertheless, the linearly declining intangible is clearly not a true representation of the mechanism of historical preference dynamics. Using this relationship in a forecast into the future would be problematic. Since the causal mechanism is undefined, the linear decline consistent with past trends will not necessarily be relevant for the future. Indeed, the linear trend cannot continue indefinitely into the future, or the intangible cost would eventually decline toward negative infinity. Further, in the context of a CIMS forecast, even if policies were implemented to encourage a shift toward lower carbon-emitting vehicles, the model would still be “hard wired” for a continued shift toward light-duty trucks.

3.4.3 Vehicle parameters with dynamic intangible costs

Applying the dynamic truck intangible to the model does allow the calibration algorithm to replicate the dynamics of the historical trends. Parameter distributions generated by the calibration algorithm are illustrated in Figure 15.

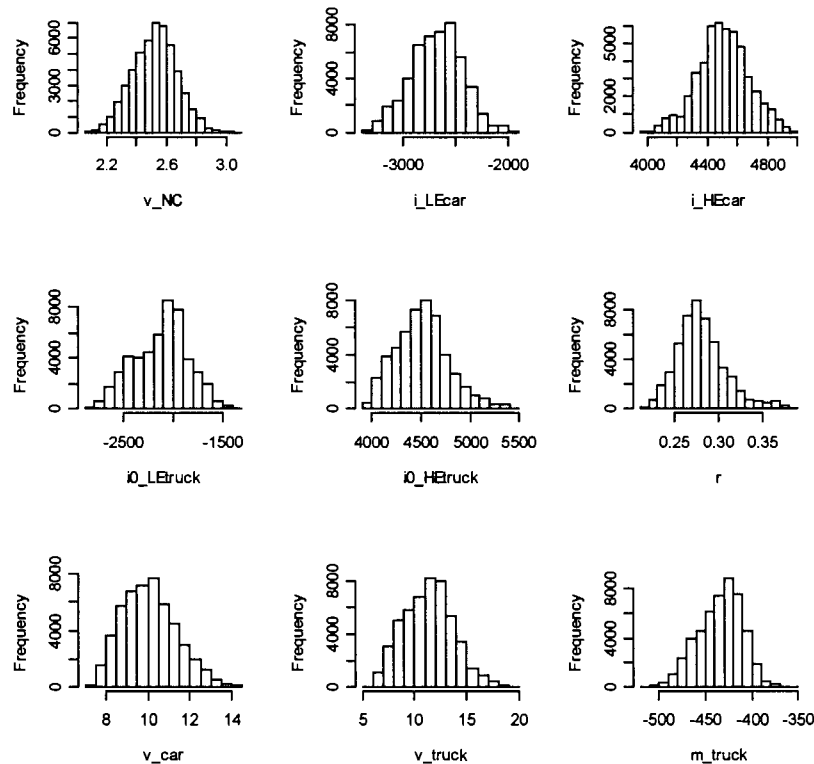


Figure 15: Unscaled marginal posterior probability distributions for behavioural parameter values. Ratio of accepted points in calibration = 0.278; number of iterations = 50,000 (additional 50,000 removed as burn-in)

Several indicators suggest that the calibration included a sufficient number of iterations to approximate the posterior distribution. First, the distribution is generally normal in shape, though some skewness is apparent. Second, as Figure 16 shows, the modes of the parameter distributions estimated do result in an approximate fit with historical data. Third, paralleling the furnace analysis, I again generated trace plots that showed good mixing. Finally, the ratio of accepted points (0.278) is small enough to indicate that each step in the random walk is big enough such that the random walk did not become

“trapped” on a local minimum in the distribution, but small enough that the true shape of the distribution could be estimated.

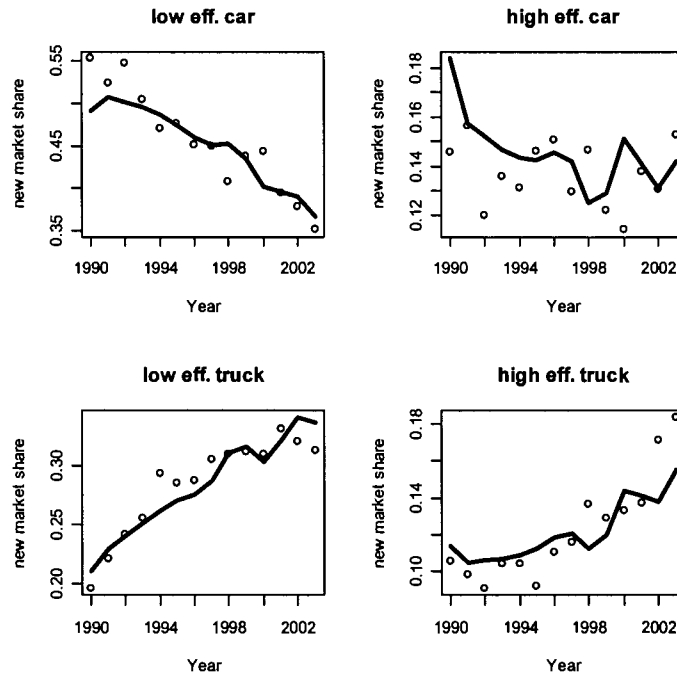


Figure 16: Comparison of historical and forecasted (using behavioural parameters equal to the mode of the estimated posterior distribution) market shares from 1990 - 2003 for gasoline vehicles

The estimated posterior distributions in Figure 15 suggest reasonable values for the behavioural parameters. The heterogeneity parameter distribution for the choice between cars and trucks, v_{NC} , has a mode of 2.4, with an approximately normally shaped marginal distribution. This value suggests substantial heterogeneity in the market, and that the split between cars and trucks in general is less sensitive to the difference in price between the two archetypes than currently represented in CIMS. This value is consistent with the results from previous empirical research, as illustrated in Table 3.

Similarly, the distributions for the heterogeneity factors describing the technological competitions between high and low efficiency cars and trucks appear reasonable. Both are normally shaped, and have means close to $v = 10$, the value used currently in the

model. While this value is quite different from the values of ν estimated in other studies (Horne, 2003; Eyzaguirre, 2004; Mau, 2005; Axsen, 2006), these studies estimated only a single ν parameter rather than the three parameters estimated here for the node-compete structure. This study seems to suggest less heterogeneity in the choice of high and low efficient vehicles than previous empirical research. As this historical shift toward SUVs illustrates, however, the node-compete split seems to play a very important role in assessing preferences. The fact that these past studies are very much consistent with the calibrated ν_{NC} suggests that past studies may have been driven more by the car – truck choice than the high efficiency – low efficiency choice. Further, these past studies included new, emerging low emission vehicles such as hybrid cars while the calibration backcast included only gasoline cars and trucks. It makes sense that consumers would be less price sensitive when faced with the option of a new, emerging technology and more inclined to base their decision on other factors such as risk. CIMS policy analyses forecast the effects of future policies on the penetration of just these kinds of new technologies. As such, the calibrated heterogeneity parameter may be too low for policy forecasting.

The values of the distributions for intangible costs are generally consistent with Horne's (2003) study, as again illustrated in Table 3, at least in terms of the approximate relative differences between archetypes. High efficiency models are small, light, and usually have poorer performance; these characteristics are represented by the high estimated intangible costs. Similarly, even though lower efficiency vehicles are more expensive and less efficient, they offer non-financial benefits such as storage space, power, and versatility of use that are reflected in the estimated negative intangible costs. (An SUV, for example, might have off-road capability where a small compact car would not). While the relative values are similar, the absolute values are not. These differences are likely the result of the parameterization issue described in Section 2.2, given that since the model calculates market share as a function of *relative* life cycle costs, multiple, scaled combinations of intangibles could result in the same market share. Further, since a different set of available vehicle technology options are modelled in Horne's (2003) study, the

calibration algorithm may estimate intangible costs that have been scaled differently. A direct comparison may not be valid.

Unlike in the furnace calibration, I also estimate the revealed private discount rate, r , for vehicle choice. The calibration routine estimated a distribution with a mode of 26.5%, which falls directly amidst the range of values estimated by previous studies (Horne, 2003; Eyzaguirre, 2004; Mau, 2005; Axsen, 2006), as illustrated in Table 3. Further, the estimated distribution ranged from about 20% to 35%, which again is consistent with the other estimations. It is also consistent with literature: Train's (1985) survey suggests revealed preference discount rates lie within a range between 2 and 41%. The fact that the estimation of consumers' time preferences matches so well with previous studies and with literature provides further evidence that the calibration algorithm successfully estimated parameter value distributions.

Table 3: Results from other CIMS parameter estimation studies for values of behavioural parameters

Study	Mean behavioural parameter values estimated						r (%)
	v (node choice)	v (vehicle choice)	i (low efficiency car)	i (high efficiency car)	i (low efficiency truck)	i (high efficiency truck)	
Horne, 2003	-	2.9	-\$2,693	\$16,058	-\$2,693	\$16,058	22.6
Eyzaguirre, 2004	-	5.2	-	-	-	-	27.6
Mau, 2005	-	2.4	-	-	-	-	21.8
Axsen, 2006	-	5.3	-\$3,420	\$6,555	\$301	-\$10,325	21.6
MCMC calibration ²²	2.525	cars: 9.8 trucks: 11.5	-\$2,850	\$4,450	-\$5,430	\$970	26.5

3.4.4 Alternative model structures

Even though the parameter values suggested by the posterior are generally reasonable, the imperfect model fit in Figure 16 suggests that the model still does not fully represent

²² For the purpose of comparing results of the calibration with other parameter estimations, the linearly declining truck intangible parameter, $m = -\$435$ / year, is not reported in Table 3, since the other studies did not use such a parameter. Instead, fixed intangibles are reported as the *average* of the dynamic intangibles over the historical period.

preferences over the historical period. Some noise may exist in the historical data that does not reflect a significant process in representing consumer preferences (note in particular the oscillating market share for high efficiency cars). However, more generally speaking, the best-fit forecasted trends also do not perfectly match the trends in the historical data. In the high efficiency truck panel in Figure 16, for example, the model trends fits the first half of the historical trend much better than the second. Similarly, the low efficiency car forecast fits the historical data very well in the latter years of the historical run, but seems much flatter than the steep historical trend in the first five years.

The inability of the model forecast to match the historical trend over the entire historical period could indicate that the data series provides conflicting information at different times in the series and therefore that the model is not capturing some significant effect evidenced by the data. This structural inadequacy of the model is hardly surprising: after all, the linear dynamic intangible was imposed on the truck archetype arbitrarily to represent the historical change in preference towards SUVs. As I discuss in Section 3.4.2, many factors likely contributed to this change, and the mechanisms for these factors are not necessarily easy to model. A simple linear trend is probably not an ideal description of historical preference dynamics. Ideally, these dynamics could be further explored by repeating the calibration for alternative parameterizations (representing the dynamic intangibles, for example, with an asymptotic exponential effect, as is suggested by the neighbour effect research (Mau, 2002), or even with non-stationary random walk parameters). The usefulness of these alternative model structures could then be compared and assessed using a Bayesian Information Criterion (BIC).²³ However, given that the dynamics in historical preferences may be totally different from those in the future, this analysis may not be helpful for forecasting. The usefulness of replicating the historical dynamics is particularly questionable given that historical preference dynamics took place in an environment without strong policies or price signals. The historical dynamics are therefore even less likely to accurately forecast consumers' responses to policies such as a carbon tax. For these reasons, a more detailed assessment of alternative intangible

²³ The BIC allows different models (i.e. parameterizations) to be compared. It takes into account how well the forecast matches the data (the residuals) as well as complexity of the parameterization (the number of free parameters estimated).

structures is outside the scope of this study. I nevertheless return to this issue in Chapter 5.

3.5 Summary

Table 4 summarizes the modes of the estimated parameter posterior distributions from each of the nodes calibrated. The parameter values estimated for these nodes are generally consistent both with past parameter estimation research using CIMS and with results from literature.

Table 4: Summary of estimated parameters through calibration

Node	Mode behavioural parameter values estimated			
	v	i	r	<i>other</i>
Refrigerators	not successfully estimated	not successfully estimated	NA	NA
	$v = 6.4$	$i(\text{med. eff. oil}) = \670	NA	NA
Residential Furnaces		$i(\text{med. eff. natural gas}) = \15 $i(\text{high eff. natural gas}) = -\585 $i(\text{heat pump}) = -\$5695$		
	$v(\text{node compete}) = 2.4$	$i(\text{low. eff. car}) = -\2650	$r = 0.265$	$m(\text{truck}) = -\$435/\text{year}$
Gasoline Vehicles	$v(\text{truck}) = 9.8$ $v(\text{car}) = 11.5$	$i(\text{high eff. car}) = \4450 $i(\text{high eff. truck}) = -\1850 $i(\text{low eff. truck}) = \4450		

Using these parameters in CIMS policy forecasts will incorporate revealed consumer preferences, which may pose problems for modelling policies such as carbon tax. In this case, the model would forecast consumers' response to apparent energy prices (as modified by the tax for carbon-intensive fuels) well outside the range represented in the historical calibration period. Calibrated parameters may not be relevant under these different economic conditions. Nevertheless, since the results of calibration can be used in support of other empirical parameter estimations generated through stated and combined stated and revealed preference approaches, they should improve the model's

representation of the consumer behaviour. Using the distributions to explicitly incorporate uncertainty into a forecast can further improve its credibility.

Even failure of calibration (as with the refrigerator node, and to a lesser extent the vehicle node) is useful in qualitatively assessing structural uncertainty in CIMS. The inability of the refrigerator mode to replicate the dynamics in historical preferences may suggest that manufacturer decisions drove an apparent autonomous increase in refrigerator efficiency over time. Historical trends implicitly suggest that demand-side management programs such as the “golden carrot” might have been more significant than anticipated in accelerating technological change. These hypotheses cannot, however, be confirmed without additional analysis of rates of technological change in previous time-periods. The refrigerator analysis also indicates that changing refrigerator-related services might have a substantial effect on consumer preferences. Consumers may tend to choose a refrigerator based refrigeration volume and the availability of features such as icemakers, rather than on the amount of energy it consumes.

Similarly, the failure of the static vehicle model to match dynamics in historical preferences also led to insight regarding possible drivers of the shift away from cars toward light trucks during the 1990s. While a linear declining intangible cost function model *was* capable of providing an approximate fit with historical trends, both the arbitrariness of the structure and the lack of a perfect fit suggest additional effects may also be relevant. These effects are not, however, necessarily relevant to forecasts of future technological change under very different economic conditions as modified by policies such as a carbon tax.

CHAPTER 4: POLICY ANALYSIS UNDER UNCERTAINTY

4.1 Introduction

Fundamentally, CIMS is a tool designed to help address an important policy problem: how can Canadian greenhouse gas emissions be reduced as part of a global effort to reduce future risks of climate change?²⁴ CIMS can help policy-makers answer this question by forecasting the effects of alternative energy policies on technological change. By predicting the potential for low carbon-emitting technologies, CIMS can help indicate which policy instruments will most effectively reduce greenhouse gas emissions. In this chapter, I apply the behavioural parameters estimated in this study to a policy analysis using CIMS forecasts and evaluate how the results of calibration affect this analysis.

4.2 Energy policy instruments

Alternative policy instruments are available to policy-makers. These options can be assessed using multiple criteria: *efficacy* measures the success of a policy in reducing emissions, *economic efficiency* is a metric of the per unit cost of reducing emissions, and *political feasibility* assesses the potential for a government to actually implement a policy, given the varied interests of an electorate.²⁵ I overview the main types of policy alternatives in this section. For more details regarding energy policy instruments designed to mitigate emissions, see Jaccard and Rivers (2007) and Jaccard (2005b).

Command and control policies, such as minimum efficiency standards for manufacturers, use government regulations to mandate specific technology characteristics. While regulations can be effective, overly prescriptive regulations can be economically inefficient, causing higher abatement costs than necessary. Regulations are only

²⁴ While CIMS can also be useful for other environmental policy problems, such as reducing emissions of “criteria air contaminants (CACs), it has been most often used in the climate change context, and I continue to focus on this primary application.

²⁵ Other policy assessment criteria might be *administrative feasibility* and *equity*.

moderately politically feasibly; some economists have suggested that regulations impose an unnecessary economic burden on firms (Jaccard, 2005b).

Financial disincentive policies, such as a carbon tax, send a price signal through the economy to consumers, encouraging the transition to lower carbon-emitting technologies. Disincentives can be both effective and economically efficient since they allow firms and consumers the flexibility to choose the lowest cost approach to reducing emissions. Taxes are often politically challenging; they can be perceived as government intrusion on the economy (Jaccard, 2005b).

Financial incentives include policies such as subsidies, low interest loans, publicly funded research and infrastructure, and tax credits. While incentives are politically feasible, studies have shown them to be less effective and economically efficient (Loughran and Kulick, 2004) relative to disincentives. Due to a free-rider problem, subsidies may reward firms or consumers who would have adopted a low-efficiency technology even in the absence of the subsidy.

Voluntarism and information policies, such as advertising, labelling, demonstration projects, and information brochures, attempt to convince firms and consumers that changing their behaviour and reducing energy use is ethical or in their best interest. The effectiveness of these programs, however, is questionable given their historical failure to reduce emissions, and their economic efficiency is usually low (OECD, 2003)

Finally, *market-based regulatory policies*, such as a “cap and trade” system or the California vehicle emissions standards, are economically efficient because they allow flexibility for abatement. For example, under a cap and trade scheme, firms for whom abating emissions is expensive can purchase emission credits from those for whom abatement is less expensive.

4.3 Stochastic policy analysis

CIMS can quantitatively forecast the impacts of these different policy instruments. By comparing forecasts of greenhouse gas emissions under a combination of policy instruments to a *business as usual* (BAU) scenario representing the status quo, the model can assess the forecasted effectiveness of the policy package in reducing emissions. Including uncertainty in a policy analysis can improve analysis. Two of Morgan and Henrion's (1990, p. 39) "ten commandments of good policy analysis" are to "be explicit about uncertainties," and to "perform systematic uncertainty analysis." By applying the probability distributions estimated through calibration to a policy *forecast*, uncertainty in the behavioural parameters can be explicitly accounted for and propagated through CIMS forecasts. Incorporating uncertainty into the CIMS model can further improve its usefulness to policy makers because risks can be quantified and a range of possible outcomes can be assessed (Morgan and Henrion, 1990).

I apply a Monte Carlo simulation technique to explicitly include uncertainty in a forecast of future greenhouse gas emissions from residential furnaces in Ontario. This approach involves: 1) randomly sampling behavioural parameter values from their joint posterior probability distributions; 2) running the model for each sample; and 3) calculating an average, or "expected value" over a number of simulation of trials. Calibration outputs are ideally suited to Monte Carlo simulation. Sampling from the joint posterior over all the parameters calibrated includes correlation between parameter values, which provides a better representation of the interactions between the behavioural parameters than using marginal distributions (as for example, in Figure 9).

Still, Monte Carlo sampling in the full CIMS model is impractical because a single CIMS run takes 10 – 15 minutes. The hundreds of runs required for a stochastic forecast make this approach infeasible. For this calibration study, however, individual CIMS technology competition nodes were modelled outside of the full CIMS model as custom-coded, individual node models.²⁶ Monte Carlo sampling is possible and practical for these

²⁶ A "node" is a technology competition for a single service (eg. heating, vehicle-km-travelled, etc. as modelled within CIMS which includes multiple nodes).

simplified models. While running policies on single nodes cannot assess system-wide GHG reductions, it does provide an opportunity to assess the impacts of incorporating uncertainty into forecasts.

4.4 Incorporating uncertainty into a CIMS policy analysis

For my stochastic policy analysis, I focus on the residential furnace node, because calibration results were most reliable for this node. Further, the node is an important one: in 2004, space heating was responsible for almost 54% of residential greenhouse gas emissions in Canada (Canada, 2006). While also running a stochastic policy for the vehicle node might also be interesting, to do so is problematic for several reasons. First, though the linearly declining intangible cost provides good fit for the historical period, no evidence suggests the linear structure will necessarily be applicable in the future. Similarly, in the historical period gasoline vehicles were the most significant technology archetypes. A future policy analysis would assess the effect of a carbon tax price regime on the penetration of alternative technologies such as hybrid, plug-in hybrid, or hydrogen vehicles which were unavailable in the historical period, and thus not informed by the calibration.

I also chose to focus on a financial disincentive in the form of a carbon tax as a policy instrument. This approach also parallels recent analysis by Bataille *et al.* (2007) for the Canadian National Roundtable on the Environment and the Economy (NRTEE). In my analysis, I simulate a carbon tax (Table 5) on the residential furnace node using the shadow prices for the NRTEE study's middle reductions trajectory in which all Canadian greenhouse gas emissions are reduced to 65% below 2005 levels. While the simulation runs the residential furnace node in isolation, the analysis implicitly includes the effects of the carbon tax on electricity generation. By separately running the electricity generation sector under the tax policy in the full CIMS model, I determine the forecasted greenhouse gas intensity of electricity in Ontario, and exogenously impose this trend on the isolated furnace simulation. Macro-economic feedback effects are not included in this run; the electricity price is not endogenously adjusted in the model to ensure equilibrium between energy supply and demand.

Table 5: Escalating carbon tax simulated in policy analysis, paralleling Bataille *et al.* (2007)

Year	Price of Carbon (\$ / tonne)
2001- 2005	0
2006-2010	0
2011-2015	18
2016-2020	88
2021-2025	176
2026-2030	284
2031-2035	317
2036-2040	317
2041-2045	317
2046-2050	317

The policy simulation compares a *business as usual* (BAU) run in which no policy is modelled, and a *policy* run in which the escalating carbon tax affects the lifecycle technology costs of each archetype in the model. These two trends are modelled under each of three modelling scenarios, with parameters as described in Table 6:

1. using un-calibrated parameter values currently used in CIMS;
2. using deterministic calibrated parameter values; and
3. explicitly incorporating uncertainty using calibrated probability distributions.

Table 6: Behavioural parameters used in forecasting for each of three modelling scenarios

Parameter	Parameter values for each modelling scenario		
	(1) uncalibrated	(2) calibrated; deterministic	(3) calibrated, stochastic (mean \pm standard deviation)
v	10	6.4	6.4 \pm 0.23
i (medium efficiency oil furnace)	\$200	\$670	\$687 \pm 40.0
i (high efficiency natural gas furnace) ²⁷	\$200	\$670	\$687 \pm 40.0
i (medium efficiency natural gas furnace)	\$0	\$15	\$14 \pm 32.7
i (high efficiency natural gas furnace)	\$0	-\$585	-\$587 \pm 29.2
i (heat pump)	\$0	-\$5,695	-\$5681 \pm 31.0

²⁷ An intangible cost for high efficiency gas furnaces could not be estimated from historical data because no such technology existed over the historical period. However, given the market growth of gas furnaces over the period, the development of a high efficiency model seems likely. For the purposes of a forecast, the intangible cost estimated for the medium efficiency oil furnace was also used for the high efficiency furnace. Given that these technologies provide a very similar service, this assumption seems reasonable. Still, the differences between the backcast and forecast model introduces additional uncertainty.

The results of this approach thus have relevance to issues of modeling in terms of the usefulness of calibration, the robustness of the deterministic model, and the potential worth of explicit uncertainty analysis. They can also inform substantive policy issues in terms of how uncertainty in consumer behaviour affects the efficacy of a carbon tax.

While CIMS can be used to assess costs of policy options (Peters, 2003), costing was not built into the single technology competition backcast models. This policy analysis is limited to assessing the efficacy of a carbon tax policy under different modelling scenarios to test and implement the results of the calibration. A detailed assessment of policy options under other evaluative criteria, such as economic efficiency or political feasibility, is outside the scope of this analysis.

4.5 Policy simulation results

Figure 17 illustrates simulation results under each of the three modelling scenarios. The leftmost panel shows forecasts under scenario 1 (un-calibrated parameters currently used in CIMS). Similarly, the middle panel shows forecasts using the mode of the estimated joint posterior probability distribution for the behavioural parameters (scenario 2). Finally, the right panel shows the results of a Monte Carlo analysis in which 300 points are sampled from the calibrated posterior (scenario 3). Ninety percent of the sampled points result in trajectories within the dotted lines on this final panel, which indicate the uncertainty in the forecasts, given uncertainty in the behavioural parameters. Comparing each of the plots in Figure 17 provides insight into the effects of incorporating calibration results of Chapter 3 into a forecast into the future.

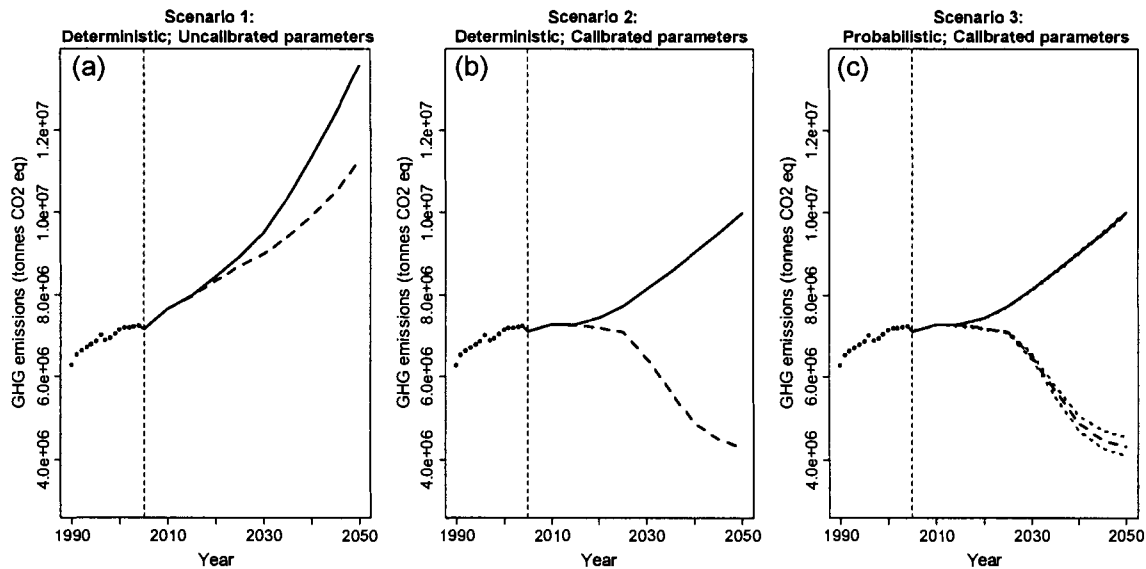


Figure 17: Historical (points; prior to 2005) and forecasted greenhouse gas trajectories for Ontario residential furnace technologies under three alternative modelling scenarios under no policy, or “business as usual” (solid) and under an escalating carbon tax (dashed). Dotted lines in panel (c) represent 90% uncertainty bounds. 300 iterations were used in the Monte Carlo stochastic run.

4.5.1 Impact of calibrated parameters

The substantial differences between the plotted forecasts in the first panel and the second and third illustrate the impact of using calibrated parameters in place of uncalibrated ones. The un-calibrated parameters, as illustrated in the first panel, result in a sharp increase of greenhouse gases, whether or not a carbon tax is imposed. This climb in emissions results from an increasing market share of oil-burning furnaces, which itself is a consequence of anticipated increases in the price of natural gas in around 2020.²⁸ The insensitivity of the model to the carbon tax price signal reflects that fact that no inexpensive electrical (low carbon) option was competed in the model as baseboard heaters were excluded.²⁹ Alternatively, a forecast using the modes of the parameter distributions actually suggests that greenhouse gas emissions from residential furnaces in

²⁸ CIMS uses an exogenous forecast for the price of energy. In this case, the forecasted price of natural gas parallels the NRTEE report (Bataille *et al.*, 2007) which adapts forecasts developed by Natural Resources Canada and the U.S. Department of Energy.

²⁹ Baseboard heaters are competed separately in the full CIMS model based on floor space and were excluded here to avoid the interactions between space heater choices and insulation technology choices. In my furnace model, baseboard heaters were accounted for by subtracting a fixed percentage of the total forecasted households in Ontario.

Ontario will level off somewhat, before oil furnace shares again begin to grow in 2020. The sharp growth illustrated in the first panel starting in 2005 seems inconsistent with the historical levelling of emissions (more detailed market share forecasts for the three scenarios are reported in Appendix B). The new parameters seem to result in a smoother transition from historical emissions to forecasted emissions under the “business as usual” conditions. This continuity between history and forecast suggests the calibrated parameters provide a better forecast than the un-calibrated parameters.

Under the calibrated parameter scenario, the carbon tax is very effective in curbing emissions, as indicated in Figure 17b. The principal difference between parameter values in the two scenarios is the large negative intangible cost associated with heat pumps in the calibrated run. This large negative value results in substantial forecasted market penetration for heat pumps. In 2050, the model suggests 27% of all new furnaces will be heat pumps, in the absence of a policy, and almost 70% under the carbon tax.

This large growth in heat pumps may be exaggerated. Part of this effect likely results because the node has been constructed to exclude baseboard heaters. The heat pump market share responds strongly to the carbon tax policy because it is the only available archetype that uses electricity (which is less greenhouse gas intensive, especially under the carbon tax, than oil or natural gas). Under a carbon tax regime, the lifecycle cost of heat pumps becomes very favourable relative to the other technology archetypes.

Extrapolation problems may also be responsible for heat pump growth. The calibration process estimates a large negative heat pump intangible cost parameter to replicate the historical growth of heat pumps (the number of heat pumps in Ontario more than doubles from 1990 to 2004). This growth continues into the forecast, resulting in the large forecasted penetration. The calibration effectively extrapolates the historical growth into the future, and the exogenous future increase in natural gas prices enhances the effect; however, such an extrapolation may not be realistic. Technology penetration often takes the form of a logistic function, eventually flattening out into an asymptote. Logistic growth might make sense for heat pumps: perhaps an environmentally-minded niche market forms a finite subset of the entire Ontario market. On the other hand, as a very

efficient technology, perhaps heat pumps will indeed develop into a dominant space heating technology; certainly at some high level of carbon tax this would be the case. The historical data only show the very beginning of heat pump penetration and can therefore give no indication of how large a maximum market share might be. Further, the calibrated parameters were estimated under historical prices; they do not necessarily reflect consumer responses to the substantially different prices that would occur under the carbon tax. These ambiguities emphasize potential limitations of calibrating a model of future behaviour from historical data. I return to this issue in Section 5.3.

A second issue arises from comparing forecasts between calibrated and uncalibrated parameters, in that the large growth in oil-fuelled furnaces may also not be realistic. Historical data show a steady growth in medium efficiency oil and high efficiency oil furnaces; by 2050 under a “business as usual” model run, oil furnaces make up 75% of new furnace sales. Since oil furnaces are more often used in rural households and rarely in urban ones, a maximum penetration might exist. In reality, the total stock of oil furnaces might actually be “capped” at a maximum penetration. Also, the isolation of the furnace node from space heating and baseboard heaters limits the effects of the carbon tax. Even if the oil furnaces are arbitrarily limited to smaller market shares, natural gas furnaces absorb the difference in market share. Without baseboard heaters, a low-cost, low carbon intensity alternative (assuming the electricity sector responds to the carbon tax) is not available for consumers. For these reasons, the growth of greenhouse gases in Figure 17a under both the business as usual and carbon tax runs may also be unrealistic.

4.5.2 Impact of explicitly incorporating uncertainty

While Figure 17b runs the forecasting model using a single set of parameters (the mode of the calibrated joint distribution over the behavioural parameters), the third panel explicitly incorporates the uncertainty in the distribution into the forecast by sampling 300 points from the distribution. Given the historical calibration data and the Bayesian prior distribution, greenhouse gas emissions have a 90% probability of falling within the dotted lines for both the *business as usual* and *policy* runs. Comparing the second and third panels reveals some interesting results.

Despite the variation in the behavioural parameter distributions (market heterogeneity, ν , for example, varies from 5.8 to 7.0 within the posterior distribution) there is very little uncertainty in the emissions forecast (Figure 17c). Under the carbon tax, the 90% probability band ranges only from 4.08 to 4.55 megatonnes of greenhouse gases (carbon dioxide equivalents), or by about 11%.

The Monte Carlo uncertainty analysis samples from the joint distribution, which is made up of the parameter combinations accepted by the calibration algorithm. Effectively, strong correlation between the parameters reduces the impact of the range of parameter uncertainty on the forecast. Though 300 different points are sampled in the stochastic policy analysis (with substantial variation in each input parameter dimension), the correlation between parameters results in small variation in forecasted outputs. The cases in which both parameters are small and in which both are large results in similar relative life cycle costs, and similar calculated market shares. A third case in which one is small and one is large results in very different calculated market shares. The strongly correlated distribution, however, indicates this third case is highly unlikely, and therefore is weighted very lightly when averaged into the stochastic run. The correlated posterior thus reduces the impact of variation in the dimensions of individual parameters.

Limited variation in forecasted emissions, might also be specific to the residential furnace node. While the variation in forecasted greenhouse gas emissions is small for both the “business as usual” and “policy” model runs, the BAU run has a particularly tight 90% probability band, which results from the larger uncertainty in the oil furnaces, particularly the medium efficiency archetype. The intangible cost parameter for medium efficiency oil furnaces is fairly uncertain. Under a “business as usual” scenario, the oil furnace archetype takes on a large market share because it has a relatively low lifecycle cost that is fairly robust to variation in the intangible cost. Even with a large intangible, the archetype is relatively inexpensive and the resulting market share relatively large. Under the policy run, however, few oil furnaces are purchased. When a carbon tax that penalizes the high emission oil furnace is included, the lifecycle cost becomes sensitive to the variation in the intangible cost modelled in the Monte Carlo uncertainty analysis.

Finally, and perhaps most significantly, the parametric uncertainty quantified by the uncertainty bounds is truly valid only under the energy price regimes in the historical calibration period. Uncertainty is likely much larger for a forecast with a carbon tax than Figure 17 suggests, since the calibration does not provide any information on how consumers respond to the much larger energy prices that would be associated with a carbon tax. Further, structural uncertainty is not quantified by this analysis or included in the uncertainty bounds. For these reasons, the bounds may underestimate the true uncertainty in emissions.

4.6 Summary

To summarize the application of parameter calibration in a policy analysis, I return to Figure 17, which illustrates an interesting paradox. First, as I have discussed, using calibrated parameters in the place of un-calibrated parameters has a very large effect on both greenhouse gas trajectories and the efficacy of a carbon tax. Clearly, the model is sensitive to values for the behavioural parameters, when the changes are substantial. Yet when uncertainty in these parameters is explicitly included in the analysis, the expected greenhouse gas trend is almost identical to the deterministic case. Though the posterior probability distributions have moderate variation in each parameter dimension independently, the distribution as a whole constrains the forecast quite substantially. Effectively, policy analysis for the furnace model seems to be strongly affected by calibrating parameters, but weakly affected by explicitly including uncertainty in parameter values. Part of this inconsistency may result from the fact that the uncertainty associated with consumers' response to a substantially different price regime under a carbon tax is not captured by the calibration. The calibration does not inform the model regarding consumers' response to prices outside the range of the historical calibration period.

CHAPTER 5: CONCLUSIONS

5.1 Modelling conclusions

In this study's calibration of three CIMS technology competition nodes, I have attempted to meet three specific objectives: to approximate probability distributions for behavioural parameters given historical data; to explicitly incorporate uncertainty into CIMS forecasts of future greenhouse gas emissions; and to identify key issues resulting in differences between historical trends and model forecasts. The extent to which I have met these objectives has several implications for future modelling work and policy analysis.

5.1.1 Estimating behavioural parameters

To meet the first objective, I used the calibration algorithm to generate posterior probability distributions for behavioural parameters for the residential furnace technology node and to a lesser extent, for the gasoline vehicle node. Posteriors incorporate both the likelihood of historical data and prior knowledge of parameter values from previous studies. They incorporate revealed preferences: when backcast over a historical period using the modes of these distributions, the model successfully replicates historical market share and stock trends. Further, these estimated distributions appear reasonable and generally consistent with past studies, even though only mildly informative priors were incorporated into the analysis. Estimated distributions for heterogeneity parameters, v , and private discount rate, r , in particular, are consistent with past results.

5.1.2 Incorporating uncertainty

I have addressed the second objective by explicitly incorporating uncertainty in behavioural parameters (as estimated through calibration) into a forecast-based analysis of a carbon tax. Interestingly, 90% probability bounds around the greenhouse gas forecasts were very narrow, especially for the business as usual case. This minimal effect of parametric uncertainty on forecasts may be due to strong correlation between parameter values. Alternatively, the minimal variation in forecasts may be specific to the

residential furnace node. The characteristic costs of technologies in this particular node may make it robust to parametric uncertainty. Calibration may also underestimate the uncertainty associated with consumers' response to higher energy prices associated with a carbon tax.

5.1.3 Assessing model structure

Analyses of the less successful calibrations of the refrigerator and vehicle node are also informative. In these cases, the results provide insight into qualitatively assessing structural uncertainty in CIMS, as per the final objective generated in Chapter 1. This insight comes primarily from the cases in which calibration proved *incapable* of matching historical trends; that is, parametric uncertainty was insufficient in explaining the model's deviation from history. For these analyses (namely the refrigerator and vehicle analysis) I have attempted to describe possible drivers for the discrepancy and thus to describe possible inadequacies in the model's structure.

In both the refrigerator and vehicle analyses, CIMS had difficulty replicating historical market shares, presumably as a result of exogenous changes in consumer preferences. In the vehicle node, for example, a linearly declining, time-dependent intangible was imposed to allow the parameter calibration to correctly replicate the historical shift in preferences toward light duty trucks and away from cars. Potential drivers for this shift include a perverse effect of the CAFE (corporate average fuel efficiency) standards, higher perceived collision safety in larger vehicles than smaller, high manufacturer profit margins, and shifting preferences.

The critical point here is that when running a forecast, a modeller could not have anticipated these exogenous shifts. In a forecast, CIMS represents only a baseline for consumer behaviour. Inevitably, it will not replicate shifts in behaviour driven by factors outside its framework of energy costs. Nevertheless, while this issue is problematic for calibration, it does not preclude the usefulness of forecasting. After all, good modellers and policy analysts explicitly recognize that a model forecast is uncertain. While unexpected exogenous effects might indeed drive actual future adoption of a given

technology well above the model forecast, other unexpected effects may instead result in future adoption trends well *below* the forecast.

5.2 Policy conclusions

Evaluation of less successful calibrations may also provide insight into the efficacy of historical policies. The inability of the backcasting model to replicate historical trends in refrigerator market shares, for example, implicitly suggests that demand-side management policies, especially those targeted at manufacturers might have resulted in an increase in refrigerator efficiency. Backcasts in both the residential refrigeration and furnace nodes suggested that government regulations mandating minimum efficiency standards may also have resulted in emissions reductions. This evidence may support assertions by Jaccard (2005b) that regulations and demand-side-management specifically targeting manufacturers can be effective. These hypotheses cannot, however, be tested within the context of this study. To do so would require showing the rate of autonomous energy efficiency increases in refrigerators between 1990 and 2003 was faster than in previous periods.

The policy analysis component of this study (Chapter 4) also has implications for policy-makers. On one hand, variation in consumer behaviour has a large effect on baseline emission trends and on the effectiveness of a carbon tax regime. Using calibrated and uncalibrated parameters in a forecast resulted in substantially different forecasts. On the other hand, explicitly incorporating uncertainty into a forecast-based policy analysis did not substantially affect the model's forecast of emissions reduced from household heating. The analysis thus may suggest that a carbon tax policy instrument could be somewhat robust to bounded uncertainty in consumer preferences ("bounded uncertainty" in this case meaning bounded by empirical data, historical prices, and the structure of the model used in calibration).

5.3 Limitations of the analysis and potential future research

Several key limitations in the analysis represent challenges to extrapolating the calibration results to CIMS forecasts. These limitations may represent opportunities for

additional analysis that might further solidify the empirical grounding of the behavioural elements of the model.

5.3.1 Accounting for the possible multiple parameter values

One limitation of this analysis is the existence of multiple parameter combinations capable of matching a backcast to historical data. All possible “good fit” parameter combinations are not specified. Choosing informative Bayesian priors, however, constrained the calibration’s search through parameter space for “good fit” combinations. These combinations of parameters are not selected exclusively as a function of the historical data; priors rely on subjective judgement. Still, the prior probability distributions are partially based on past empirical research (Rivers, 2003; Mau, 2005; Axsen 2006).

Applying a sequential experimental design methodology could provide new insight into the range of acceptable parameter combinations. This approach estimates a contour through parameter space showing all “good fit” parameter combinations in a given range. Some work is already progressing in customizing a sequential experimental design approach for calibration of CIMS. Dr. Pritam Ranjan, currently with the statistics department at Acadia University is developing code and generating a contour for the behavioural parameters in the residential heating node. These results will be compared to the results of this study, and further research on this alternative methodology could potentially be pursued as a result.

5.3.2 Data limitations and revealed preferences

A second issue stems from the limited historical market share data available. In many cases some assumptions were made in order to derive full historical time trends. For the vehicle node, for example, detailed model breakdowns from the year 2000 were extrapolated over more general time trends to determine historical market shares over the entire historical period. The quality of parameter posterior distributions is directly tied to the quality and quantity of the data to which they were calibrated.

Data also limited the CIMS nodes that could be calibrated. Calibrating a node in the industrial sector, in which purchase decisions were made by firms, rather than individuals and households, might have provided interesting contrasts in estimated parameters. Firms might be less heterogeneous and might have smaller discount rates. Unfortunately, no sufficiently detailed market share data for industrial equipment were available.

In calibrating the behavioural parameters to match model outputs to historical data, the model is essentially calibrating consumer's responsiveness to price signals. The historical period of 1990 – 2004, however, did not represent a period of extreme fluctuations in fuel price. While calibrating to a longer historical time period would provide substantial price variation, finding the detailed market share data required for calibration for periods prior to 1990 is even more challenging. The historical period is also very short. Capital stock inertia (*i.e.*, the fact that consumers often do not purchase new equipment until their existing car, furnace, or refrigerator is ready to be retired) suggests that significant time might be required for purchase data to reflect the impact of price on consumer preferences. The limited period of calibration (15 years) may be too short to capture this effect. For these reasons, the calibration may not fully represent the impact of price on preferences, and the calibrated model may fail to represent a consumer response to more dramatic price changes in the future if, for example, a carbon tax were introduced.

The issue of limited data emphasizes that calibration relies on a revealed preference approach. Policy analyses incorporating calibrated behavioural parameters can only accurately model consumer behaviour under similar economic conditions as existed in the historical calibration period. In this case, since energy prices did not vary substantially, the calibrated parameters do not necessarily represent consumers' responses to the strong price signal associated with a carbon tax. Similarly, the calibrated model will not necessarily represent consumers' choices when they are faced with new and emerging technologies (such as hybrid vehicles) that were not available in the historical calibration period.

One solution to this problem may be to extend the calibration to multiple geographic regions with multiple price regimes, and assume similar purchasing preferences over

these regions. The additional empirical data from multiple regions involved in this approach would add additional constraints to the calibration, potentially allowing for better assessment of the impact of price signal on consumer preferences. Further, consumers' actual responses to carbon taxes could be evaluated from data from regions such as Norway, Finland, or the Netherlands, which enacted carbon taxes in the 1990s. Added data may also better constrain the problem and reduce the need for informative priors.³⁰

5.3.3 Structural uncertainty and issues in extrapolation

An inability to reproduce historical data for the refrigeration and vehicle node highlights a third potential issue for calibration. The model's lack of the required dynamics under *any possible* parameter combination suggests the structure of the model is not an adequate representation of the mechanisms of consumer behaviour. In itself, this fact is not surprising; all models require assumptions and simplifications. And indeed, the failure of the model to replicate dynamics for refrigerator and vehicle choices allowed for an assessment of possible drivers for failure (such as the effects of difficult-to-model policies, dynamic intangible costs, or changing non-energy characteristics) and thus a qualitative assessment of model structure.

This identifiable structural inadequacy, however, emphasizes possible implications for more subtle effects of structural uncertainty. Oreskes (2003) argues that in calibration, even a structurally flawed model's parameters could potentially be tuned such that a model's forecasts successfully match historical observations. Because this calibration process deals only with parametric uncertainty, uncertainties in the structure of the model, or even in the data used for calibration, are ignored. Calibrating a model by adjusting parameter values forces the calibrated parameters to account for *all* uncertainty,

³⁰ I did in fact experiment with this approach by modeling residential furnaces in both Quebec and Ontario simultaneously. This calibration failed to generate meaningful results, however, as no parameter combinations existed that could match historical trends in *both* regions: combinations that resulted in a good fit for some technologies in one region resulted in a poor fit for other technologies in the other region. The calibration became over-determined. Formal spatial modeling techniques could be explored in future studies. More dramatic differences in price regimes (using for example, Canadian and European data) might produce more meaningful results, though differences in preferences due to culture might prove to be a barrier in this case.

not just parametric uncertainty. As such, it can be dangerous to interpret the calibrated parameter values as estimates of true values (Kennedy and O'Hagan, 2001).

Caution must therefore be applied in extrapolating these posterior distributions to a forecast over a future time period. For example, Section 3.4 describes how a linearly declining intangible cost on light trucks can help the model match historical trends, even though the linear structure is arbitrary and not based on a clear real-world mechanism. This structural uncertainty inhibits calibrating to “true” parameter distributions for the vehicle node. Over a future forecast period, the linear intangible structure might not be appropriate (indeed; it *cannot* be appropriate in the long term as the intangible would continue to decline forever; instead, some kind of asymptotic behaviour is more likely).

Future research might attempt to better assess the issues of structural uncertainty regarding intangible cost dynamics by running a calibration under several alternative “structures”. For example, the dynamic intangible could also be modelled as an inverse exponential function rather than a linear function to parallel previous research on the “neighbour effect” (Mau, 2005). This approach would involve calibrating additional parameters and creating additional model complexity. The tradeoffs between additional complexity (*i.e.*, additional model structure) and improving the model’s ability to recreate historical data could be assessed using a Bayesian Information Criterion (BIC). The BIC provides a quantitative metric for comparing alternative structures given historical data.

On a related note, this study’s assessment of structural uncertainty is qualitative only. In the cases of less successful calibrations, I have explored potential effects not included in the model as possible explanations for an insufficient representation of changing consumer preferences. In the refrigeration analysis, for example, I suggested that demand-side management policies targeted at manufacturers might have explained the shift toward more efficient fridges. While this might indeed be the case, the evidence presented here in the context of CIMS inability to reproduce this historical shift is circumstantial. As demonstrated in the refrigerator analysis, part of the problem in evaluating preference dynamics from historical data is that distinguishing the rate of autonomous energy efficiency from explanatory factors, such as policies and prices, is

difficult. As energy is an input to the economy with a positive cost, over time a competitive market will experience energy productivity gains through natural increases in efficiency.

Additional analysis would be required to confirm the efficacy of the “golden carrot” policy relative to the impacts of changing consumer preferences, or the significance of the CAFE regulations on the growth of SUVs relative to the significance of perceptions of size and safety. Formal multivariable econometric regression analysis, or evaluating rates of energy efficiency increases in earlier historical periods might provide further insight as to the most important historical causal effects.

5.3.4 No autocorrelation in parametric uncertainty

A fourth potential limitation in this study arises from the application of probability distributions in a stochastic policy forecast. The calibration algorithm estimates only a single probability distribution for the behavioural parameters. In including this uncertainty in a forecast, the same distribution is applied at each time interval in the forecasting simulation. However, if preferences are dynamic (as the results of this study seem to indicate is often the case), uncertainty should actually increase the farther the model projects into the future. Uncertainty about behaviour in 2040 should be much greater than uncertainty in 2010 given that the probability distributions were estimated from a historical period 1990 – 2005. In the stochastic forecast, this effect is not included, and the impacts of behavioural parametric uncertainty might therefore be under-represented.

An alternative approach that might handle this issue would be to calibrate parameters while assuming *non-stationarity* in the intangible costs. This approach would involve allowing the intangible costs to change slightly from their previous values in each simulated time-period using a constrained random walk formulation. The random walk would ensure that each new intangible value was a function of the previous value, ensuring autocorrelation in dynamic intangibles and allowing uncertainty to propagate

through the forecasted time trends. Parameters for the stochastic random walk function could also be calibrated from historical data.

5.4 Assessment of historical calibration for technology explicit models

Calibration in general has provided reasonable estimates for parameter values for the residential furnace and gasoline vehicle nodes. Nevertheless, issues such as insufficient price variation in historical data, an absence of revealed preference data for emerging technologies, and uncertainty in model structure are challenges for using calibration to inform consumer behaviour in policy forecasts. Behavioural realism is clearly important to the credibility of energy-economy models. Yet given the challenges, is historical calibration a useful way for modellers to establish an empirical basis for consumer behaviour in a technologically explicit model?

Since it relies on historical data, calibration through backcast faces all of the advantages and all of the disadvantages of a revealed preference approach. It is constrained to historical context and thus is based on real-world, unbiased data. Calibration is limited, however, in how much it can inform a model about future consumer choices, in which energy prices or available technology options may be very different. This issue was apparent in Chapter 4: I was able to apply only the residential furnace joint posterior probability distribution directly to a meaningful forecast, and even in this best case, issues of uncertain consumer responses to prices and pragmatic simplifications in model structure resulted in suspect forecasts and policy analyses. Stated preference approaches, on the other hand, rely on hypothetical survey responses subject to well-documented biases. The stated preference approach, however, is flexible and can explore consumer responses in various possible future scenarios. This flexibility would allow for much easier integration of estimated parameters into a meaningful policy forecast.

Revealed and stated preferences thus have complementary advantages and disadvantages. As suggested by Axsen (2006), empirical estimates of behavioural parameters are perhaps most credible when supported by both approaches. In the case of this study, the estimated parameters, particularly v and r , are consistent with other studies relying on

stated or combined approaches. Together, the correspondence between the results of this study and those from previous ones improves the credibility of the parameter values used in CIMS. Given the importance of representing real-world behaviour in energy economy models, a revealed preference approach to parameter estimation does therefore provide valuable – though not definitive – insight to modellers.

The specific methodology used in this study also has both advantages and disadvantages. I selected the MCMC approach to parameter estimation because it seemed to have two key advantages over other approaches. First, it took a Bayesian approach that allowed me to integrate prior knowledge regarding parameter values to supplement calibration with results from expert judgement and previous stated preference research. As I have suggested above, using both stated and revealed preference approaches improves the usefulness of empirical parameter estimation for forecasting. Second, the MCMC approach allowed me to generate posterior probability distributions for parameters and then to explicitly incorporate this parametric uncertainty into a policy analysis. The benefits of attempting to explicitly include parametric uncertainty in forecasts, however, may be outweighed by the costs. Though calibration generated a joint probability over all uncertain behavioural parameters, explicitly including correlation between parameters, the distribution is very difficult to apply to a forecast if any differences exist in the availability of technologies. Since backcasts inevitably do not include emerging technologies, this will almost always be the case. Further, the uncertainty analysis was only for a single node. Performing a similar analysis for the full CIMS model, or even a larger sub-component of the model, would require calibration of many parameters and a computationally and time-intensive sampling approach. Finally, CIMS is a tremendously complex model with many other uncertain parameters and many other types of uncertainty that might be more significant than the uncertainty in the behavioural parameters. Indeed, this analysis suggested the parametric uncertainty, bounded by calibration, did not substantially affect forecasted greenhouse gas emission. Given these facts, widespread explicit uncertainty analysis in CIMS seems neither practical nor useful. Analysts would be better served exploring parametric uncertainty through sensitivity analysis focused on identifying critical parameters and evaluating the ranges

of values of these parameters to which the model is robust. While this approach would not account for correlation between parameters, it is relatively simple to implement.

Because all model forecasts are wrong, especially ones based on human behaviour, no approach to forecasting future consumer behaviour is perfect. Historical calibration of technologically explicit models would be particularly useful (and most directly applicable to forecasting) if historical data are available for long time series with extensive price variation for technology options similar to those available in future periods. Overall, however, historical calibration is a useful tool for energy-economy modellers even if only limited data are available. Applying calibration in combination or in support of stated preference approaches, whether by comparing parameter estimates separately or integrating them in Bayesian approach, can help overcome the limitations of both a pure revealed preference approach and a pure stated preference approach.

Dowlatabadi and Oravetz (2006, p. 3251) raise the fundamental issue in historical calibration of whether the “past is a guide to the future.” Past behaviour is not necessarily consistent with future behaviour of consumers. Yet CIMS can still be a useful tool for policy makers without trying to anticipate every potential shift in consumer preferences. Shifts in preferences are inevitably exogenous to the model, and entirely uncertain. Future preference shifts might be toward low-emission technologies as a “green,” environmental ethos establishes itself. Alternatively, however, growing wealth might further increase consumptive values and societal greenhouse gas emissions. A useful model therefore must forecast in terms of a baseline; forecasts should provide a reasonable estimate of future behaviour. Given the lack of other available empirical data, generating “reasonable” forecasts means generating forecasts for the future that are consistent with the past and with the real-world revealed preferences of consumers. As such, this calibration takes a small step toward making CIMS forecasts more “reasonable,” more credible, and more useful to policy makers.

APPENDICES

Appendix A: The Bayesian framework in limited mathematical detail

Bayes' Rule

Applying Bayes' rule (Equation 6) we can generate a *posterior probability distributions*, $p(\theta|y)$, by combining a *prior probability distribution*, $p(\theta)$, with a *likelihood function* $p(y|\theta)$ for the historical data.

$$p(\theta | y) = \frac{L(y | \theta)p(\theta)}{p(y)} \quad \text{Equation 6}$$

Where, in the discrete case:

$$p(y) = \sum_{\theta \in \Theta} p(\theta) * P(y | \theta) \quad \text{Equation 7}$$

Or, in the continuous case:

$$p(y) = \int_{\theta \in \Theta} p(\theta) * P(y | \theta) \quad \text{Equation 8}$$

Such that:

$\theta =$ a vector of the behavioural parameters, $[v, i, r]$ where each of v, i, r is a sub-vector consisting of specific parameters for individual technology competitions.

$y =$ historical data; i.e. historical market shares of all technologies in a given node of CIMS

$p(\theta | y) =$ the posterior probability distribution for θ given the data y

$p(\theta) =$ the prior probability distribution for parameter values θ

$\Theta =$ the parameter space of all possible θ

$L(y|\theta) =$ the likelihood function (unscaled probabilities) of parameter values θ (the likelihood of data given parameters; a function of θ).

Prior probability distributions

Prior probability distributions (priors) are *a priori* representations of the probability of the ‘correctness’ of values of the behavioural parameters. That is, the prior expresses probability density as a function of parameter values.

$$p(\theta) = f(v, i, r) \quad \text{Equation 9}$$

For this study, normal distributions (within transformed parameter space, Θ) were used for priors, $p(\theta_i)$. The means of these distributions were set to the values currently used in CIMS, as determined through expert opinion and previous stated and revealed preference studies. Table 7 describes these values. Standard deviations for the priors were set such that the magnitude of the log-prior was small relative to the magnitude of the log-likelihood so as to ensure the priors were only mildly informative and did not dominate the posterior distribution.

Table 7: Previous parameter values used in CIMS

Node	Previous parameter values used CIMS for calibrated nodes		
	v	i	r
Refrigerators	10	0	0.6
Residential Furnaces	10	$i(\text{oil furnaces}) = \200 $i(\text{natural gas}) = \0 $i(\text{heat pump}) = \$0$	0.3
Gasoline Vehicles	10	$i(\text{low. eff. vehicle}) = -\3420 $i(\text{high eff. vehicle}) = \6555	0.22

Likelihood functions

The likelihood function, $p(y|\theta)$, for a given combination of behavioural parameters, θ_i can be determined from a historical run for the given parameter combination. Equation 10 represents one example of a common likelihood function, obtained by assuming normal distributions.

$$p(y_i | \theta_j) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{d_i^2}{2\sigma^2}\right)} \quad \text{Equation 10}$$

Such that:

$d_t^2 =$ The squared deviation between the historical market share and the model-forecast market share at time t .

$\sigma =$ The standard deviation of the market share data as

$\theta_j =$ Given values for v, i, r

As described in the text and shown in Equation 4, however, a modified expression for the posterior was developed that was not a function of the nuisance parameter, σ

Appendix B: Model results of forecast analysis

Table 8 and Table 9 illustrate the results of the models forecasts under the three modelling scenarios in greater detail. Table 8 reports forecasted market shares for the specific technology types at the beginning of the forecast (2005) under each scenario. Table 9 reports forecasted market shares for the final year of the forecast (2050) under both the business as usual (BAU) and carbon tax policy (POL) conditions. The two tables provide a sense of how the different parameters affect the forecasts as well as the trends of the forecasts from 2005 to 2050. Note that for scenario 3, the mean market share from the stochastic simulation is reported.

Table 8: Forecasted technology new market shares in 2005 under each of three modelling scenarios

Technology	New market shares for each modelling scenario in 2005 (%)		
	1) uncalibrated	2) calibrated; deterministic	3) calibrated; stochastic
Low efficiency oil	0.0	0.0	0.0
Med efficiency oil	13.5	7.4	7.2
High efficiency oil	0.0	0.0	0.0
Low efficiency natural gas	0.0	0.0	0.0
Medium efficiency natural gas	54.5	30.4	30.4
High efficiency natural gas	32.0	51.1	51.5
Heat pump	0.0	11.1	10.9

Table 9: Forecasted technology new market shares in 2050 under each of three modelling scenarios

Technology	New market shares for each modelling scenario in 2050 (%)					
	1) uncalibrated		2) calibrated; deterministic		3) calibrated; stochastic	
	BAU	POL	BAU	POL	BAU	POL
Low efficiency oil	0.0	0.0	0.0	0.0	0.0	0.0
Med efficiency oil	46.2	13.2	19.2	2.7	19.1	2.7
High efficiency oil	29.6	17.5	14.8	3.2	14.7	3.2
Low efficiency natural gas	0.0	0.0	0.0	0.0	0.0	0.0
Medium efficiency natural gas	13.0	28.4	14.9	6.0	15.0	6.1
High efficiency natural gas	11.1	39.5	27.2	12.1	27.5	12.3
Heat pump	0.0	1.4	23.8	75.9	23.6	75.7

REFERENCE LIST

- Automotive news data centre: U.S. car and light-truck sales 2004.* (2005).
www.automotivenews.com
- Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29(3), 155-173.
- Axsen, J. (2006). *Combining stated and revealed choice research to inform energy system simulation models: The case of hybrid electric vehicles.*, Burnaby B.C.: Simon Fraser University.
- Balakrishnan, S., Roy, A., Ierapetritou, M. G., Flach, G. P., & Georgopoulos, P. G. (2003). *Uncertainty reduction and characterization of complex environmental fate and transport models: An empirical bayesian framework incorporating the stochastic response surface method.* Unpublished manuscript.
- Bataille, C., Jaccard, M., Nyboer, J., & Rivers, N. (2007). Towards general equilibrium in a technology-rich model with empirically estimated parameters. *The Energy Journal*, (Special Issue), 93-112.
- Bataille, C., Rivers, N., & Peters, J. (2007). *Pathways and policies for long-term greenhouse gas air pollutant emissions reductions* (Final Report for The National Roundtable on the Environment and the Economy. Vancouver, Canada: J & C Nyboer.
- Bergman, L., & Henrekson, M. (2003). *CGE modeling of environmental policy and resource management.* Stockholm, Sweden: Stockholm School of Economics. Retrieved 02/25, 2005, from <http://www.ictp.trieste.it/~eee/workshops/smrl533/Bergman%20-%20Handbook-1.doc>
- Bohringer, C. (1998/6/1). The synthesis of bottom-up and top-down in energy policy modeling. *Energy Economics*, 20(3), 233-248.
- Boonekamp, P. G. M. (2007). Price elasticities, policy measures and actual developments in household energy consumption -- A bottom up analysis for the netherlands. *Energy Economics*, 29, 133-157.
- Canada. (2005a). *Comprehensive energy use database.* Ottawa, Canada: Natural Resources Canada: Office of Energy Efficiency.
- Canada. (2005b). *Energy consumption of major household appliances shipped in canada: Trends for 1990 - 2001.* Ottawa, Canada: Natural Resources Canada: Office of Energy Efficiency.

- Canada. (2006). *Energy use data handbook :1990 and 1998 to 2004* (5th ed.). Ottawa, ON: Natural Resources Canada.
- Canada (2007). *Guide to Canada's energy efficiency regulations*. Retrieved 09/09, 2007, from <http://oee.nrcan.gc.ca/regulations/bulletin/gas-furnace-mar2007.cfm?text=N&printview=N>
- Carraro, C., & Hourcade, J. C. (1998). Climate modelling and policy strategies. the role of technical change and uncertainty. *Energy Economics*, 20, 463-471.
- Casman, E. A., Morgan, M. G., & Dowlatabadi, H. (1999). Mixed levels of uncertainty in complex policy models. *Risk Analysis*, 19(1), 33-42.
- Challenor, P. (2006). *Uncertainties in the integrated assessment process -- A Tyndall Centre project*. Retrieved 12/2, 2006, from <http://www.floodrisknet.org.uk/a/2002/08/challenor-iam-uncertainty/>
- Consumer Reports. (2006). Refrigerators: Cooler ways to chill. *Consumer Reports*, 71(1), 40.
- Consumer Reports. (2004). Refrigerators: Playing it cool. *Consumer Reports*, 69(2), 36.
- Consumer Reports. (2002). Refrigerators test: For the chill of it. *Consumer Reports*, 61(1), 44.
- Consumer Reports. (2000). Refrigerators: Cold choices. *Consumer Reports*, 65(6), 41.
- Consumer Reports. (2000). Profiles of the 2000 cars. *Consumer Reports*, 65(4), 34.
- Consumer Reports. (1999). Refrigerators 1999. *Consumer Reports*, 64(8), 45.
- Consumer Reports. (1997). Refrigerators update test: The cold, hard facts. *Consumer Reports*, 62(2), 32.
- Consumer Reports. (1995).
The new refrigerators: How much energy do they save? *Consumer Reports*, 60(5), 310.
- Consumer Reports. (1994). Refrigerators for a wiser world. *Consumer Reports*, 59(2), 80.
- Consumer Reports. (1992). Refrigerators: A comprehensive guide to the big white box. *Consumer Reports*, 57(7), 456.
- Consumer Reports. (1990). Summary judgements of the 1990 cars. *Consumer Reports*, 55(4), 235.
- Cooley, T. F. (1997). Calibrated models. *Oxford Review of Economic Policy*, 13(3), 55-69.
- Denison, D. G. T., Holmes, C. C., Mallick, B. K., & Smith, A. F. M. (2002). *Bayesian methods for nonlinear classification and regression*. Chichester, England: John Wiley & Sons, Ltd.

- Dowlatabadi, H., & Oravetz, M. A. (2006). US long term energy intensity: Backcast and projection. *Energy Policy*, 34, 3245-3256.
- Eyzaguirre, J. (2004). *Simulating market transformation dynamics using a hybrid energy economy model: a look at the adoption of hydrogen fuel cell vehicles.*, Burnaby B.C.: Simon Fraser University.
- Frank, R. H. (1999). *Luxury fever :Why money fails to satisfy in an era of excess.* New York, NY: Free Press.
- Freedman, V. L., Waichler, S. R., Cole, C. R., Vermeul, V. R., & Bergeron, M. P. (2005). Identifying the potential loss of monitoring wells using an uncertainty analysis. *Ground Water*, 43(6), 916.
- Gelman, A., Carlin, J. B., Stern, H. A., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Godek, P. (1997). The regulation of fuel economy and the demand for "light trucks". *Journal of Law and Economics*, 40, 495-509.
- Groves, S. *Addressing the evolution of residential energy service demand for technologically explicit energy-economy models.* Unpublished manuscript.
- Hassett, K., & Metcalf, G. (1993). Energy conservation investment: Do consumers discount the future correctly? *Energy Policy*, (June), 710-716.
- Higdon, D., Kennedy, M., Cavendish, J. C., Cafeo, J. A., & Ryne, R. D. (2004). Combining field data and computer simulations for calibration and prediction. *Society for Industrial and Applied Mathematics Journal of Scientific Computing*, 26(2), 448-466.
- Hilborn, R., & Mangel, M. (1997). *The ecological detective :Confronting models with data.* Princeton, N.J.: Princeton University Press.
- Horne, M. *Incorporating preferences for personal urban transportation technologies into a hybrid energy-economy model.*, Burnaby, B.C.: Simon Fraser University.
- Hourcade, J. C., Jaccard, M., Bataille, C., & Ghersi, F. (2007). Hybrid modeling: New answers to old challenges. *The Energy Journal*, (Special Issue), 1-12.
- International Panel on Climate Change. (2001). *Modeling and cost estimate.* Retrieved 02/25, 2006, from
- Jaccard, M. (2005a). Hybrid energy-economy models and endogenous technological change. In R. Loulou, J. -. Waaub & G. Zaccour (Eds.), *Energy and environment* (pp. 1-29). New York: Springer.
- Jaccard, M., (2005b). *Sustainable fossil fuels :The unusual suspect in the quest for clean and enduring energy.* Cambridge, UK: Cambridge University Press.

- Jaccard, M., Nyboer, J., Bataille, C., & Sadownik, B. (2003). Modeling the cost of climate policy: Distinguishing between alternative cost definitions and long-run cost dynamics. *The Energy Journal*, 24(No. 1), 49-73.
- Jaccard, M., & Rivers, N. (2007). *Estimating the effect of the canadian Government's 2006-2007 greenhouse gas policies* C.D. Howe Institute Working Paper.
- Jaccard, M., & Rivers, N. (2007). *Canadian policies for deep greenhouse gas reductions (draft chapter)* Canadian Priorities Agenda Project: Institute for Research on Public Policy.
- Jaccard, M., Rivers, N., & Horne, M. (2004). The morning after: Optimal greenhouse gas policies for canada's kyoto obligations and beyond. *C.D. Howe Institute Commentary*, (No. 197), 1-35.
- Jager, W., Janssen, M., De Vries, H. J. M., De Greef, J., & Vlek, C. A. J. (2000). Behaviour in commons dilemmas: Homo economicus and homo psychologicus in an ecological economic model. *Ecological Economics*, (35), 357-379.
- Johnson, B. (1994). Modeling energy technology choices: Which investment analysis tools are appropriate? *Energy Policy*, 22(10), 877-883.
- Kann, A., & Weyant, J. P. (2000). Approaches for performing uncertainty analysis in large-scale energy/economic policy models. *Environmental Modeling & Assessment*, 5(1), 29.
- Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of Royal Statistical Society. Series B (Statistical Methodology)*, 63(3), 425-464.
- Koomey, J., & Sanstead, A. (1994). Technical evidence for assessing the performance of markets affecting energy efficiency. *Energy Policy*, 22(10), 826-832.
- Koomey, J., Sanstead, A., & Shown, L. Energy efficient lighting: Market data, market imperfections and policy success. *Contemporary Economic Policy*, XIV(July), 99-111.
- Kuhner, M. K. (1995). Estimating effective population size and mutation rate from sequence data using metropolis-hastings sampling. *Genetics*, 140, 1421-1430.
- Levine, M., Koomey, J., McMahon, J., Sanstead, A., & Hirst, E. (1995). Energy efficiency policy and market failures. Roundtable on energy efficiency and the economists -- six perspectives and an assessment. *Annual Review of Energy and the Environment*, (20), 535-555.
- Liu, J., Arndt, C., & Hertel, T. W. (2004). Parameter estimation and measures of fit in A global, general equilibrium model. *Journal of Economic Integration*, 19(3), 626-649.
- Loughran, K., & Kulick, J. (2004). Demand-side management and energy efficiency in the united states. *The Energy Journal*, 25(11), 19-43.

- Louviere, J., Hensher, D., & Swait, J. (2000). *Stated choice methods: Analysis and applications*. Cambridge: Cambridge University Press.
- Mau, P. (2002). *Modelling consumer preferences and technological change: Survey of attitudes to hybrid vehicles.*, Burnaby, B.C.: Simon Fraser University
- Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge, U.K.: Cambridge University Press.
- Moxnes, E. (2004). Estimating customer utility of energy efficiency standards for refrigerators. *Journal of Economic Psychology*, 25(6), 707.
- Norton, B., Constanza, R., & Bishop, R. (1998). The evolution of preferences: Why sovereign preferences may not lead to sustainable policies and what to do about it. *Ecological Economics*, (24), 193-211.
- Nyboer, J., & Bataille, C. (2007). *Energy equipment purchasing behaviour: A review of the literature*. Burnaby, Canada: Canadian Industrial Energy End-use Data and Analysis Centre.
- OECD (Organization for Economic Cooperation and Development). (2003). *Voluntary approaches for environmental policy: Effectiveness, efficiency, and usage in the policy mixes*. Paris: OECD.
- Ontario Ministry of Energy. (2007). *Heating and cooling your home: Chapter 5: Energy sources & equipment options*. Retrieved 09/09, 2007, from <http://www.energy.gov.on.ca/index.cfm?fuseaction=conservation.guide5>
- Oreskes, N. (2003). The role of quantitative models in science. In C. D. Canham, J. J. Cole & W. K. Lauenroth (Eds.), *The role of models in ecosystem science* (pp. 13-31). Princeton: Princeton University Press.
- Patz, R. J., & Junker, B. W. (1999). A straightforward approach to markov chain monte carlo methods for item response models. *Journal of Educational and Behavioral Statistics*, 24(2), 146-178.
- Peters, J. (2006). *The road ahead for energy-economy policy models: Integrating market dynamics, expectations and welfare costs*. Burnaby, B.C.: Simon Fraser University.
- Peterson, S. (2006). Uncertainty and economic analysis of climate change: A survey of approaches and findings. *Environmental Modeling and Assessment*, 11, 1-17.
- Ranjan, P., Bingham, D., & Michailidis, G. (2007). *Sequential experimental design for contour estimation from complex computer codes*. Unpublished manuscript.
- Rivers, N. J. (2003). *Behavioural realism in a technology explicit energy-economy model : The adoption of industrial cogeneration in Canada.*, Burnaby, B.C.: Simon Fraser University.

- Roberts, B. M. (1994). Calibration procedure and the robustness of CGE models: Simulations with a model for Poland. *Economics of Planning*, 27, 189-210.
- Ruderman, H., Levine, M., & McMahon, J. (1987). The behaviour of the market for energy efficiency in residential appliances including heating and cooling equipment. *Energy Journal*, 8(1), 101-123.
- Smith, L. A. (2003). Predictability past predictability present. *ECMWF Seminar on Predictability*, London School of Economics, London, UK. from http://lse.ac.uk/collections/cats/papersPDFs/56_PredictPastPredictPresent_2002.pdf
- Sutherland, R. (2000). "No cost" efforts to reduce carbon emissions in the U.S.: An economic perspective. *Energy Journal*, 21(3), 89-112.
- Sutherland, R. (1991). Market barriers to energy-efficiency investments. *Energy Journal*, 12(3), 15-34.
- Tanner, M. A. (1996). *Tools for statistical inference: Methods for the exploration of posterior distributions and likelihood functions* (3rd ed.). New York: Springer.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Thompson, P. (2002). Consumer theory, home production, and energy conservation. *Contemporary Economic Policy*, 20(1), 50-59.
- Train, K. (2002). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.
- Train, K. (1985). Discount rates in consumers' energy-related discussions: A review of the literature. *Energy*, 10(12), 1243-1253.
- Trucano, T. G., Swiler, L. P., Igusa, T., Oberkampf, W. L., & Pilch, M. (2006/0). Calibration, validation, and sensitivity analysis: What's what. *Reliability Engineering & System Safety*, 91(10-11), 1331-1357.
- Tschang, F. T., & Dowlatabadi, H. (1995). A bayesian technique for refining the uncertainty in global energy model forecasts. *International Journal of Forecasting*, 11, 43-61.
- U.S. Department of Transportation. (2004). *Summary of fuel economy performance* No. NVS-220). Washington, DC: U.S. Department of Transportation.
- United States Department of Energy. (1994). The MARKAL model. *National Energy Options for Reducing CO2 Emissions*, 1, 1-20-1-24, A-1 through A-9.
- Urban, G. L., Weinberg, B. D., & Hauser, J. R. (1996). Pre-market forecasting of really-new products. *Journal of Marketing*, (60), 47-60.
- Vanderheiden, S. (2006). Assessing the case against the SUV. *Environmental Politics*, 15(1), 23-40.

- Walters, C., & Ludwig, D. (1994). Calculation of Bayes posterior probability distributions for key population parameters. *Canadian Journal of Aquatic Sciences*, 51, 713-722.
- Wene, C. O. (1996). Energy-economy analysis: Linking the macroeconomic and systems engineering approaches. *Energy*, 21(No. 9), 809-824.
- Wing, I. S. (2006). Representing induced technological change in models for climate policy analysis. *Energy Economics*, 28, 539-562.