Analysis and Calibration of Social Factors in a Consumer Acceptance and Adoption Model for Diffusion of Diesel Vehicle in Europe

by

Qi Zhang

B.S. Materials Science

Fudan University, China, 2004

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Technology and Policy

at the

Massachusetts Institute of Technology

February 2008

Signature of Author..   ..
Technology and Policy Program, Engineering Systems Division
January 18th, 2008

Certified by...
John D. Sterman
Jay W. Forrester Professor of Management and Engineering Systems
Director, System Dynamics Group
Thesis Supervisor

Accepted by...
Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program
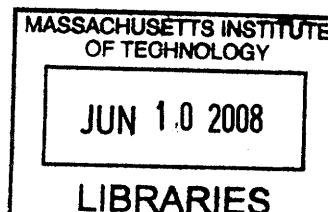
Analysis and Calibration of Social Factors in a Consumer Acceptance and Adoption Model for Diffusion of Diesel Vehicle in Europe

by

Qi Zhang

Submitted to the Engineering Systems Division

on January 18th, 2008 in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Technology and Policy

Abstract

While large scale diffusion of alternative fuel vehicles (AFVs) is widely anticipated, the mechanisms that determine their success or failure are ill understood. Analysis of an AFV transition model developed at MIT has revealed that AFV diffusion dynamics are particularly sensitive to consumer consideration as influenced by social exposure to AFVs. While some empirical research in this area exists, uncertainty regarding these social exposure parameters remains high. Following principles of partial model testing, this research examines social exposure parameters, with a focus on empirical accounts of diffusion involving diesel passenger vehicles in Europe. The research uses the historical data of diesel sales in six European countries. To complete diffusion datasets the research generates synthetic data in early stages of diffusion. The results from the calibrations yield parameters that are in line with other marketing studies. These findings help reduce uncertainty regarding social exposure parameters in the automotive industry. Further, bootstrapping confidence intervals are conducted to test the reliability of the parameter estimate. Challenges and avenues about building confidence in parameter estimate and data analysis are discussed.

Thesis Supervisor: John D. Sterman

Jay W. Forrester Professor of Management and Engineering Systems

Director, System Dynamics Group

2

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

Introduction

This thesis addresses the challenge of parameter estimation and confidence building in modeling consumer acceptance and adoption behaviors. Focused on empirical accounts of diesel passenger vehicle diffusion in Europe, it derives insights about alternative fuel vehicle transition. It also improves approaches in reconstructing data series and testing confidence intervals of parameter estimation.

Motivation

As the single largest vehicle market in the world, the United States heavily relies on petroleum fuels. Sixty-eight percent of petroleum fuels are consumed annually by the transportation sector (EIA 2006). The diversity of transportation fuel types is limited (Figure 1). As fuel demand continues to rise and domestic supplies diminish, the US depends more and more on fossil fuel imports, which have raised serious concerns of energy security in the public policy community. In addition, the increasing awareness of the environment impact of petroleum fuel emissions inspires scientists, researchers and policy analysts to explore alternative clean fuel options for the automotive market. These options include biodiesel, compressed natural gas (CNG), ethanol, hydrogen, and clean diesel.
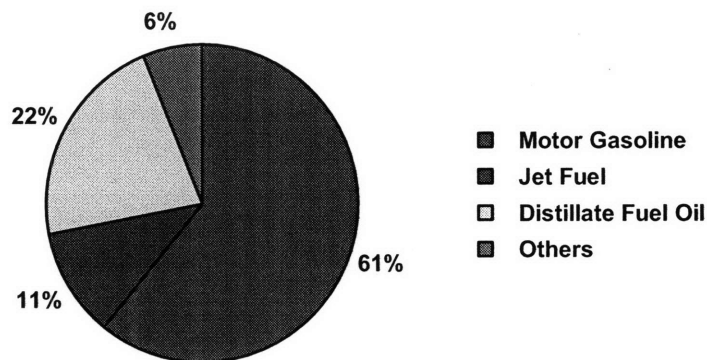
Figure 1 US Fuel Assumption in Transportation Sector in 2006
Source: EIA 2006

While large scale diffusion of alternative fuel vehicles (AFVs) is widely anticipated, the mechanisms that determine their success or failure are poorly understood mainly for two reasons. First, the diffusion process is a highly complex system with many stakeholders, strong feedback dynamics and long time delays, making it a significant challenge for the transition to a self-sustaining level. Second, there are only a few AFV diffusion cases worldwide that provide the empirical field for studies. Among the few cases, diesel diffusion in Europe is one of the success stories that has achieved large scale penetration while compressed natural gas (CNG) in New Zealand faltered after initial subsidies expired.

Analysis of an AFV transition model (AVMT) developed at MIT (Struben 2006, Struben and Sterman 2007) has revealed that AFV diffusion dynamics are particularly sensitive to consumers' decisions influenced by their social exposure to AFVs. While some empirical research in this area exists, uncertainty about these parameters remains high. A detailed empirical study is needed to reduce the uncertainty of these social exposure parameters, and to build confidence in this model before applying it to simulate other alternative fuel transitions.

Diesel diffusion in Europe is the natural model choice. The successful diffusion of diesel vehicles in Europe over the past thirty years provides rich data resources to make the calibration of social exposure parameters possible. The study of diesel diffusion patterns and dynamic behaviors could benefit research communities and policy analysts in terms of introducing other alternative fuel vehicles.

This work focuses on solving challenges in data completeness, parameter calibration and confidence tests of a dynamic behavioral model, and derives lessons for alternative fuel vehicle diffusion.

Audience

The approaches to data completion and confidence tests used in this thesis are helpful to SD researchers in academic communities who face challenges of building credibility in dynamic behavioral models. The derived insights about alternative fuel diffusion are of interest to

members of the public policy world who are concerned about fossil fuel dependence in the automotive market, including but not limited to policy analysts, legislators, and agency program managers, as well as business strategists who want to explore the future of alternative fuel vehicles in the corporate world.

Research Objectives

The study aims to (1) explicate the various patterns of diesel diffusion in Europe, particularly the role of social factors in the diesel transition, to (2) apply calibration methods to improve robustness and build confidence in parameter estimation in a complex dynamic model, and to (3) solve challenges in data completion, generation and analysis in an empirical research.

Chapter 2 Alternative Fuel Vehicle Transition: A System Dynamic Approach

Introduction to System Dynamics

System dynamics (SD) is an approach to interpreting the combinatorial complexity of problematic behaviors of a system, which often involve multiple stakeholders, dynamic interacting feedbacks, nonlinearities and time delays. To learn the dynamic behaviors of a complex system, SD researchers continue the process of formulating hypotheses, building models and simulating behaviors. The insights derived from simulation experiments can enhance the understanding of real world problems and help researchers design and evaluate solutions for behavior improvement.

Behaviors of a system are driven by the underlying structure of feedback loops (Forrester 1969). This structure is composed of information feedbacks, stocks and flows and nonlinearities introduced by the decision making of the agents involved (Sterman 2000). The theory that explains how endogenous consequences of the structure generate observed behaviors is the Dynamic Hypothesis (DH). The DH is the foundation of an SD simulation model and needs to be improved by model testing. SD model testing is an experimental process of building confidence in the DH so that the DH can be responsible for the structure of feedback loops and observed

behavior patterns (Oliva 2003). With enough confidence in the DH, researchers can derive insights from simulation experiments for problems of interest, design and evaluate policy solutions and help communicate this understanding to clients and other modelers.

Though no model is perfect due to people's subjective, limited, and simplified understanding of the real world (Sterman 2000), a robust model – one that matches observed behaviors and is confident in its DH and structure – is able to better inform the client and influence important decisions. Model testing is an iterative and ongoing process necessary in building a robust model. A substantial body of literature discuss a variety of tests on SD models (Forrester and Senge 1980, Barlas 1989, 1990, 1996, Sterman 2000, Oliva 2003) including boundary assessment, structure assessment, parameter assessment, extreme conditions, behavior reproduction, sensitivity analysis and calibration. Chapter 4 discusses the details of the model tests used for this study.

The following discussion provides a brief introduction to the field of system dynamics.
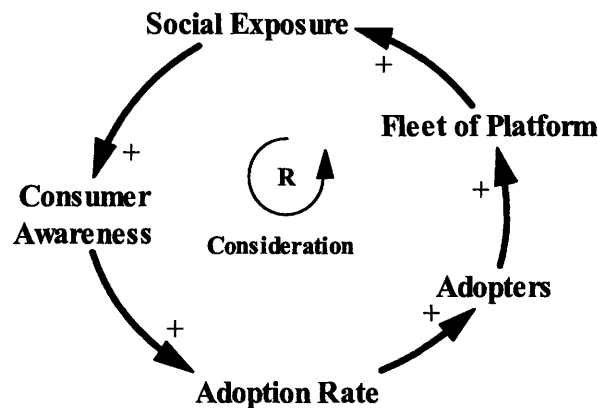
Figure 2 Feedback Loops — Reinforcing Loop

Feedback

Complex behaviors of a system arise from the interactions among stakeholders. The interactions are composed of two types of feedback loops: reinforcing loops and balancing loops. Reinforcing

9

loops amplify the existing influence in a system while balancing loops counteract the change in a system. Figure 2 shows an example of a reinforcing loop in a vehicle adoption system.

In the example, the larger installed base of a vehicle platform, the more social exposure a platform has to consumers, the more consumer awareness of the platform, and vice-versa. Consumer awareness drives consumer's transition to the platform at a fractional rate. Figure 2 and 3 are called casual loop diagrams. The directions of the arrows indicate the causal relationships. The + sign at the arrowhead indicates the effect is positive. R represents the polarity of a reinforcing loop.

Figure 3 shows an example of balancing loops. The more adopters for a product, the less potential consumers will adopt the product in the future. While the potential adopter pool shrinks, the negative effects kick in through the balancing feedback. This loop, B, counteracts changes in the causal chains. All the systems consist of a network of reinforcing loops and balancing loops. The interactions among these loops drive the dynamic behaviors of a system.
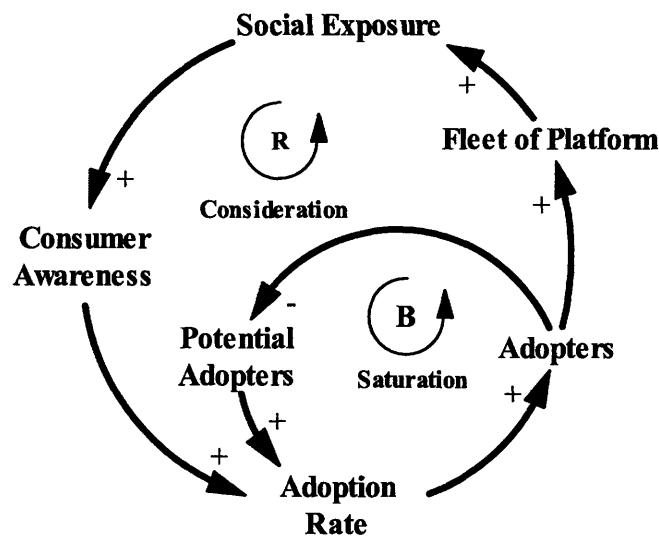
Figure 3 Feedback Loops – Reinforcing Loop and Balancing Loop

Stocks and Flows

SD conceptualizes system components as state variables (stocks) and rate variables (flows). Stocks are the accumulation of net flows (inflows less outflows). As a memory of previous dynamics, stock variables introduce a key dynamic concept into a system – time delays. Figure 4 applies stocks and flows structure to the vehicle adoption example.
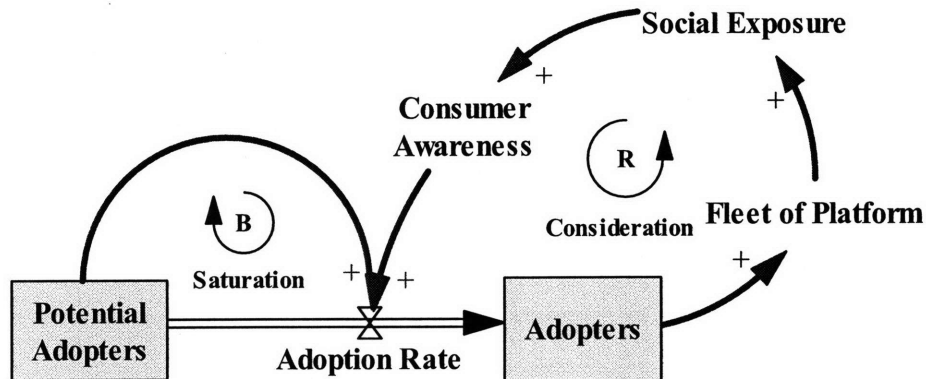
Figure 4 Feedback loops with Stock and Flow Structure

Challenges in Alternative Fuel Vehicle Transition

The increasing concerns of energy dependence, particularly petroleum imports, and greenhouse gas emissions from conventional vehicles, or internal combustion engine (ICE) vehicles, have made AFVs a favorable solution in the public policy domain. However, the successful diffusion of AFVs in the automotive marketplace faces more difficult challenges than many other durable goods for several reasons.

The disadvantage of being a late entrant is significant in this case. Emerging at the end of the 19th century, ICE has dominated the automotive market for over one hundred years. The vast ICE fleet and the enormous automotive industry create a wide range of powerful positive feedback effects on the dominance of ICE technology. Key positive feedbacks include economics of scale reflected in cost reduction and performance improvement, consumer awareness through word of mouth and marketing channels, learning by doing, R&D, technology spillovers and complementary services such as fuel infrastructure and maintenance.

11

As the ICE vehicle installed base achieves a sizeable market share, world of mouth from ICE drivers increases the consumers' awareness and adoption of the platform, and more investments are channeled into R&D, vehicle model development, infrastructure expansion and other areas so that ICE becomes even more attractive to consumers. On the contrary, the adoption of AFVs faces the "Chicken and Egg Problem" (Farrell et al. 2003, Ogden 2004, Bentham 2005, Struben 2006). For instance, automakers and component suppliers will not invest significantly in a new product unless they perceive a promising market potential. Consumers will not go after AFVs without a demonstration of their features and the costs of these investments. The chicken-and-egg dilemma is even severe for some AFV technologies that require new infrastructure networks, such as hydrogen.

Time delay is another important factor contributing to the challenges of the AFV transition. Because of the long lifespan of a vehicle, 16.9 years for the most recent model (Davis and Diegel 2006), consumer replacement purchases are delayed. The effects of consistent AFV transition efforts from the government and automakers may be visible only after decades. Time delays also exit in consumer awareness and adoption. It takes time for consumers to attain their knowledge, grow interest and develop emotional attachment to new vehicle platforms through personal experience, word of mouth and marketing channels. The existence of long physical and informational time delays requires policymakers and automakers to take a long-term view when designing AFV policies. But it is difficult to achieve due to people's cognitive limits such as linear and often short-term thinking.

Different from many durable goods, passenger cars serve not only a transportation tool, but also as a sign of social status and personal identity. Except for vehicle attributes, consumer choice is strongly influence by personal experience, social interaction and culturally shared views (Kay 1997, Hard and Knie 2001, Miller 2001). Consumer's perceptions are often heterogeneous in a physical and socio-economic space, which means that a successful AFV diffusion pattern in one geographic place may not work well in other areas. Other challenges in AFV transition include negative AFV innovation spillovers that may benefit conventional ICE technologies, huge experience lags between AFV and ICE technologies among production and service professionals,

and last but not least, the capital-extensive nature of AFV R&D and infrastructure investments. Existing diffusion studies, particularly the Bass model, perform poorly in capturing the complex dynamics in the AFV transition where multiple platforms compete and multiple stakeholders participate. The complexity of the AFV transition calls for an explicit diffusion model.

Alternative Vehicle Market Transition (AVMT) Model

To address these challenges in AFV diffusion, the AFV study at MIT System Dynamics Group has developed a behavioral, dynamic integrated model with a broad model boundary (Struben 2006; Struben and Sterman 2007), named Alternative Vehicle Market Transition (AVMT) model. This AVMT model describes the diffusion and competition among AFV platforms driven by endogenous interactions of consumer awareness, product attractiveness, infrastructure complementarities, learning by doing, R&D and innovation spillovers. The broad boundary of AVMT model enables analyses of the wide range of feedbacks that determine the AFV diffusion dynamics. The model builds on theory in a variety of fields including behavioral dynamic models (Forrest 1969, Sterman 2000), innovation diffusion (Rogers 1962, Bass 1969, Lekvall and Wahlbin 1973, Norton and Bass 1987, Urban et al. 1990, Urban et al. 1996) and discrete consumer choice (McFadden 1978, McFadden 2001, Ben-Akiva and Lerman 1985).

Figure 5 provides a conceptual overview of the model boundary and key feedback loops.
In this model, consumers make choices among platforms (e.g., ICE, hybrid, CNG) depending on the relative attractiveness of each platform in their consideration set. Individuals extend their consideration set to include a particular option only when they become sufficiently familiar with it. Familiarity is increased by social exposure through driver experience, word of mouth and marketing channels. Attractiveness of each platform depends on vehicle attributes, such as performance, cost, range, etc. Vehicle attributes are improved endogenously through learning by doing, R&D and complementary assets. With a broad boundary, the AVMT model represents the AFV transition with fidelity, and provides an experimental platform for in-depth tests of designed intervention policies at many points in the system. Struben and Sterman offer detailed model descriptions and documentations in their paper.

13
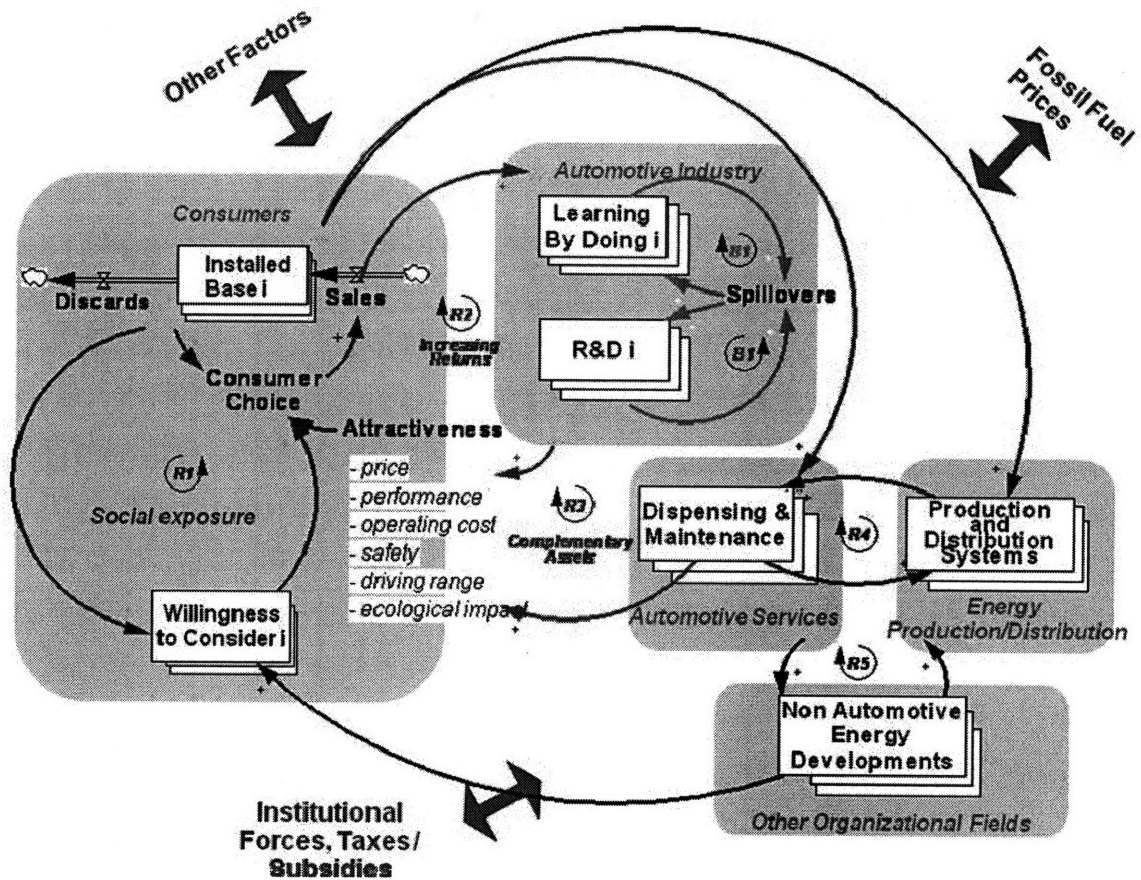
Figure 5 The AVMT Model Boundary

(Jeroen Struben 2006)

The dynamics of the AVMT model arise from several characteristics. First, this model captures the physical evolution of the vehicle installed base, vehicle technology, and refueling infrastructure. It also endogenously tracks the intangible changes in consumer perception and the learning experiences of automakers and infrastructure operators. The model structure of physical asset and information evolution brings time delays into the systems, which are particularly important in the real world AFV transition. Different from many durable consumer goods, vehicle diffusion is a slow process with time delays at various points. One influential delay in AFV transition is the time-consuming vehicle replacement that follows the vehicle life cycle. It takes years for any progress in consumer awareness, product attributes or policy incentives to produce visible effects on the market share of AFVs. Significant time delays also exist in the

14

formation of consumer opinions, the expansion of infrastructures, and the development of technologies. Taking all these time delays into account, the AVMT model is capable of reproducing a vehicle diffusion pattern in which time delays play a key role.

Secondly, the AVMT model is unique in addressing important spatial heterogeneity by endogenously modeling the co-evolution of infrastructure supply and vehicle demand, which is the underlying mechanism of the chicken-and-egg dynamics AFV transition. Jeroen Struben (2006) conducts a detailed analysis of the co-evolutionary dynamics with strategically locating fuel station entrants, which reveals important findings in a spatial context. For instance, the urban adoption clusters which often speed local diffusion in the early AFV diffusion stage can obstruct the formulation of a self-sustaining market on a larger spatial level.

Further, the AVMT model illustrates nonlinearities in the system, such as fuel availability on individual trip demand. The marginal benefits of a few additional fuel stations is much lower in the early diffusion period for consumers who want to make trips, but will increase dramatically as the number of fuel stations and vehicles increases, and returns to zero when the station space distributions are saturated. Therefore, the correlation between the marginal benefits of fuel availability and consumer trip efforts is more like an S shape than linear. The nonlinearities increase the level of difficulty for AFV diffusion to achieve a self-sustaining level in the early stage of market formation.

Last but not least, AVMT is a behavioral model capturing decisions of various stakeholders explicitly. Rather than assuming rational decision making with perfect information, this model represents how decisions are actually made by various stakeholders who use simple decision rules based on uncertainty or incomplete information. The behavioral elements in the model include consumers' choice to adopt an AFV platform, their trip plan, and their decision to go out of the way to find fuel, as well as a fuel station operator's decision to choose station entry and location, adjust capacity, and so on. Grounded in observed behaviors, the modeling of explicit decision making derives important insights about the dynamic complexity of AFV transition credited to multiple agents.

Consumer Consideration

The core structure of the AVMT model is consumer acceptance conditioned by consumer consideration and product attractiveness. The willingness of consumers to consider a vehicle platform is driven by their social exposure to that platform. Social exposures come from word of mouth and marketing channels.

Struben and Sterman (2007) demonstrate that word of mouth effects play an important role in the overall AFV diffusion. The larger the installed base of an AFV platform, the greater the social exposure to and familiarity with that platform among potential adopters, increasing the chances that they will choose that platform. Further, Sruben and Sterman point out that the word of mouth effects from non-AFV drivers are significant. These non-AFV drivers tell others about what they saw and thought about AFVs. The indirect effects of word of mouth from non-drivers may not have been as significant as those from AFV drivers at one time. But because of the long vehicle replacement cycle, the word of mouth effects from non-drivers have accumulated significantly over the years so that consumer awareness about the new technology continues to increase and spread even though these individuals do not need a vehicle at that time.

More specifically, Struben and Sterman find the existence of a critical threshold that the adoption of AFV must exceed to achieve self-sustaining. The threshold is particularly sensitive to the values of three factors – marketing programs for AFVs, the word of mouth impacts from AFV drivers and the word of mouth impacts from non-AFV drivers. These factors are crucial in determining whether the AFV adoption can be successful or not. However, the uncertainty regarding the magnitude of these effects remains high. An empirical study is needed to quantify and confine the values of these factors.

Diesel Diffusion in Europe

Diesel has become a major transportation fuel in the European market. Since their take-off in the mid-twentieth century, the market share of diesel passenger cars in many European countries has notably increased. The diesel share of new registrations in France and Spain exceeds 65% each year. Diesel penetration in Germany and Italy has grown at an average rate of 20% since 1991.

However, the diesel diffusion pattern is not uniform across European countries. Sweden did not experience an obvious transition from petrol to diesel, and its diesel share has struggled at below 10% for decades.
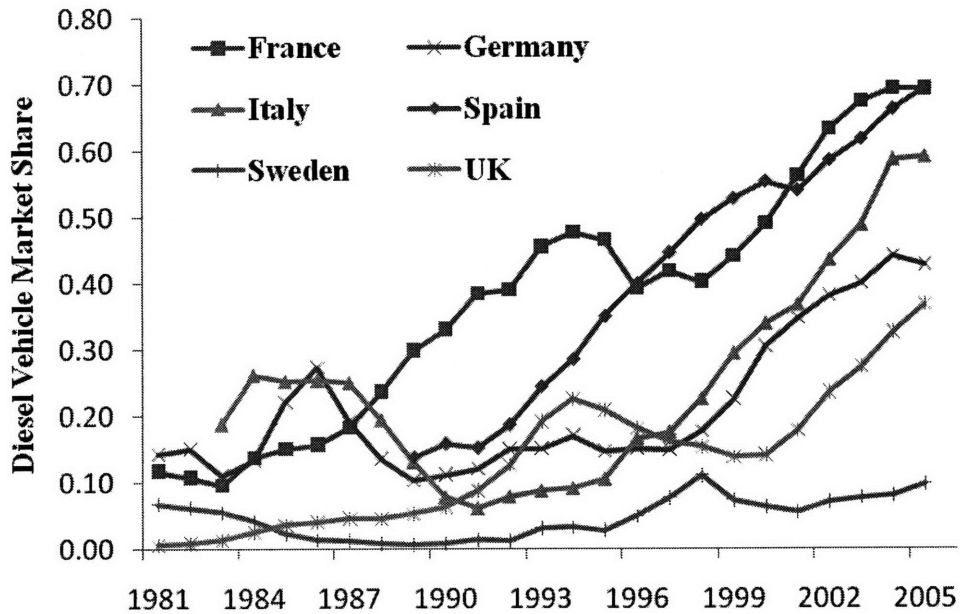
Figure 6 Diesel Share in Passenger Vehicle New Registrations
Source: PFC Energy

The rising share of diesel as an alternative to petrol fuel is credited to policy incentives and consumer acceptance in these countries. Many European governments are known to leverage policies to favor diesel fuel. As a result, diesel fuel is significantly cheaper than petrol in those countries, typically at a 20% discount (Chen and Sperling 2004). The substantial fuel expenditure savings as well as tax benefits associated with purchasing and owning a diesel vehicle offset its higher retail price. Consumer consideration is another key contributor to the remarkable rise of the diesel population. Consumers in European countries are generally familiar with diesel vehicle technology from their extensive exposure to diesel from word of mouth among fellows as well as marketing programs. Consumers' common perception of diesel as dirty, noisy, heavy and polluting has changed as technology advances in improving vehicle performance. The increased utility of a diesel platform attracts an ever broader range of consumers, which economically motivates automakers and infrastructure operators to build

17

business around the diesel platform. The attractiveness of a diesel platform expands as diesel models become widely available, vehicle performance is improved through R&D, and refuel infrastructure network is well constructed. Driven by these powerful positive feedbacks, the diesel population in countries like France, Italy and Spain has outsold petrol models.

However, diesel shares are not uniform across Europe. In Sweden, higher diesel fuel prices are discouraging diesel sales as well as tax policies favoring gasoline over diesel (Chen and Sperling 2004).

The long history of the diesel transition in European markets provides quality data for an empirical study. The similar social-economic environments in these countries diminish the economical and cultural reasons on the difference of diffusion patterns. Vehicle attractiveness conditioned by country specific policies as well as consumer consideration driven by local marketing channels contribute to the main difference in the diffusion paths taken by these countries. Moreover, the similarity of mobility behaviors of consumers in European countries – how often consumers drive, their driving distance, the way they communicate with fellow drivers – enables the research on the general, non-country specific factors, such as word of mouth effects across different diffusion patterns.

This study collects and reconstructs diesel sales empirical data from France, Germany, Italy, Spain, Sweden and the United Kingdom from 1970 to 2005 (Chapter 3). Using the AVMT model, the study explores the different patterns of diesel diffusion in these six countries, particularly focusing on the role of the social factors in the diesel transition. Model calibration is conducted to estimate the influence of social factors in determining the diffusion pattern (Chapter 4). Confidence intervals of the parameter estimations are tested using the bootstrapping method (Chapter 5). Conclusions and recommendations are provided in Chapter 6.

Chapter 3 Data Collection and Generation

Data Collection

This study requires empirical data beginning in 1970's, when diesel diffusion started to take off in Europe. So the year 1970 is considered to be a reasonable starting year for estimation. Both the installed base and new registrations can be calibration targets though the installed base data, a stock value, are less volatile and less noisy than new registrations data, a flow variable.

This work collects data on new registrations, installed base and other related data of France, Germany, Italy, Spain, Sweden and the United Kingdom from various resources including international data organizations, such as Eurostat, and independent vehicle industry information agents, such as Polk and PFC Energy. Periods of available data vary: new registrations from 1981 to 2005 from Polk; installed base and new registrations from 1992 to 2004 from PFC Energy; installed base from 1979 to 2004 from Eurostat.

Table 1 Data Resource

Platform	Source	Usage	Data Available	Data Needed	Data Gap
Diesel					
New Registrations	Polk	Calibration	1981-2005	1970-2005	1970-1980
Installed Base	PFC Energy	Calibration	1992-2004	1970-2005	1970-1991
Total					
New Registrations	Eurostat	Model Input	1979-2004	1970-2005	1970-1978
Installed Base	Eurostat	Model Input	1979-2005	1970-2005	1970-1980

As shown in Table 1, there are substantial gaps between numerical data that are available and data that are needed: Eurostat data not available at fuel type level; Polk and PFC Energy data missing early diesel diffusion. The data gaps can reduce accuracy and credibility of calibration results. In order to fill the data gap during the early adoption stages, this study designs two approaches to generate synthetic datasets of diesel new registrations and installed base in the early diffusion.

Generation of Synthetic Datasets

This study uses two approaches to estimate diesel new registrations and diesel installed base in early stages. The first approach is a manual trend extrapolation while the second approach is an automated computation. Approach One is often called "calibration by hand." This approach sets a parameter, and then adjusts the value of this parameter to generate outputs that match with historical data. The steps of this approach are described as follows.

Approach One: Trend Extrapolation

(1) Use diesel new registrations data from 1981 to 2005 from Polk. Set twelve years since 1981 as the reference period, new registrations at 1992 as the reference price. Derive the growth trend of diesel new sales based on the reference period and the reference price. The growth rate is adjusted by parameter g_s.

(2) Assume the initial diesel installed base V_0 at 1990 is equilibrium, meaning replacement purchases equally compensate vehicle discards. Set the average vehicle life at 15 years. Therefore, the initial diesel installed base at 1970 is equal to new registrations at 1970 multiplied by average vehicle life T adjusted by the growth rate of the installed base g_V during a vehicle lifespan.

(3) Add the previous year's installed base V_{t-1} to new registrations s_t and distract discards of that year d_t to get the installed base for the years after 1970, V_t.

(4) Assume discards d_t from 1971 and 1985 (one vehicle life cycle after 1971) is equal to new registration s_t minus previous year's installed base (V_{t-1}) multiplied by the growth rate of installed base g_V. Discard after 1985 is equal to new registrations s_{t-T} one vehicle life cycle ago. (5) Manually fit synthetic diesel installed base to historical data from PFC Energy by altering values of g_s and g_V. Repeat step (1) to (5) till a reasonable fit between data and model output.

Figure 7 shows the manual calibration results for Germany. Though Approach One is a straightforward method, it has several limitations. First, the extrapolation sales trend is conditioned by the choice of the reference price and the reference horizon. The subjective picking of reference points can introduce biases into the estimate. Second, "calibration by hand"

is often not accurate when there are multiple parameters. A manual manipulation is not able to optimize several parameters simultaneously, which limits its ability to search through all the value space to find the best fit scenario. Finally, it is difficult to repeat the results of a manual calibration. Since the best fit is judged by visual observations, a replication effort can hardly derive original parameter estimations.

In order to avoid the problems of "calibration by hand", this study designs a structure to generate synthetic datasets by automatic computation. The idea is to build a data derivation model in Vensim and take advantage of Vensim's optimization function to search for the best estimation for the growth rate of diesel new registrations prior to 1981. With an estimated diesel sales growth rate and new registration between 1970 and 1981, initial diesel installed base at 1981 can be derived, serving as a start point for Chapter 4 calibration.

The core of this data derivation model is a vehicle aging chain (Figure 8). This structure divides total vehicles into three categories, or cohorts: new, mature and old. The stock of new installed base is an accumulation of new registrations less discards. Aging vehicles are transitioned to the mature installed base, and then the old installed base. Vehicles surviving after the first two cohorts are inputs into old installed base and falls through discards eventually. The transition rate is determined by the average aging time for this stock and vehicle survival rate. Vehicle survival rate for new installed base is 0.97 and for mature installed base, 0.83 (Cohen and Greenspan 1999).

**Total New Registrations
Germany**

**Diesel New Registrations
Germany**
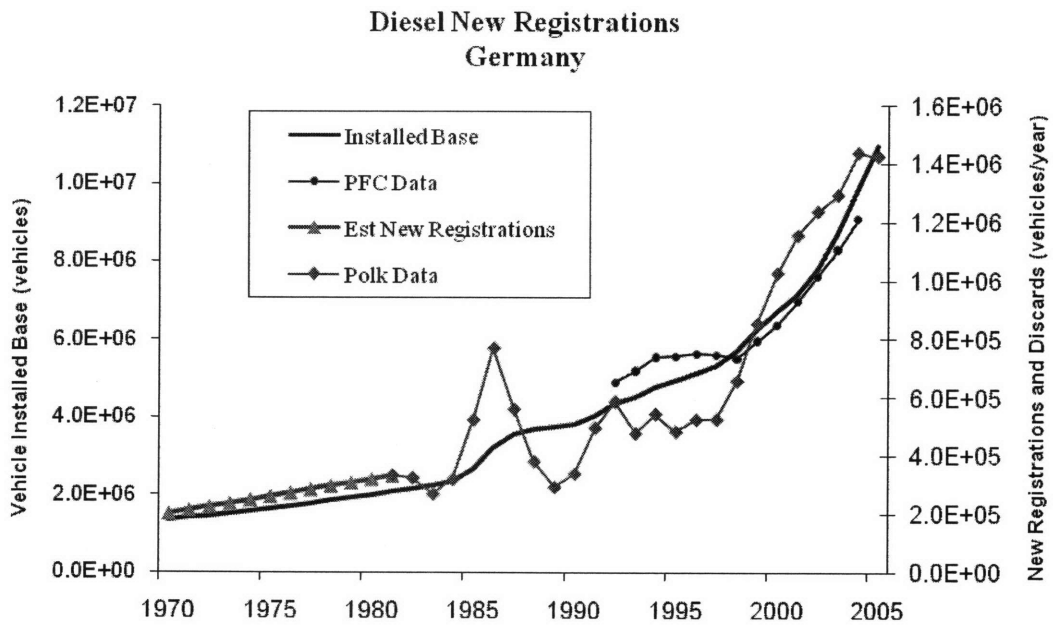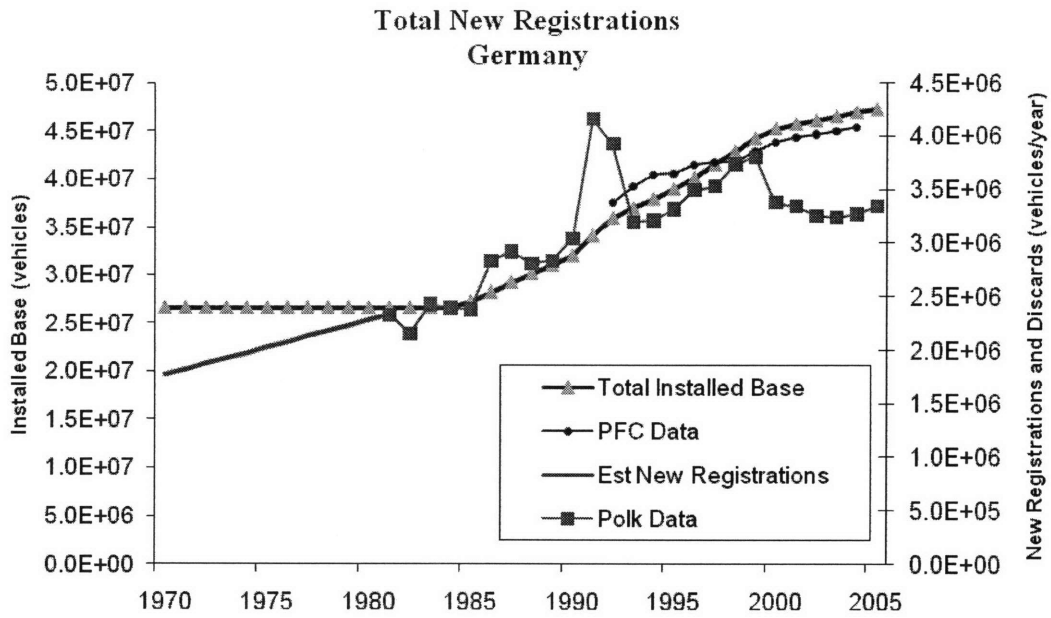
Figure 7 Manual Calibration Results

Figure 8 Vehicle Aging Chain

Installed Base Total p

Installed Base New

New Registration

New Vehicle Aging

New Discard

Vehicle Cohort Time

New Vehicle Scrappage Rate

New Vehicle Survival Rate

Installed Base Mature

Mature Discard

Mature Vehicle Aging

Mature Vehicle Scrappage Rate

Mature Vehicle Survival Rate

<Vehicle Cohort Time>

Installed Base Old

Old Vehicle Discard

Old Vehicle Cohord Time

Figure 8 Vehicle Aging Chain

Figure 9 Data Derivation Structure

Diesel Sales Growth Rate pre Polk

UPLOAD New Registrations Diesel

UPLOAD Installed Base Total Next Year

New Registrations Diesel

Installed Base Diesel

Discard Diesel

Installed Base Total

Discard Rate

New Registrations Petrol

Installed Base Petrol

Discard Petrol

New Registrations Total

Discard Total

Figure 9 Data Derivation Structure

23

Figure 8 is the diagram of the data derivation structure. The vehicle aging chain applies to modeling both diesel and petrol installed base. Steps of Approach Two are as follows.

Approach Two: Data Derivation

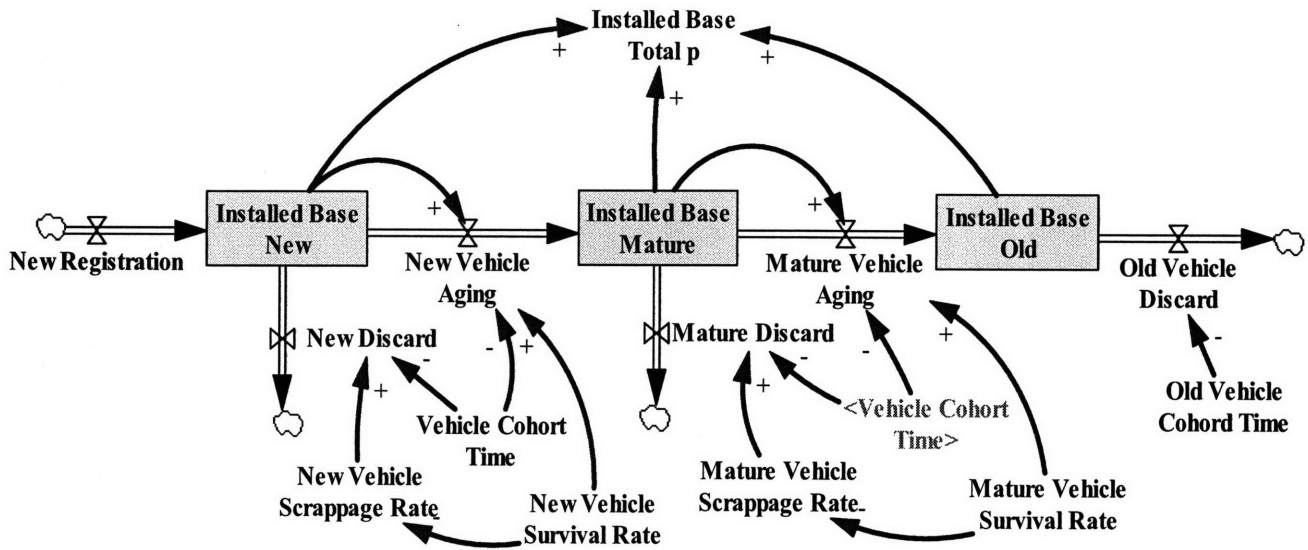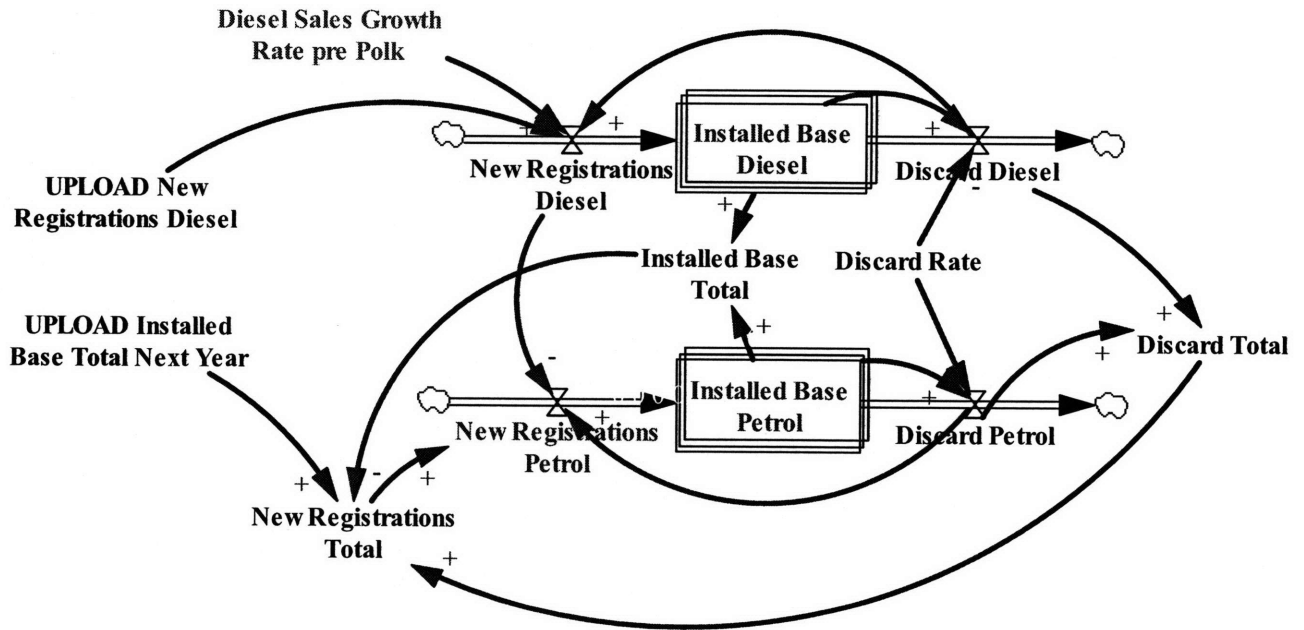(1) Set the diesel sales growth rate prior to first year of Polk data, 1981, as a parameter g_{pp}. Extrapolate the diesel sales trend prior to 1981 using g_{pp} and the diesel sales data at 1981. Upload both synthetic datasets and historical diesel data into the diesel aging structure.

(2) Set petrol new registrations as total new registrations minus diesel new registrations. Total new registrations is equal to total installed base at year t+1 minus total installed base at year t plus total discards. Total installed base at year t is the sum of diesel installed base and petrol installed base.

(3) Define payoff elements in Vensim's optimization function. Payoffs are the variables to fit. In this case, compare petrol new registrations to uploaded petrol new registrations, compare diesel installed base to uploaded diesel installed base, and compare petrol installed base to uploaded petrol installed base. Set 1 to the weight of each payoff element.

(4) Choose Diesel Sales Growth Rate prior to Polk, g_{pp}, for Vensim's optimizer to vary, along with the maximum search bound, 1, and the minimum bound, 0, for it.

(5) Run the optimization based upon least squares principles. The difference between the data and the model variable is multiplied by the weight specified and this product is then squared. This number is then subtracted from the payoff. The optimization process is to maximize the payoffs.

As a result, the optimized diesel growth rate prior to 1981 for France is 0.10, Germany 0.15, Italy 0.35, Spain 0.50, Sweden 0.15 and the UK 0.05. The diesel installed base in 1970 for France is 600,000, German 600,000, Sweden 40,000 and the UK 50,000. Italy and Spain do not have a diesel presence in 1970, so both are zero. Diesel new registrations prior to 1981 are calculated using the optimized growth rate and new registrations data in 1981.

Figure 10 shows that Approach Two generates a smooth trend of diesel new registrations growth from 1970 to 2005. The estimated diesel share (diesel installed base divided by total installed base) in Figure 11 also follows the historical pattern pretty well (Figure 2). Therefore, the further analyses will use the synthetic data estimated by Approach Two.
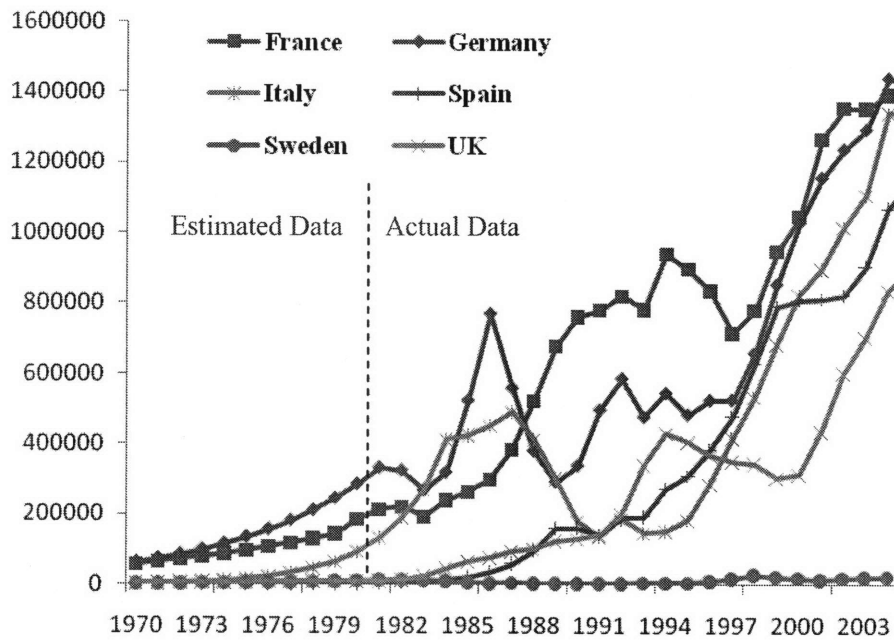
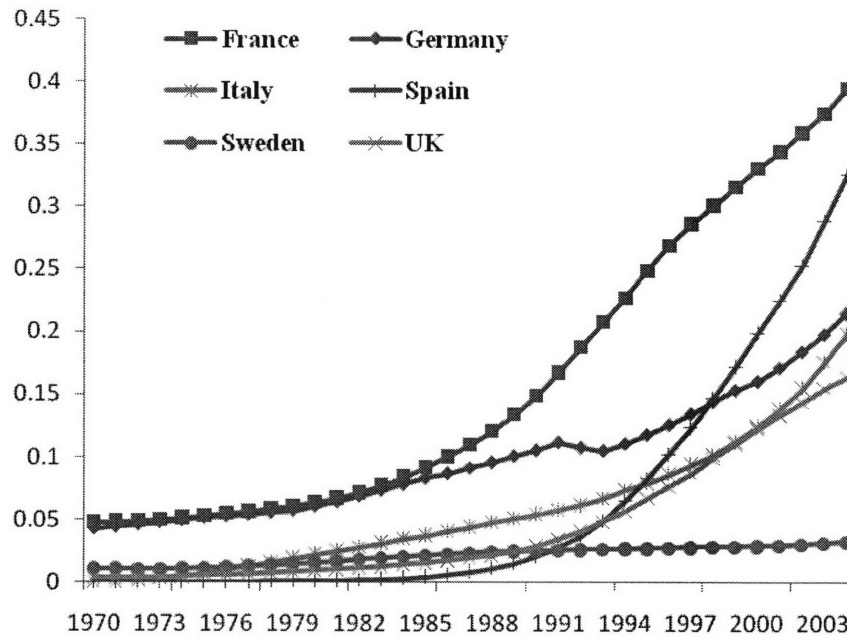**Figure 10 Diesel New Registrations – Actual Data and Estimated Data**

Figure 11 Diesel Share Installed Base – Estimated Data

Chapter 4 Model Calibration

Purpose of Calibration

SD practitioners build simulation models to learn how decision making and casual chains generate problematic behaviors observed in the real world and to design solutions for solving problems. Modeling, a powerful tool of learning, is an iterative process of building confidence in the dynamic hypothesis (DH) - formulating the DH, building models, testing models and revising the DH. The purpose of model testing is not to show how successfully one model is able to pass various tests, but to discover the flaws and limits in the DH and the model structure, and improve the suitability of the model for the study purpose (Sterman 2000).

Many studies discuss a variety of tests on SD models (Forrester 1973, Forrester and Senge 1980, Barlas 1989, 1990, 1996, Sterman 2000). Among this wide range of tests, parameter assessment is the one most often receiving modelers' central attention. Parameters are variables with

constant values. The formation of parameters is grounded in real life meanings, whose values impact simulation outcomes.

Model calibration is the process of estimating model parameters by fitting model simulations to empirical data. Empirical data can be numerical data points or judgments grounded in the knowledge of real world practices. Calibration in SD models can be done by hand (Lyneis and Pugh 1996). However, the manual calibration is not feasible in situations where multiple parameters are to be estimated or multiple data series are to be fit. Popular SD software, such as Vensim, now provides automatic computation packages that enable the search for parameter values in a given space for a best fit of model outputs to empirical data. Based on optimization methods such as nonlinear least squares and Kalman filtering, the automated computation is able to handle nonlinear feedback in SD models. The automated computation for parameter estimation was used in Chapter 3 to restructure synthetic datasets. This chapter will apply this method for estimating social factor in the AVMT model.

The AVMT Model for Diesel Study

The focus of this study is to understand social factors involved in diesel diffusion. For this purpose, the study develops a simplified version of the AVMT model. The consumer consideration loops are isolated from the full AVMT model following partial model test principles (Homer 1983, 1996, 1997). The simplified model reduces treatments of the exogenous vehicle attractiveness loops by assigning fixed values to vehicle attributes. The network effects and other feedbacks are cut off. Figure 12 shows the elaborated loops for the diesel diffusion case. In the simplified AVMT model, consumers' transition from petrol to diesel is purely influenced by their social exposure to diesel and the diesel's attractiveness relative to petrol. Total social exposure to diesel consists of word of mouth from diesel drivers, word of mouth from non-diesel drivers and marketing program effects. Word of mouth effects are endogenous while marketing effects are exogenous. This simplified model is referred to as the AVMT consideration model in the following context.
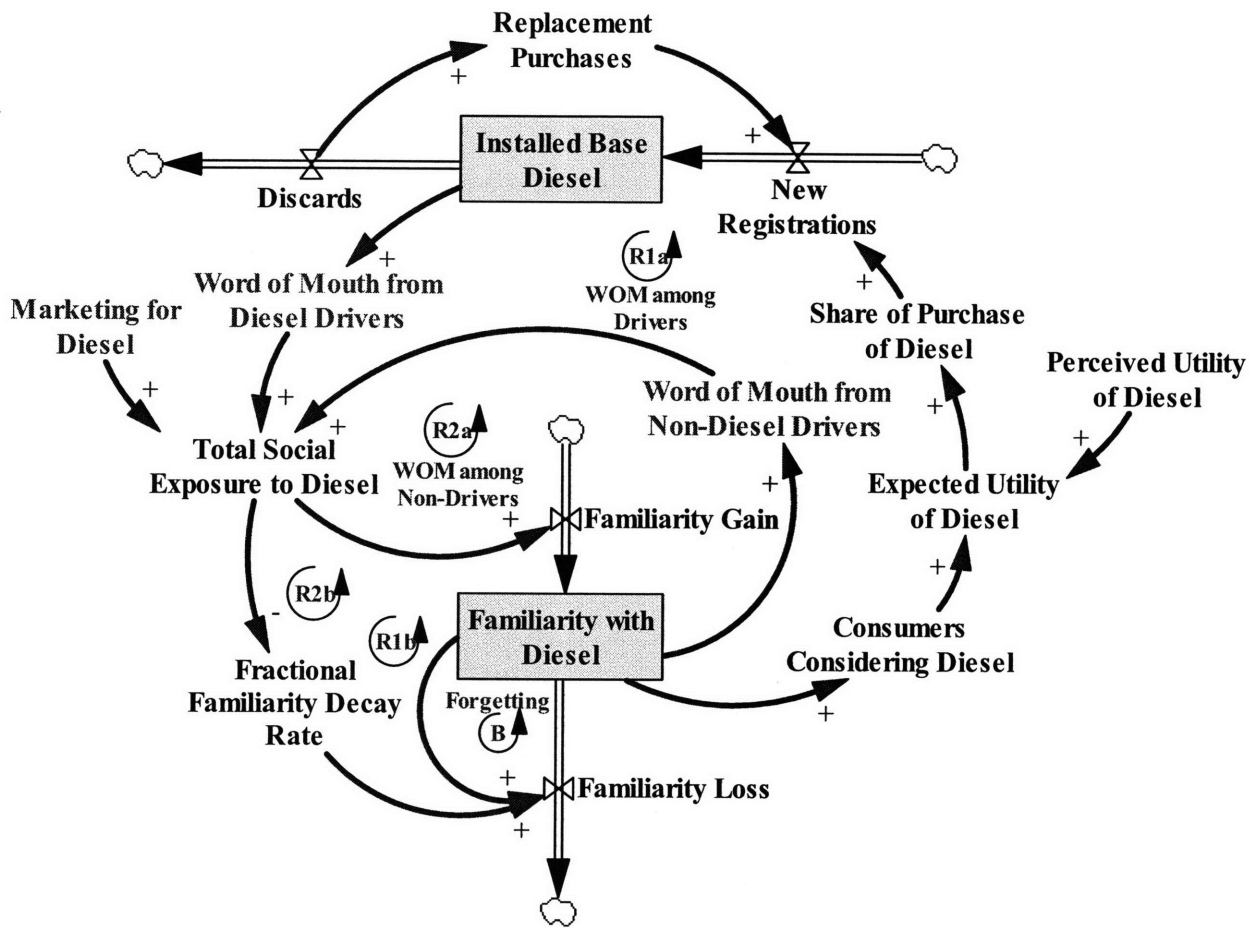
Figure 12 AVMT Consideration Loops
(Modified from Jeroen Struben's model)

R1 Word of Mouth from Diesel Drivers

R1a: The likelihood of consumer considering diesel grows once consumers become familiar with this platform. The greater likelihood that consumers will consider diesel, the higher the expected utility of diesel is, which determines the share of new purchases to diesel platform. The size of diesel installed base drives word of mouth effects from diesel drivers and thus the total social exposure to diesel contributing to familiarity gains. The accumulation of familiarity gains and loss determines the likelihood that consumer will consider diesel.

R1b: People's memory and attention to the diesel platform, adding another positive effect on familiarity.

R2 Word of Mouth from Non-Diesel Drivers

R2a: The greater the familiarity with the platform, the more word of mouth effects among non-diesel drivers arise, which contributes to the total social exposure effects and familiarity gains.

R2b: An increasing word of mouth effect also prevents consumers' attention of the diesel platform from decaying, adding another positive effect on familiarity accumulation.

B Forgetting

Familiarity loss is conditioned by familiarity decay rate and familiarity level. The relationship between familiarity loss and social exposure is nonlinear. Familiarity fades rapidly with little social exposure. The decay rate will reduce as social exposure increases, and will reach zero when social exposure is really frequent.

Social Exposure and Vehicle Attributes

Struben and Sterman (2007) introduce the detailed formulation of consumer consideration. The following is a brief review of this structure modified to this diesel case. For simplicity, this analysis only considers the competition between platform petrol (p) and diesel (d). Consumers can choose from two platforms $j = \{p, d\}$. The vehicle installed base for diesel, V_d, is an accumulation of new vehicle sales, s_d, less discards, d_d.

$$\frac{dV_d}{dt} = s_d - d_d \tag{1}$$

Consumers can choose to switch between diesel and petrol platform when they purchase a new vehicle. Diesel new sales come from two resources - diesel drivers and drivers switched from petrol models. This model assumes that new vehicle sales are all replacement purchases, which is often seen in developed economics. Assume no population growth, $\sum_j s_j = \sum_j d_j = N$.

$$s_d = \sigma_{pd} d_p + \sigma_{dd} d_d \tag{2}$$

where σ_{pd} is the share of new purchases made by petrol drivers to diesel platform, and σ_{dd} is the share of diesel drivers continue purchase diesel vehicles. As an indicator of diesel transition, σ_{pd} is determined by the expected utility of diesel perceived by petrol drivers as a fraction of total expected utility of platforms.

$$\sigma_{pd} = \frac{u_{pd}^e}{u_{pd}^e + u_{dd}^e} \tag{3}$$

The expected utility depends on two factors: familiarity and vehicle utility. As mentioned earlier, familiarity represents the likelihood that a platform will be considered by consumers. Vehicle utility depends on vehicle attributes of a platform perceived by drivers.

$$u_{pd}^e = F_{pd} * u_{pd} \tag{4}$$

where F_{pd} is the familiarity of petrol drivers to diesel platform, and u_{pd} is diesel utility perceived by petrol drivers. Familiarity increases as social exposure increases and decays over time. Hence

$$\frac{dF_{pd}}{dt} = n_{pd}(1 - F_{pd}) - \rho_{pd}F_{pd} \tag{5}$$

$$F_{pd} = F_0 \text{ when } t = 0$$

Hence n_{pd} is the effect of social exposure on petrol drivers' familiarity about diesel, and ρ_{pd} captures the fractional loss of familiarity about diesel among petrol drivers. Social exposure of petrol drivers to diesel consists of three elements: marketing effectiveness, word of mouth effects from diesel drivers, and word of mouth effects from petrol drivers. Hence

$$n_{pd} = \alpha_d + c_{pdd}F_{pd}(V_d / N) + c_{pdp}F_{pd}(V_p / N) \tag{6}$$

where α_d is the effectiveness of marketing efforts for diesel platform. Word of mouth effects from diesel drivers happen through contacts with diesel drivers. The possibility of meeting diesel drivers is (V_d / N). C_{pdd} is the effective contact rate of petrol drivers with diesel drivers. The direct word of mouth effects from diesel drivers are captured in $c_{pdd}F_{pd}(V_d / N)$. Besides, word of mouth effects can also come from petrol drivers who see diesels on the road, and meet and

discuss the platform with fellow petrol drivers. $c_{pdp}F_{pd}(V_p / N)$ represents the indirect word of mouth effects from petrol drivers.

Aside from familiarity, diesel utility (perceived by consumers), u_{pd}, is another variable shaping consumer choice. u_{pd} is conditioned by the availability of vehicle models and the attractiveness of vehicle attributes (Figure 13). The diesel portfolio measures the number of diesel vehicle models available on the market. Vehicle attributes consist of vehicle price, vehicle performance, fuel availability, operation cost and diesel technology effect. Vehicle price takes registration tax into account. Operation cost include the effect of fuel cost, maintenance cost and ownership tax. Diesel technology effect captures the dummy factors that are not modeled explicitly.

Sensitivity of perceived diesel utility to vehicle attributes measures how consumers' perception of diesel changes with changes in price (diesel price and operation cost), performance (diesel performance and fuel efficiency), convenience (fuel availability), choice (diesel portfolio), and other factors (diesel technology effect). Though individual sensitivity to these factors may vary due to numerical reasons, the average sensitivity of a group of people is often prejudiced by common views. Because of the close cultural and economic connections between the six European countries, the AVMT consideration model assumes the sensitivities to price, performance, convenience and choice are the same across countries. However, the diesel technology effect, a parameter measuring dummy factors that impact consumers' choice, is country specific so that the difference between countries, if any, is still captured in the model. Table 2 includes the descriptions and values set for each vehicle attribute using Italy as an example.
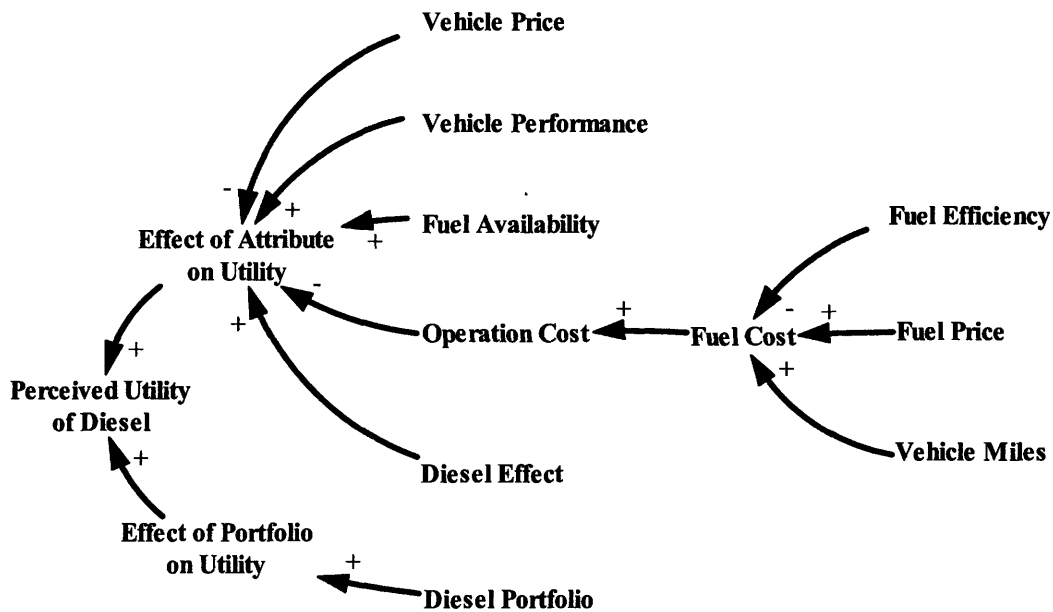
Vehicle Price

Vehicle Performance

Fuel Availability

Fuel Efficiency

Effect of Attribute on Utility

Operation Cost

Fuel Cost

Fuel Price

Perceived Utility of Diesel

Diesel Effect

Vehicle Miles

Effect of Portfolio on Utility

Diesel Portfolio

Figure 13 Perceived Utility of Diesel Platform conditioned by Vehicle Attributes

Table 2 Vehicle Attributes — Italy

	1980 Petrol	1980 Diesel	2005 Petrol	2005 Diesel	Unit	Description	Level
Vehicle Price	18650	17150	18650	17150	$/vehicle	Assumed fixd on typical 2002 model (Peugeot 307) including tax	Country
Vehicle Performance	1	0.786	1	1	dimensionless	Assumed that diesel closes gap with petrol linearly; diesel 0.7 in 1970	Uniform
Fuel Efficiency	10	12	10	12	kilometers/liter	Assumed diesel is 20% more efficient than petrol	Uniform
Fuel Price	1.12	0.522	1.35	1.12	$/liter	Data available	Country
Vehicle Portfolio	300	59	710	262	models	Data available	Country
Fuel Availability	1	1	1	1	dimensionless	Partly captured in diesel technology effect	Uniform
Other Operation Cost	-	-	-	-	$/year	Partly captured in diesel technology effet	Uniform
Taxes	-	-	-	-	-	Included in vehicle and fuel prices	Country

As demonstrated, the effective contact rate with diesel drivers C_{pdd}, effective contact rate with petrol drivers C_{pdp} as well as marketing effectiveness α_d determines a consumer's purchase decision via the magnitude of their social exposure to diesel. The larger the values of these parameters, the more likely that the consumers will include diesel in their consideration set and

contribute to the diesel market share by switching from petrol to diesel. Since the three parameters play a key role in consumer considerations, their influence on simulated model outcome deserves a detailed calibration. The three variables are referred to as social exposure parameters in the following context.

Calibration of Social Exposure Parameters

The purpose of this calibration is to search for the optimum value for three social exposure parameters - marketing effectiveness α , effective contact rate with diesel drivers C_d, effective contact rate with petrol drivers C_p by fitting model output to actual data (the subscript is simplified for clarity). The three social exposure parameters are estimated by minimizing the sum of squared errors between actual diesel sales data s_t and sales diesel model output $\hat{s}_t$.

$$\underset{\alpha, c_d, c_p, F_0, f}{Min} \sum_{t=1}^{36} (s_t - \hat{s}_t)^2 \tag{7}$$

subject to

$$0 \leq \alpha \leq 0.1$$

$$0 \leq c_d \leq 1$$

$$0 \leq c_p \leq 1$$

$$0 \leq F_o \leq 0.2$$

$$-5 \leq f \leq 5$$

The vehicle markets of the six European countries are highly connected in terms of the similarity of automotive models produced and the mobility of drivers. Therefore, effective contact rates with diesel drivers C_d and petrol drivers C_p are assumed to be same across countries. Marketing effectiveness α is bound by country borders because of the local nature of marketing channels and language barriers. The calibration also includes the initial familiarity F_0 to reflect the starting familiarity level in the simulation. In addition, the sensitivity of utility to diesel technology f is calibrated together for the purpose of exploring dummy factors that are

contributed to the diesel utility.

Four calibrations are designed to estimate values of these four parameters using AVMT consideration model.

Calibration One: Base Scenario (BASE)

    (1) Use initial diesel installed base at 1970 derived in Chapter 3 and total vehicle installed base data as model inputs.

    (2) Match the model output, sales diesel model, with diesel new registrations from 1970 to 2005, a combination of synthetic and real data derived in Chapter 3.

    (3) Search for best fit values for four social exposure parameters - effective contact rate drivers, effective contact rate non-drivers, marketing effective and initial familiarity with diesel, and one vehicle attribute parameter - sensitivity of utility to diesel effect. As discussed earlier, marketing effective and initial familiarity are country specific parameters while the others are the same across the six countries. Use the Vensim optimization function to conduct the automated search.

Calibration Two: Base Scenario with Country Specific Diesel Effect (BASE DEC)

    (1) Same as (1) of Calibration One.

    (2) Same as (2) of Calibration One.

    (3) Estimate same parameters as (3) of Calibration One. Except for marketing effectiveness and initial familiarity, the sensitivity of utility to diesel effect is also country specific.

Calibration Three: New Registrations (NVR)

    (1) Use initial diesel installed base data derived from Chapter 3, and total new registrations data as model inputs.

    (2) Same as (2) of Calibration One.

    (3) Same as (3) of Calibration One.

Calibration Four: New Registrations with Country Specific Diesel Effect (NVR DEC)

    (1) Same as (1) of Calibration Three.

(2) Same as (2) of Calibration One.

(3) Estimate same parameters as (3) of Calibration One. The sensitivity of utility to diesel effect is country specific as well as marketing effectiveness and initial familiarity.

Further, in order to compare AVMT consideration model to the Bass diffusion model in terms of diffusion behavior and parameter value, one scenario is designed to calibrate Bass model to diesel data. Table 3 summaries the four AVMT calibrations as well as the Bass model calibration.

Calibration Five: BASS

(1) Same as (1) of Calibration One.

(2) Fit sales diesel in the Bass model with diesel new registrations data from 1970 to 2004.

(3) Estimate effective contact rate with drivers and marketing effectiveness.

Table 3 Calibration Scenarios

|	BASS	Base	Base DEC	Base NVR	Base DEC NVR
Scalar Parameter	Contact Rate Drivers	Contact Rate Drivers Contact Rate Non-Drivers Diesel Effect	Contact Rate Drivers Contact Rate Non-Drivers	Contact Rate Drivers Contact Rate Non-Drivers Diesel Effect	Contact Rate Drivers Contact Rate Non-Drivers
Country Specific Parameter	Marketing Effectiveness	Marketing Effectiveness Initial Familiarity with Diesel	Marketing Effectiveness Initial Familiarity with Diesel Diesel Effect	Marketing Effectiveness Initial Familiarity with Diesel	Marketing Effectiveness Initial Familiarity with Diesel Diesel Effect
Data Input	Total Installed Base Initial Diesel Installed Base	Total Installed Base Initial Diesel Installed Base	Total Installed Base Initial Diesel Installed Base	Total New Registrations Initial Diesel Installed Base	Total New Registrations Initial Diesel Installed Base
Calibration Data **Model Output**	Diesel New Registrations Sales Diesel Bass	Diesel New Registrations Sales Diesel Model	Diesel New Registrations Sales Diesel Model	Diesel New Registrations Sales Diesel Model	Diesel New Registrations Sales Diesel Model

Behavior Reproduction Tests

As commonly used tools for model testing, behavior reproduction tests can assess a model's ability to reproduce certain behaviors observed in reality. Based upon qualitative observations and point-to point statistics measures, behavior reproduction tests compare the AVMT consideration model to the BASS model in terms of their ability to reproduce diesel diffusion patterns, as well as compare the replicating abilities of the four calibrations of the AFV consideration model. Figure 14 shows the best fit simulation results of each calibration scenarios.

Figure 14 indicates that BASS model lags behind the diesel sales trend after the first twenty years of simulation for France, Germany and Italy, and thirty years for the UK. In particular, the

BASS model performs poorly on data for Sweden, whose diesel diffusion has low penetration and sharp inflections. The BASS model also consistently over-predicts the diesel share for all of the six European countries.

As a comparison, the Base calibration of the AVMT consideration model is capable of replicating the diesel sales trend for all the six countries. Though the amplitude of the simulated sales is less fluctuating than actual data in countries like Italy and UK, this mismatch is not substantial considering the fact that the model smoothes out noises in a short time frame. The AVMT consideration model generally performs much better than BASS model in reproducing various diesel diffusion patterns, not matter whether is a successful case (France, Germany, Italy, Spain and UK), or a failed case (Sweden).

Aside from intuitive observations, this study uses common statistical metrics– R2, MAE, MAPE, MSE/RMSE and Theil's Inequality Statistics – to interpret the model fit. R2, the most commonly used measure of model fit, is the square of the correlation coefficient between model and data series. R2 is equal to one when the model exactly replicates the actual data series while zero means the model output is constant. MAE, the mean absolute error, MAPE, the mean absolute percent error, MAE/Mean, mean absolute error as a percent of the mean, and MSE/RMSE (root MSE) all measure the average error between the model output and actual data. Theil (Theil 1966) introduces "inequality statistics" to measure three components of MSE: bias (Um), unequal variation (Us) and unequal covariation (Uc). A Um reveals that a model has a systematic error in matching average output to data series. A Us indicates that the model output has a different trend to data trend. A Uc represents the error concentrating on unsystematic point-to-point difference. The sum of Um, Us and Uc is equal to one. Sterman (2000) interprets the Theil statistics for different situations.
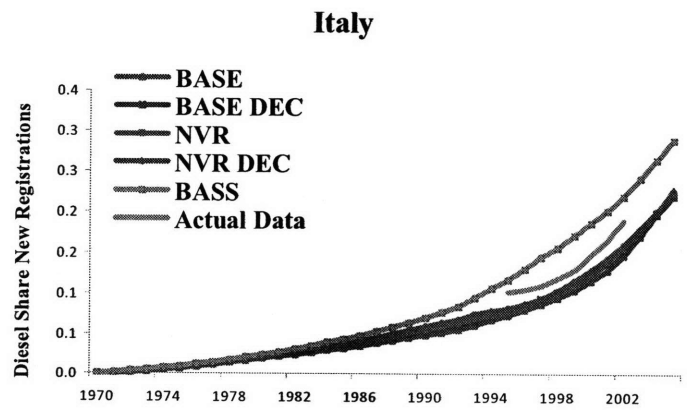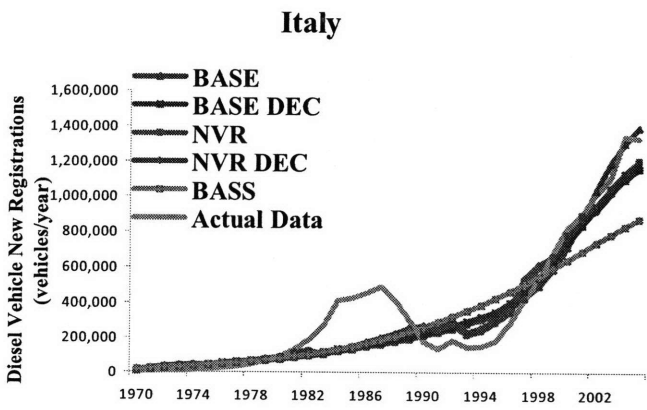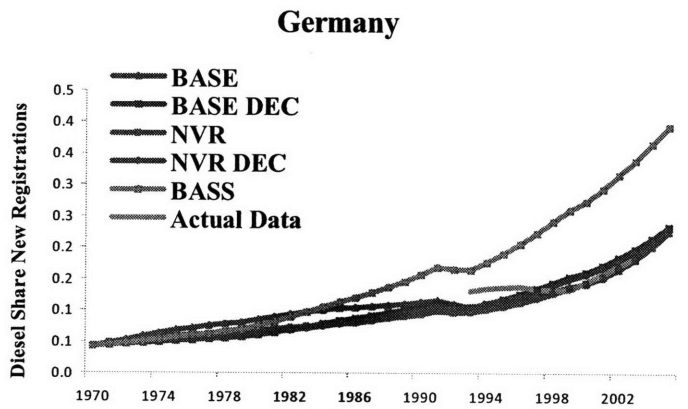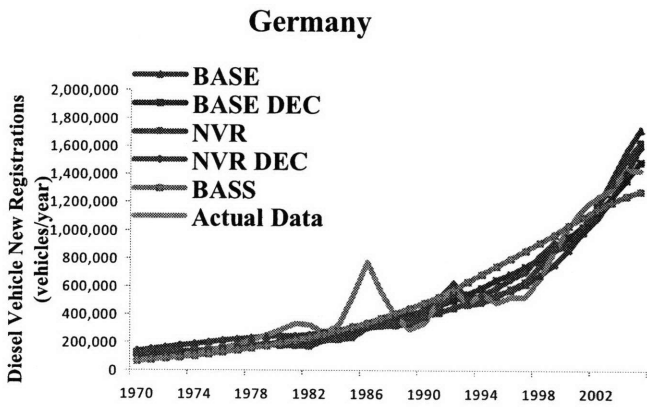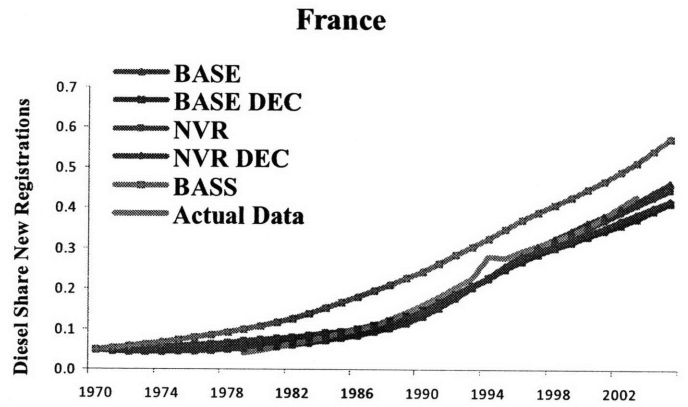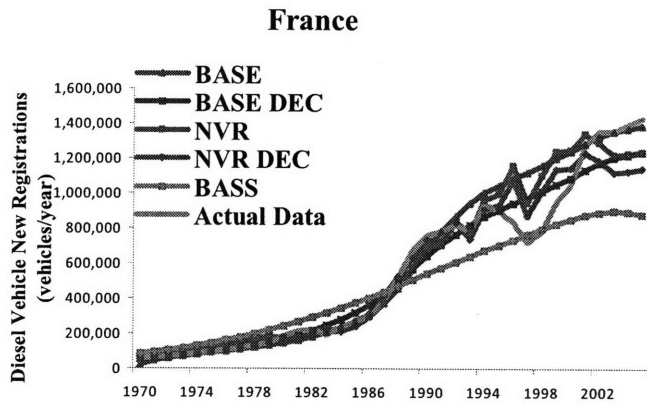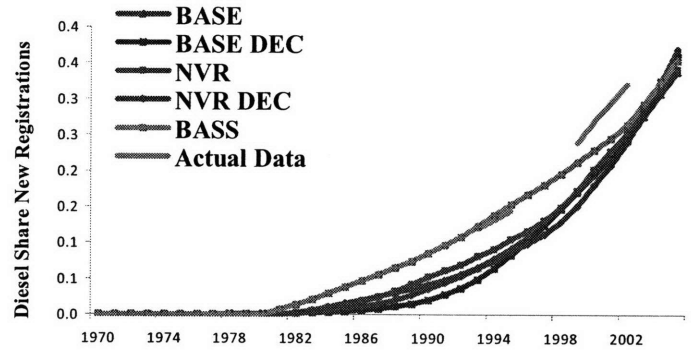
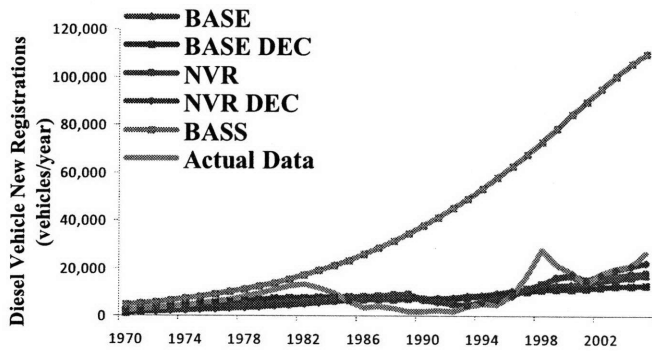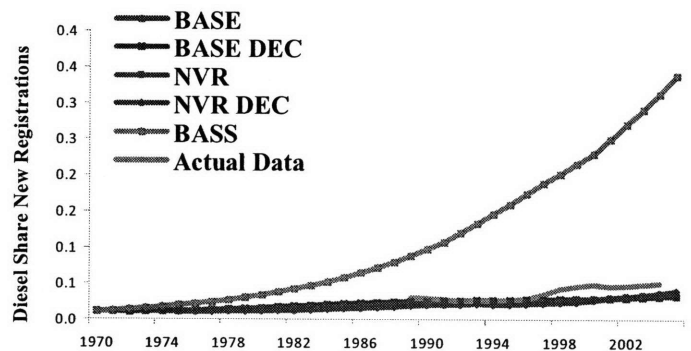**Figure 14 Diesel Diffusion Pattern Reproduction — France, Germany and Italy**
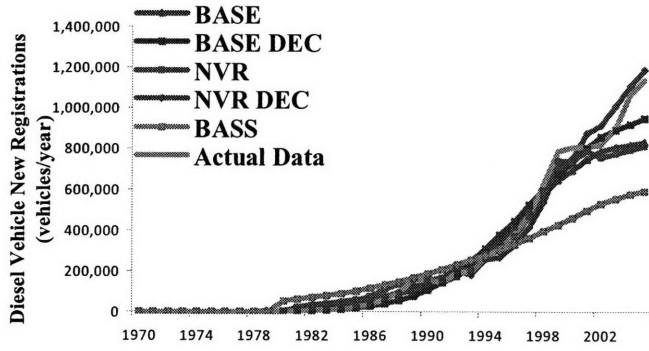
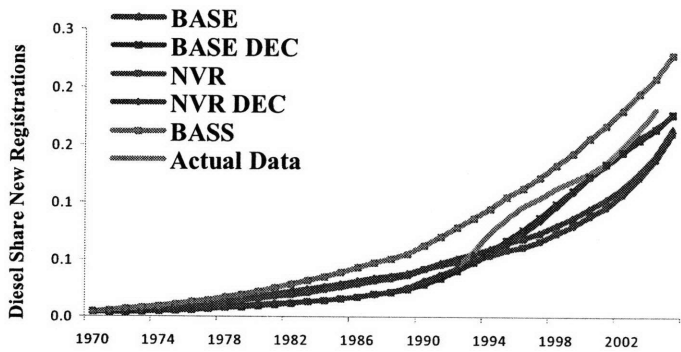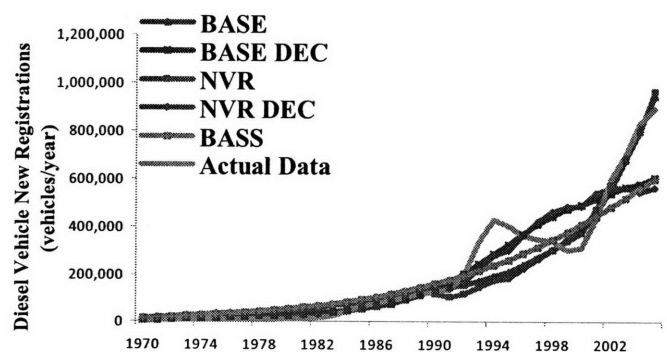**Figure 15 Diesel Diffusion Pattern Reproduction — Spain, Sweden and UK**

Table 4 lists the results of statistics measures for BASS model, and the four calibrations of AVMT consideration model. The BASS model presents poorly on all the indicators of average error – MSE, RMSE and MAE/MEAN. The big Us (0.198) and Uc (0.802) indicates that the errors of BASS model concentrate on unequal variations between model output and data, and on unequal covariance due to mismatch at point-by-point fit, e.g. France. On the contrary, all the four calibration scenarios of the AVMT consideration model have much smaller values of MSE, RMSE and MAE/MEAN and larger values of R2, all of which suggest a better fit to the empirical data. All the four scenarios have zero or close to zero Um value, and small Us value. With a close to zero Um and a small Us, the majority of MSE error is on unequal covariance, which is not a systematic error.

A high Us value may suggests that the errors may be due to problems in model structure (Sterman 2000). The high value of unequal variation indicates that the BASS model has structure difficulties in modeling diesel diffusion. This implies the importance of social exposure parameters and familiarity in capturing key behavior patterns of diesel diffusion.

Among the four calibration scenarios of AVMT familiarity model, country specific scenarios perform much better than corresponding non-country specific ones (Base vs. Base DEC, and NVR vs. NVR DEC) through higher R2 values and lower MSE, RMSE and MAE/MEAN values. The better performance of country specific models suggests that more country specific data, such as vehicle attributes, can potentially improve statistic results.

Table 4 Behavior Reproduction Tests

	BASS	BASE	BASE DEC	NVR	NVR DEC	Units
Basic Report						
Scalar Parameters	1	3	2	3	2	-
Country Specific Parameters	1	2	3	2	3	-
Number of Countries	6	6	6	6	6	-
Number of Fitted Coefficients	7	15	20	15	20	-
Number of Simulations	>1,000,000	>1,000,000	>1,000,000	>1,000,000	>1,000,000	-
Mean Absolute Error Fraction MAE/MEAN	0.319	0.188	0.189	0.189	0.192	dimensionless
Mean Square Error (MSE)	2.18E+10	9.07E+09	8.46E+09	9.64E+09	8.51E+09	(vehicles/year)^2
Root Mean Square Error (RMSE)	1.48E+05	9.52E+04	9.20E+04	9.82E+04	9.23E+04	vehicles/year
Adjusted R2	0.848	0.934	0.937	0.930	0.937	dimensionless
Theil Statistics						
Um - bias	0.000	0.000	0.002	0.001	0.000	dimensionless
Us - unequal variance	0.198	0.042	0.029	0.012	0.040	dimensionless
Uc - unequal covariance	0.802	0.957	0.973	0.987	0.960	dimensionless

Calibration Results: Social Exposure Parameters

Multiple calibrations estimate the value of social exposure parameters (Figure 16). The dot line represents the parameter values used by Struben (2006) which are from the literatures or assumptions. The calibration results suggest that the optimized value of marketing effectiveness is in line with values that are used by Struben. The effective contact rate with non-drivers is higher than the initial assumption of 0.15 while effective contact rate with diesel drivers tend to be smaller than initial value of 0.25. Initial familiarity matches with market situations in 1970 when diesel had already presented in German market while Spain and UK just started to learn about diesel platform. The calibrated social exposure parameter values are generally in line with original assumptions and further increase their numerical accuracy.
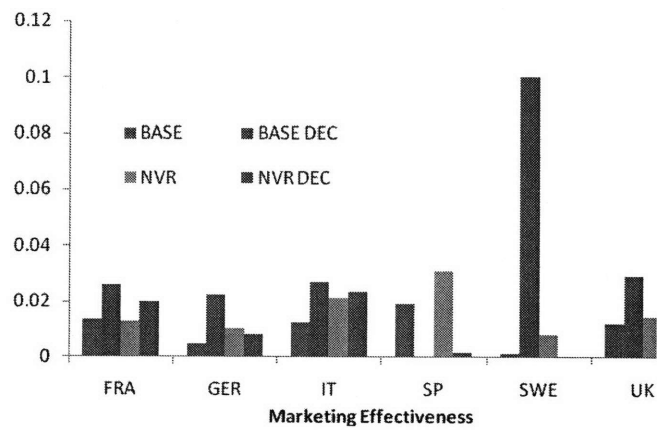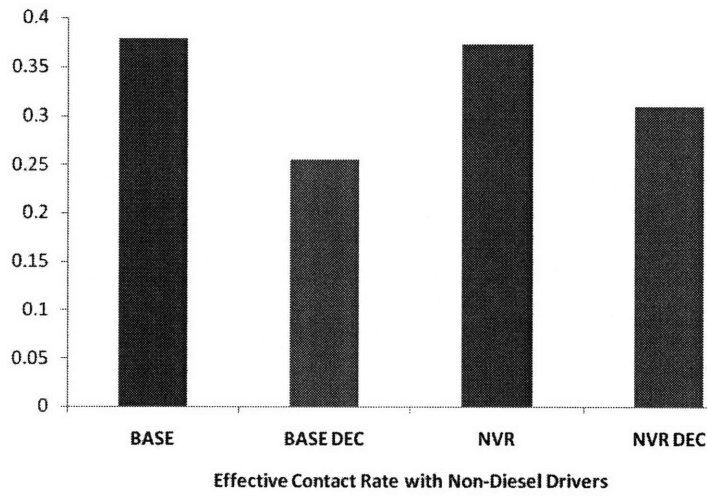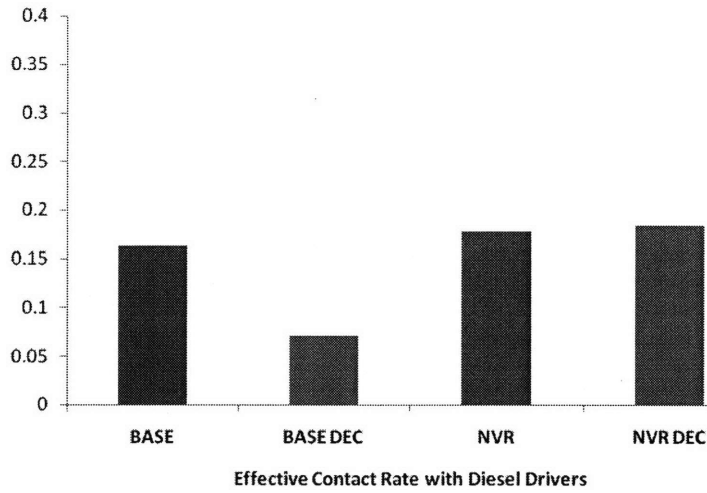
Effective Contact Rate with Diesel Drivers

Effective Contact Rate with Non-Diesel Drivers

Figure 16 Social Exposure Parameter Calibration Results

Chapter 5 Confidence Intervals

After calibrations of the AVMT consideration model confine the values of social exposure parameters, the question that follows relates to the level of confidence in these best-fit parameter estimates. A common statistical indicator of the reliability in estimate is Confidence Intervals (CI). Instead of a single value estimate, CI is an interval of possible estimates. The more likely it is for the interval to contain one parameter value, the wider the interval is. A result with a small CI is more reliable than a result with a large CI. Without the CI test, it is difficult to determine whether social exposure parameter values can be significant different from the estimated values, and therefore also difficult to assess the reliability of the estimate. In order to test the reliability of the calibrated social exposure parameter values, it is necessary to compute confidence intervals for each of them.

Bootstrapping Theory

The CI test normally does not attract a lot of attention from SD practitioners for two reasons. First, dynamic behavioral models with nonlinear feedback systems often violate the assumptions of the asymptotic methods that are commonly used for calculating CI, such as the likelihood ratio. For instance, the observations of a sample are not independent and identically distributed (i.i.d.) and the numbers of observations are small. The asymptotic methods can yield a narrower CI for SD models than it should (Dogan 2007). Furthermore, popular system dynamic software, such as Powersim, does not include CI estimate packages. Though Vensim has CI estimate capability, its algorithms are based upon asymptotic assumptions. The unavailability of an automated CI package for SD models makes CI tests difficult to implement as a regular model test procedure though its importance and necessity as a step of model calibration is unquestionable (Oliva 2003).

As for the present study, the sample of diesel sales data violates the i.i.d. assumption. By nature, the diesel sales at time t are influenced by the sales in the previous years. For instance, the peak of vehicle sales at one year can result from a sales pause in previous years. In addition, the sample size is small – one time series of 36 annual data points from 1970 to 2005 for each of the six European countries. Non-country specific parameters are estimated from 180 data points (36

42

annual data * six countries) while country specific parameters are calculated with 36 data points. Therefore, the likelihood method is not appropriate for a CI test of this study.

Dogan (2007) suggests that bootstrapping is a better CI estimate method for SD models because it does not impose restrictive assumptions as the likelihood ratio method does. Bootstrapping is randomly drawing a number of samples from a population and calculating the value of the estimator of interest for each sample to estimate the population distribution. With the bootstrap method, the sampling distribution of the chosen estimator, social exposure parameters in this case, is computed relying on the fact that the sample's distribution is a good estimate of the population distribution. Bootstrapping can use either parametric or nonparametric approach (Hinkley 1988). If the datasets fit a known probability distribution well, new datasets can be generated according to the known probability distribution. If the probability distribution of the original datasets is unknown, it is suitable to use a nonparametric method which fabricates new datasets by randomly drawing from the original datasets with replacement. Confidence intervals of parameters can be estimated with bootstrapping samples.

This study applies the residual based bootstrapping method proposed by Dogan (2007) for SD models. The general procedures are as follows:
1) Fit the SD model to actual data to search for optimum parameter values;
2) Derive original error terms by subtracting model output from actual data given the optimized parameter values;
3) Validate the maintained hypotheses of asymptotic methods. If hypotheses are violated and autocorrelation exits, determine the autoregressive process and recover the underlying white noise by removing the autocorrelation;
4) Resample new error terms by randomly drawing from the original error terms. If autocorrelation exits in error terms, resample the underlying white noise instead, and derive new error terms by adding autocorrelation back to resampled white noise. Choose either a parametric or nonparametric resampling method depending on the distribution of the resampled sample.
5) Add the new error terms to the model output to obtain fabricated data sets;
6) Search for the best-fit parameter value for the fabricated data sets;

43

7) Repeat 3), 4) and 5) to generate a large sample of new datasets, calculate parameter values for each datasets and obtain the statistics of interest.

Bootstrapping Process

The residual based bootstrapping process starts with a sample of error terms. Error terms are the discrepancy of actual data and simulated model output (equation). The calibration in Chapter 4 has provided a handful of simulated diesel sales from four calibration scenarios (BASE, BASE DEC, NVR, NVR DEC). Since the analysis in Chapter 4 indicates that the country specific calibration BASE DEC is statistically more capable of fitting empirical data, the bootstrapping process utilizes the simulated model output from the Base DEC calibration. Figure 17 to Figure 22 show the actual sales data s_t, simulated output of Base DEC $\hat{s}$, and the related error terms e_t, $e_t = s_t - \hat{s}_t$ for each data series. The error terms do not have obvious trends or seasonable cycles.
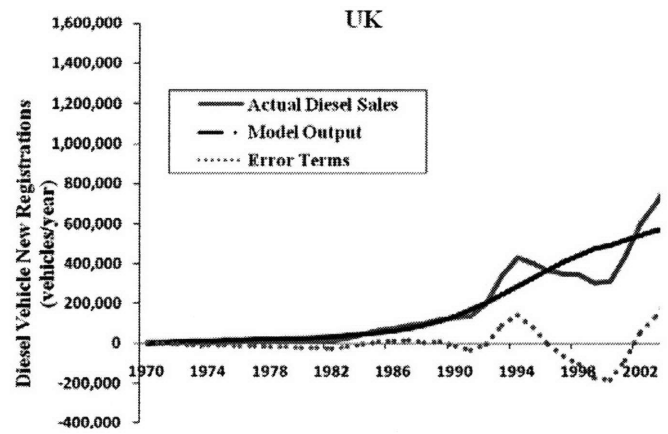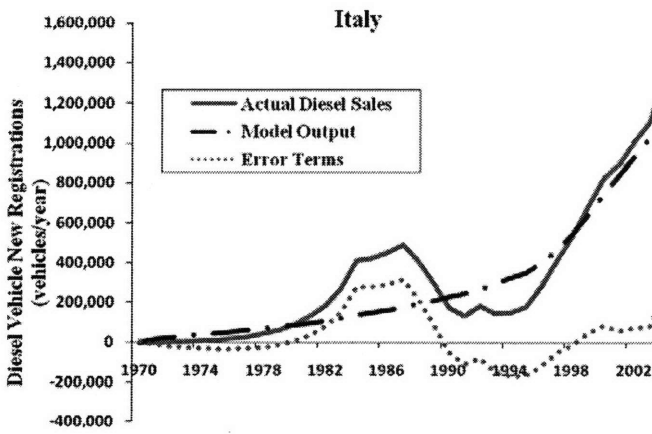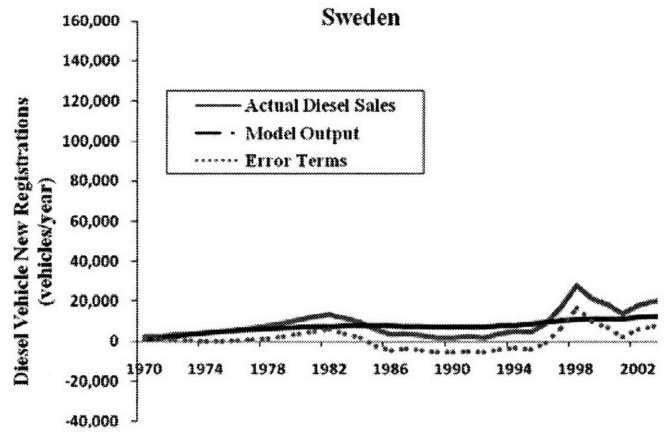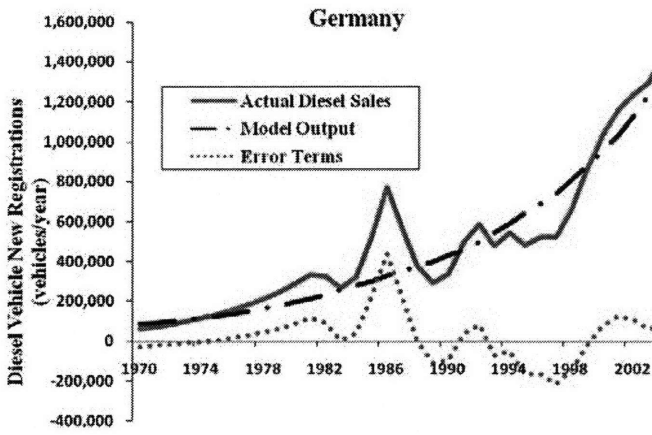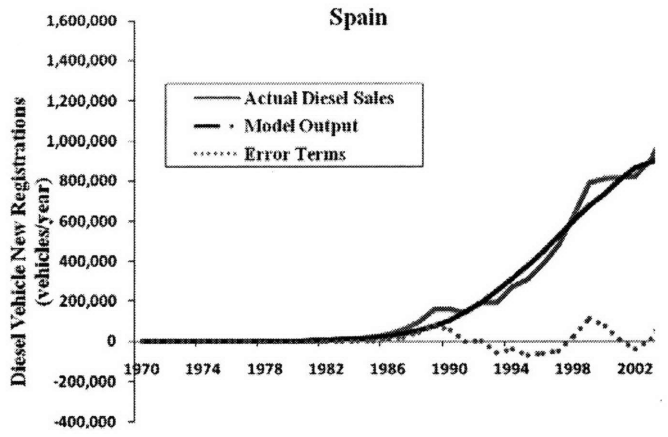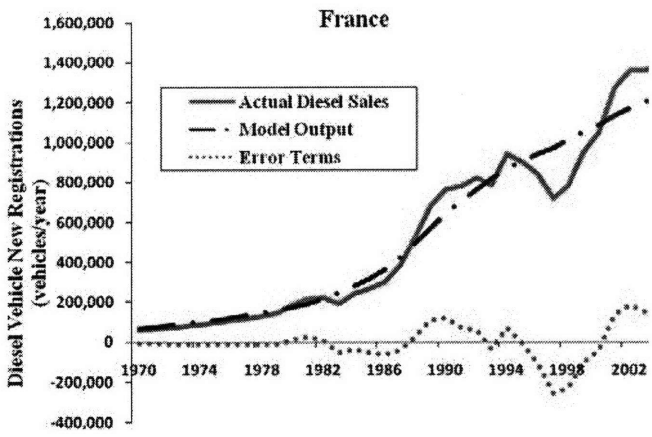
Figure 17 Error Terms

45

The next step is to validate the assumptions of the asymmetric methods – the error terms are independent and identically distributed (i.i.d). Since autocorrelation is a typical violation of i.i.d assumption, the study uses the autocorrelation function to determine the relationship between error terms.

Box and Jenkins (1976) suggest that the estimate of the kth lag autocorrelation $\rho(k)$ for the residuals, $k = \{0, 1, \ldots, N\text{-}1\}$, is

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} \tag{8}$$

where $\rho(0) = 1$. The autocovariance $\gamma(k)$ at lag k is given by

$$\gamma(k) = Cov(k) = \frac{1}{N} \sum_{i=1}^{N-k} (e_i - \bar{e})(e_{i+k} - \bar{e}) \tag{9}$$

In order to determine whether the autocorrelation values are significantly different from zero, define the autocorrelation function variance estimator (Barlas 1989) by

$$Var(\rho(k)) = \frac{1}{N(N+2)} \sum_{i=1}^{N-k-1} (N-i)(\rho(k-i) + \rho(k+i) - 2\rho(k)r(i))^2 \tag{10}$$

where $r(k)$ is autocorrelation function value for lag k, N is number of data points. The null and alternative hypotheses are

$$H_0: \quad \rho(i) = 0, \qquad \forall i \in \{1,2,\ldots,M\}$$

where M is the maximum lag and $M < (N\text{-}1)$

$$H_1: \quad \exists \, \rho(i) \neq 0, \qquad \forall i \in \{1,2,\ldots,M\}$$

The test statistic for the individual autocorrelation function value is

$$t(k) = \frac{\rho(k)}{\sqrt{Var(\rho(k))}} \tag{11}$$

This equation is an individual test of the hypothesis that autocorrelation is not significantly different from zero at lag k. A reasonable significant level (α) for an individual test is 0.05. While testing the hypothesis at all lags, the significant level should use a smaller number (Barlas 1989) then the significant level of an individual test. This study uses 0.01 for the multiple tests. Assuming a normal distribution of the autocorrelation function values, the null hypothesis is

46

rejected if $|t(k)| > 2.58$ given $\alpha = 0.01$. The test statistic would be calculated for $k=0,1,\ldots,K$ where K is smaller than N/4 (Box and Jenkins 1976).

The autocorrelation function is calculated for lag $k=1$ to 8 for each of the six countries. Results are shown in Table 5. Significant autocorrelation of Germany (GER) and Spain (SP) occurs at lag 1, France (FRA), Sweden (SWE) and UK (UK) at lag 1 and 2, and Italy (IT) at lag 1, 2 and 3. The autocorrelation at multiple lags suggests that the residuals violate the independence assumption of the asymmetric method. In order to conduct residual based bootstrapping, the autocorrelation needs to be identified so that the underlying white noise can be recovered by removing the autocorrelation (Dogan 2007).

The stochastic process to model such autocorrelation is called autoregressive (AR) model (Box and Jenkins 1976). An AR process of p order, AR(p), is defined by

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \varepsilon_t \tag{12}$$

X_t is a weighted sum of its past values, ϕ_t is coefficient parameter, and the ε_t is a white noise process with zero mean and variance σ^2. The AR model defines the correlation between the series $\{X_t, X_{t-1}, \ldots X_{t-p}\}$. However, the autocorrelation can exit not only in X_t, but also in the white noise ε_t. A more general model to capture the autocorrelation is the autoregressive moving average (ARMA) model (Box and Jenkins 1976). The ARMA(p,q) model is a linear model mixing moving average effects and the autoregressive process

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \tag{13}$$

where $\varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}$ is the moving average process of q order. θ_t is the coefficient parameter. ε_t is random noise with zero mean and variance σ^2. The ARMA model extends the boundary of AR model by including autocorrelation terms about ε_t. So it can fit with data with autoregression both in X_t and ε_t.

The test statistic results in Table 5 imply that all the six data series have autocorrelation at the first lags. It is assumed that the data series can be fitted to an AR model with p order: FRA AR(2), GRE AR(1), IT AR(3), SP AR(1), SWE AR(2) and UK AR(2). With known error terms and AR orders, the coefficient ϕ_t can be estimated by minimizing the sum of squares of residuals

– actual error terms less the fitted values of AR model. (Janacek and Swift 1993). The calculations use automatic computational modules in the statistical software SAS.

It is necessary to do a further diagnostic checking to ascertain whether the model is appropriate for the data. For the purpose of determining autocorrelation in error terms, the diagnostic checking is to test for serial correlation in the residuals. Compute the test statistic for ε_t. The results (Table 5) show that GER and SP series still have a first order autocorrelation in residuals which violate the assumptions that AR model make about ε_t. This finding suggests that a pure AR process is not sufficient to characterize the autocorrelation in GER and SP series. It is necessary to use a more complex stochastic process to represent the autocorrelation in error terms as well as in residuals for GER and SP series. The alternative choice is the ARMA model.

After fitting ARMA(1,1) to GER and SP series, calculate parameter values in SAS using the least squares method, and then derive underlying residuals w_t – actual error terms less the fitted value of ARMA model. Test statistic results indicate that the derived w_t are not autocorrelated. Therefore, the overall autocorrelation of GER and SP is identified along with other four data series.

The analysis seeks an appropriate technique to resample a large size of white noise signals w_t. The choice of bootstrapping techniques depends on the probability distribution of the recovered white noise w_t. If it is normally distributed, the bootstrapping can generate new white noise signals according to the known parametric distribution. If not, the nonparametric approach that resample w_t with replacement is applicable. Visual observations (Figure 18) strongly reject the hypothesis of the normality of w_t. So the nonparametric approach is chosen to resample i new white noise series $w_{i,t}$ with replacement from original w_t, i=1000. In this approach, a resampling is implemented in Matlab.

Table 5 Test Statistic Results

Test Statistic t(k) - Error Terms

Lag	1	2	3	4	5	6	7	8
FRA	4.81	5.27	0.30	-1.36	-1.89	-2.19	-2.54	-2.41
GER	6.82	1.28	-0.22	-0.28	0.33	0.52	-0.03	-0.40
IT	5.36	6.05	5.91	1.69	-0.52	-1.56	-2.20	-2.74
SP	6.88	0.04	-0.81	-0.46	0.48	0.22	-0.76	-1.12
SWE	8.32	3.72	1.54	1.20	1.29	0.58	-0.01	-0.98
UK	4.58	4.87	-0.73	-1.70	-2.30	-2.70	-1.14	0.03

Test Statistic t(k) - Residuals with AR model

Lag	1	2	3	4	5	6	7	8
FRA	-0.42	0.76	-0.53	-0.46	0.13	-0.85	0.18	-1.70
GER	5.29	-0.28	-1.27	-0.78	0.51	0.82	-0.06	0.06
IT	-0.11	-0.48	-0.48	0.52	0.73	-0.90	-0.38	0.50
SP	5.74	-0.47	-1.35	-0.90	0.78	0.96	0.04	-0.59
SWE	0.13	0.27	-1.02	0.42	0.61	-0.60	1.07	0.10
UK	-0.40	-1.26	-0.22	1.29	-0.54	-1.31	-0.67	0.93

Test Statistic t(k) - Residuals with ARMA model (GER and SP)

Lag	1	2	3	4	5	6	7	8
FRA	-0.42	0.76	-0.53	-0.46	0.13	-0.85	0.18	-1.70
GER	0.40	0.13	-1.27	-0.35	0.25	0.82	-0.65	0.68
IT	-0.11	-0.48	-0.48	0.52	0.73	-0.90	-0.38	0.50
SP	0.54	0.29	-1.55	-0.89	0.99	0.37	0.49	-0.75
SWE	0.13	0.27	-1.02	0.42	0.61	-0.60	1.07	0.10
UK	-0.40	-1.26	-0.22	1.29	-0.54	-1.31	-0.67	0.93

*Highlights indicate autocorrelation with a significane level of 0.01
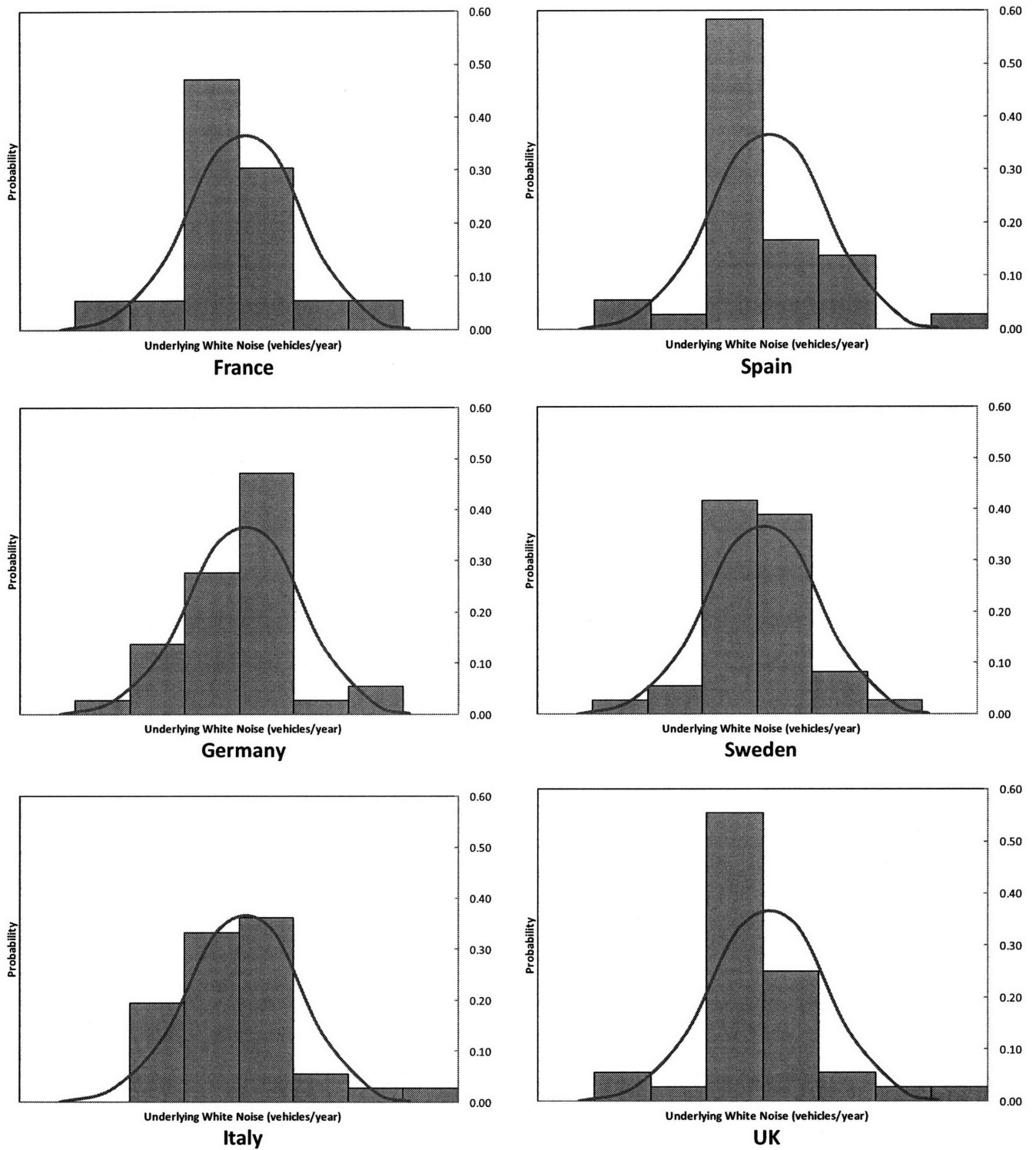
Figure 18 Histogram of the Underlying White Noise

An ith series of fabricated error terms, $e_{i,t}$, consists of three components: the AR process of $e_{i,t}$, the MA process of $\varepsilon_{i,t}$, and the resampled white noise $w_{i,t}$

$$e_{i,t} = \phi_1 e_{i,t-1} + ... + \phi_p e_{i,t-p} + \varepsilon_{i,t} + \theta_1 \varepsilon_{i,t-1} + ... + \theta_q \varepsilon_{i,t-q} + w_{i,t}$$

The equation for the data series with a AR process, such as FRA, IT, SWE and UK, can be simplified as

$$e_{i,t} = \phi_1 e_{i,t-1} + ... + \phi_p e_{i,t-p} + w_{i,t}$$

Initial value $e_{i,g}$ when 1<g<p+1 is equal to $w_{i,g}$. The fabrication of new error terms $e_{i,t}$ is straightforward for the four AR processes. Given the estimated parameter values, it involves simply adding the identified autocorrelation of $e_{i,t}$ to resampled white noise.

The equation suitable for an ARMA(1.1) process for GER and SP series is

$$e_{i,t} = \phi_1 e_{i,t-1} + \varepsilon_{i,t} + \theta_1 \varepsilon_{i,t-1} + w_{i,t}$$

Where ε_t is random noise stream with zero mean and variance σ^2. σ^2 of an ARMA (1,1) is calculated by (Box and Jenkins 1974)

$$\sigma^2 = \frac{(1 - \phi_1^2)}{1 + \theta_1^2 - 2\phi_1\theta_1} \gamma_0$$

γ_0 is autocovariance $\gamma(0)$. Assume ε_t is normally distributed, generate i^{th} datasets $\varepsilon_{i,t}$ with the calculated variance σ^2. The new error terms are fabricated by adding the autocorrelation of $e_{i,t}$ and $\varepsilon_{i,t}$ to the resampled white noise $w_{i,t}$. Sum the generated error terms and simulated diesel sales to yield an ensemble of 1000 bootstrapping data series of diesel sales, and estimate the social exposure parameters by fitting the AVMT consideration model to each of the 1000 fabricated sales data series.

It is important to test biases in the generated datasets. Dogan (2007) discusses that generated error terms – synthetic sales data less model output in 1000 simulation runs – tend to be smaller than the original error terms. The non-negativity constraint on the synthetic sales is a typical cause. If generated error terms are smaller than original ones, the original white noise should be

inflated. Dogan (2007) proposes the heuristic to use the standard deviation of error terms to check if an inflation is necessary. Table 6 shows the comparison of the mean of the standard deviation of generated error terms and the standard deviation of original error terms. Since the difference is evident, the analysis multiplies each white noise by a inflation factor, which is the standard deviation of original error terms divided by the mean of standard deviations of generated error terms. Repeat the bootstrapping resampling and derive the inflated error terms. However, the mean of the standard deviation of error terms after the first inflation is still quite small for GER and SP (GER 74% of original error terms, SP 80%). So this analysis inflates the white noise for the second time, and iterates the same procedure to compute the social exposure parameters for the 1000 simulations.
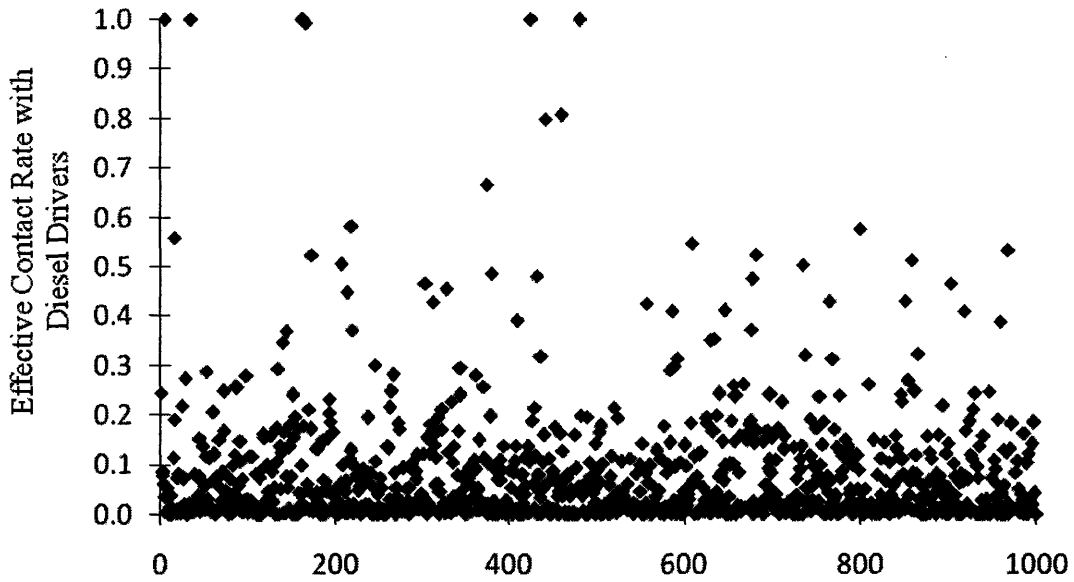
Table 6 Error Terms Inflation

	Original Error Terms STD	Generated Error Terms Mean STD	Inflated Error Terms Mean STD	Inflation Factor Fraction
FRA	97321	33714	98486	2.85
GER	119611	43803	111948	2.92
IT	129195	41875	130431	3.06
SP	52864	22344	50253	2.49
SWE	5390	2029	5194	2.76
UK	89542	47663	88059	1.91

1000 estimates for the social exposure parameters are available to test confidence intervals. This analysis uses the percentile method to compute CI (Dogan 2007). Confidence interval limits of the social exposure parameters are shown in Table. A 95% confidence interval include the initial estimated Effective Contact Rate with Non-Diesel drivers c_d (0.26). The 95% CI strongly rejects the hypothesis that c_d, is equal to 0, suggesting that the contact rate with non-diesel drivers, and thus the word of mouth effect from non-diesel drivers, is influential in increasing consumers' social exposure to the diesel platform. Furthermore, a 95% confidence interval include the estimated value of Effective Contact Rate with Non-Diesel Drivers c_p (0.07). However, neither a 95% CI nor a 90% CI can reject the hypothesis that c_p is 0. The relatively wide intervals of c_p may indicate that more data should be collected before a definite conclusion can be drawn about the parameter value. Either a 95% or a 90% confidence interval of Marketing Effectiveness, α,

covers the full search range for five countries with SP being an exception. This finding calls for a broader search range for α in the five countries. The estimated α for SP is in the range of (0, 0.03) with a 95% confidence level.

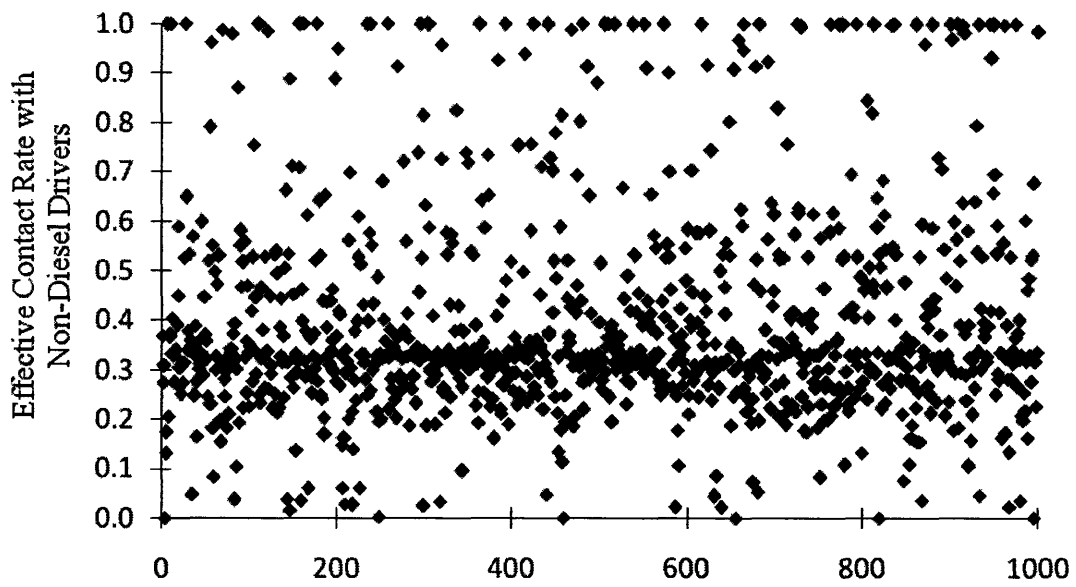**Table 7 Bootstrap Confidence Interval Results**

95% Confidence Intervals

		Lower Limit	Parameter	Upper Limit	Width of Interval
c_d		0.00	0.07	0.45	0.45
c_p		0.06	0.26	1.00	0.94
α	FRA	0.00	0.03	0.10	0.10
α	GER	0.00	0.02	0.10	0.10
α	IT	0.00	0.03	0.10	0.10
α	SP	0.00	0.00	0.03	0.03
α	SWE	0.00	0.10	0.10	0.10
α	UK	0.00	0.03	0.10	0.10

90% Confidence Intervals

		Lower Limit	Parameter	Upper Limit	Width of Interval
c_d		0.00	0.07	0.29	0.29
c_p		0.16	0.26	0.99	0.83
α	FRA	0.00	0.03	0.10	0.10
α	GER	0.00	0.02	0.10	0.10
α	IT	0.00	0.03	0.10	0.10
α	SP	0.00	0.00	0.03	0.03
α	SWE	0.00	0.10	0.10	0.10
α	UK	0.00	0.03	0.10	0.10

Empirical Cumulative Distribution Function

Effective Contact Rate with Diesel Drivers
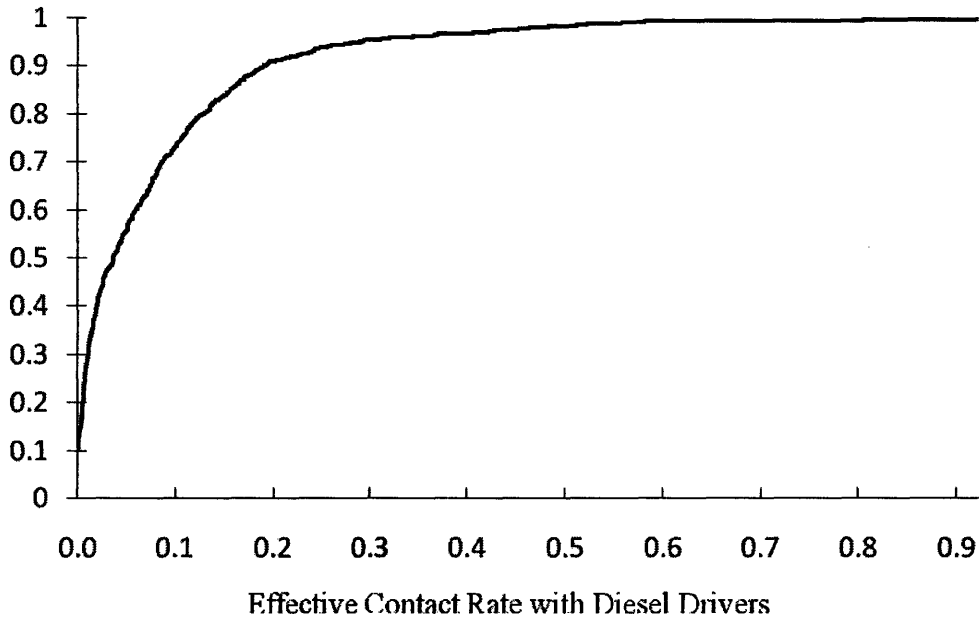
Figure 19 Bootstrapping Confidence Intervals - Effective Contact Rate with Diesel Drivers

Empirical Cumulative Distribution Function

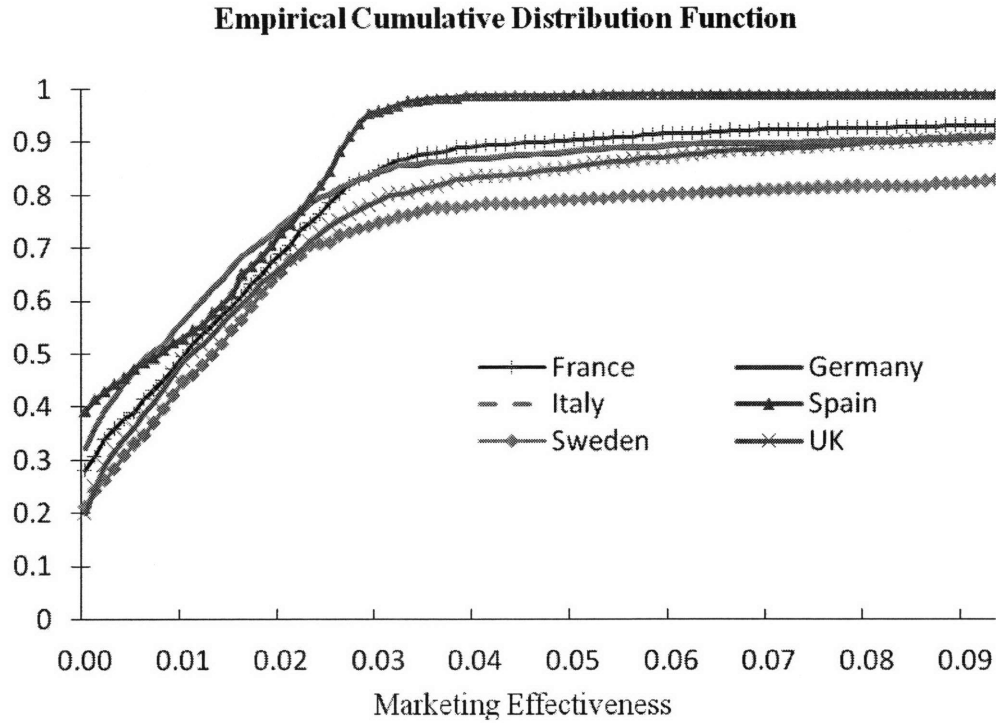Effective Contact Rate with Non-Diesel Drivers
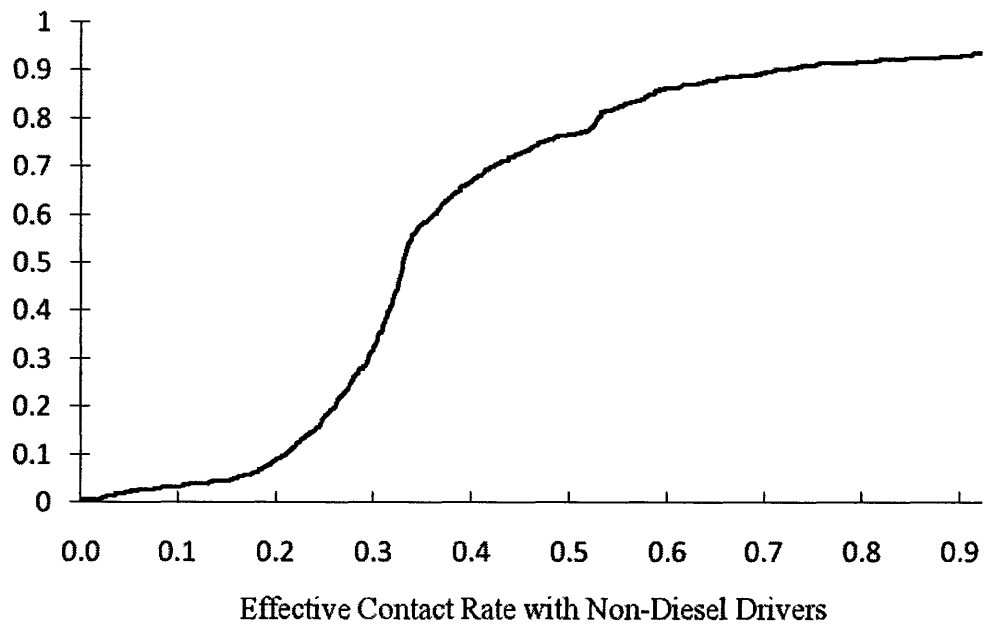
Figure 20 Bootstrapping Confidence Intervals - Effective Contact Rate with Non-Diesel Drivers

Empirical Cumulative Distribution Function

Figure 21 Bootstrapping Confidence Intervals – Marketing Effectiveness

Chapter 6 Conclusions

General Conclusion

This analysis calibrates the AVMT consideration model to the diesel diffusion in Europe. Using this simplified model with the consumer consideration structure, this analysis replicates the various diesel diffusion patterns observed in the six European countries, indicating the model's ability in capturing the dynamics brought by consumer behaviors in the AFV transition. Compared to the Bass diffusion model, the AVMT consideration model performs significantly better in simulating a wide range of diesel diffusion patterns, for either the sustainable or unsuccessful case. The simulations under multiple calibration scenarios confirm that the behavior reproduction of AVMT consideration model is repeatable and robust.

This study is the first attempt to calibrate a social exposure model to the automotive industry. The calibration of social exposure parameters – Effective Contact Rate from Diesel Drivers, Effective Contact Rate from Non-diesel Drivers, and Marketing Effectiveness – sheds light on the range of the parameter values for automotive products, and more generally, provides a parameter value reference for durable goods. Using the bootstrapping method, the test of confidence intervals specifies the reliability of the estimate of these parameters.

Data Analysis
This analysis develops practical approaches to solving data challenges that often occur in empirical studies. The challenges include the collection of historical data, the completeness of a data series, the organization of a data sample, and the interpretation of a data trend.

The search for empirical data qualified for a study need is often an arduous and time-consuming process. This study faces the difficulty of collecting complete data series of the historical diesel sales, installed base and vehicle attributes over a 30—year period from six countries. Though the diesel sales and vehicle in use are frequently included in a country's fiscal statistics, the availability of early data is poor. Even when the data is accessible, the collected data sample often has quality problems such as abnormal noises resulting from changing statistics standards over time. Furthermore, conceptual variables, like diesel performance, are hard to explain by a single indicator in the real world, and therefore hard to quantify and compare. Under such circumstances, appropriate techniques to improve the data quality are valuable. This study develops a data derivation model to generate synthetic datasets in order to complete a data series with missing data points. Following the least squares method, the idea of this structure is to calibrate the parameter which determines the simulated data trend by fitting the model to existing data. This technique takes advantage of automatic computation so that it can avoid the speculative and subjective assumptions made in a manual trend extrapolation effort.

Furthermore, the analysis organizes data for to serve different calibration purposes. For instance, the boostrapping process is conducted at country level. but the bootstrapping generated data are grouped all together to compute the confidence intervals of the two effective contact rates parameters, because they are assumed to be same across the six countries. A calibration with

multiple data series and a handful of parameters. The calibration with multiple parameters could mathematically lead to an implausible estimate of parameter values. An analyst needs to make proper assumptions to cut off margin factors and decrease the number of parameters estimated in one calibration while maintaining the calibration quality.

This study provides an example of conducting bootstrapping confidence intervals for data series with autocorrelations. The process entails identifying autocorrelation in the data sample, recovering the underlying white noise, resampling the white noise, inflating the white noise, fabricating new datasets, and producing the distribution of parameter values based on the generated datasets. As the core of the whole process, the autocorrelation of the six data series are identified by two stochastic models, AR and ARMA: the first one is straightforward; the latter one needs extra mathematical efforts. The diagnostic checking should take place at each step.

Policy Insights

Many vehicle attributes in the six European countries are almost identical: same vehicle models are available throughout these markets; vehicle performance is comparable; the difference in production cost is marginal; and refuel infrastructures are well constructed. Not like most of AFVs, the fuel efficiency of diesel is considerably higher than petrol. The equivalence of these attributes means that other factors contribute to the different diesel diffusion patterns in the six countries – social exposure through marketing programs, as well as fuel and vehicle prices. Social exposure is crucial in shaping consumers' consideration of a diesel platform. In the countries with a shift to diesel, a solid familiarity base usually exists before the diesel shares take off. Aggressive marketing policies raise consumers' familiarity with the diesel platform, and therefore increase the possibility that they will switch from a gasoline model.

Another insight gained from the comparison analysis is that taxation policies on vehicles and fuels are important for a diesel transition. Since many attributes of diesel vehicle are identical in the six European countries, the transition paths of the countries are particularly sensitive to policy incentives through taxation on vehicles and fuels. A policy favoring diesel and diesel model is a key reason that consumers adopt this platform.

These insights are applicable to other AFV platforms, such as hybrid, hydrogen and CNG. However, diesel is arguably the alternative platform with the least switching cost and the fewest hurdles in the transition among the wide variety of AFVs because diesel is a similar alternative to petrol in terms of many attributes. Therefore, the marginal factors in diesel diffusion may have considerable influence in the transition to other AFVs.

Future efforts can focus on improving the vehicle replacement structure of the AVMT consideration model. The analysis shows that the simulated installed base deviates from the historical data, which indicates room for improvement in the vehicle replacement structure. Particular attention should be given to the used car market structure which would change the speed of the AFV diffusion.

References

Bass, F. (1969) A New Product Growth Model for Consumer Durables, Marketing Science 15(5): 215-227.

Bass, F. M. (1980) The Relationship between Diffusion Rates, Experience Curves, and Demand Elasticities for Consumer Durable Technological Innovations, Journal of Business 53(3): S51-S67.

Bass, F. M., T. V. Krishnan, et al. (1994) Why the Bass Model Fits without Decision Variables, Marketing Science 13(3): 203-223.

Barlas Y.(1989) Multiple Tests for Validation of System Dynamics Type of Simulation Models, European Journal of Operational Research 42(1): 59-87

Barlas Y., (1996) Formal Aspects of Model Validity and Validation in System Dynamics Models, System Dynamics Review 12(3): 183-210

Barlas Y., (1990) An Autocorrelation Function Test for Output Validation, Simulation 55(1): 7-16

Ben-Akiva, M. and S. Lerman (1985) Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MA: MIT Press.

Box, G. and Jenkins, G. (1976) Time Series Analysis: Forecasting and Control. San Francisco, CA: Holden-Day.

Brownstone, D., D. Bunch, and K. Train. (2000) Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles, Transportation Research Part B—Methodological 34(5): 315-338.

Chen B. and Sperling D. (2004) Analysis of Auto Industry and Consumer Response to Regulations and Technological Change, and Customization of Consumer Response Models in Support of AB 1493 Rulemaking Case Study of Light-Duty Vehicles in Europe. Institute of Transportation Studies, University of California, Davis. Paper UCD-ITS-RR-04-14.

Davis, S. C. and S. W. Diegel (2006) Transportation Energy Data Book. Edition 25. Oak Ridge, TN: Center for Transportation Analysis, Oak Ridge National Laboratory (ORNL)

Davison A.C. and Hinkley D.V. (1997) Bootstrap methods and their application. Cambridge, England: Cambridge University Press.

Dogan G. (2007) Bootstrapping for Confidence Interval Estimation and Hypothesis Testing for Parameters of System Dynamics Models, System Dynamics Review 23(4): 415-436

Farrell, A., D. Keith and J. Corbett (2003) A Strategy for Introducing Hydrogen into Transportation, Energy Policy 31(13): 1357-1367.

Forrester, J.W. (1969) Urban Dynamics. Waltham, MA: Pegasus Communications.

Forrester, J.W. and P.M. Senge (1980) Test for Building Confidence in System Dynamics Models. TIMS Studies in the Management Science, 14. New York: North-Holland, 209-228.

Greenspan, A. and D. Cohen (1999) Motor Vehicle Stocks, Scrappage, and Sales, Review of Economics and Statistics 81(3): 369-383.

Hard, M. and A. Knie (2001) The Cultural Dimension of Technology Management: Lessons from The History of the Automobile, Technology Analysis & Strategic Management 13(1): 91-103.

Homer, J. (1983) Partial - Model Testing as a Validation Tool for System Dynamics. Proceedings of the 1983 International System Dynamics Conference. Chestnut Hill, MA: System Dynamics Society.

Homer, J. (1996) Why We Iterate: Scientific Modeling in Theory and Practice, System Dynamics Review 12(1): 1-19.

Homer, J. (1997) Structure, Data and Compelling Conclusions: Notes from the Field (1997 Jay W. Forrester Award Lecture), System Dynamics Review 13(4): 293-309.

Hinkley, D.V. (1988) Bootstrap Methods, Journal of the Royal Statistical Society Series B Methodological 50(3): 321-337.

Janacek, G. and Swift. L. (1993) Time Series: Forecasting, Simulations, Applications. England: Ellis Horwood.

Kay, J. (1997) Asphalt Nation: How the Automobile Took over America, and How We Can Take It Back. New York, Crown Publishers.

Lekvall, P. and C. Wahlbin (1973) A Study of Some Assumptions Underlying Innovation Diffusion Functions, Swedish Journal of Economics 75: 362-377.

Lyneis, J.M. & Pugh, A.L. (1996) Automated vs. hand calibration of system dynamic models:

An experiment with a simple project model. Proceedings of the 1996 International System Dynamics Conference.

McFadden, D. (1978) Econometric Models of Probabilistic Choice, Structural Analysis of Discrete Data with Econometric Applications. Cambridge, MA: MIT Press.

McFadden, D. (2001) Economic Choices, American Economic Review 91(3): 351-378.

Miller, D. (2001) Car Cultures. Oxford; New York, Berg.

Norton, J. and F. Bass (1987) A Diffusion-Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products, Management Science 33(9): 1069-1086.

Oliva, R. (2003) Model Calibration as a Testing Strategy for System Dynamics Models, European Journal of Operational Research 151(3): 552-568.

Ogden, J. (2004) Where will the Hydrogen Come From? System Considerations and Hydrogen Supply. The Hydrogen Economy Transition: Moving Toward the Post Petroleum Age in Transportation. Amsterdam: Elsevier Academic Press.

Rogers, E. (1962) Diffusion of Innovations. New York: Free Press.

Urban, G., J. Hauser and J. Roberts (1990) Prelaunch Forecasting of New Automobiles, Management Science 36(4): 401-421.

Urban, G., B. Weinberg, and J. Hauser (1996) Premarket forecasting really new products, Journal of Marketing 60(1): 47-60.

Sterman, J. (2000) Business dynamics: systems thinking and modeling for a complex world. Boston, Irwin/McGraw-Hill.

Struben, J. J. R. (2006) Essays on Transition Challenges for Alternative Propulsion Vehicles and Transportation Systems. PhD Dissertation. MIT Sloan School of Management. Cambridge, MA.

Struben, Jeroen J.R. and Sterman, John (2007) Transition Challenges for Alternative Fuel Vehicle and Transportation Systems. MIT Sloan Research Paper No. 4587-06.

Theil, H. (1966). Applied Economic Forecasting. Amsterdam: North Holland Publishing Company.

Acknowledgements

I would like to express my sincere gratitude to Jeroen Struben for his tireless guidance. Jeroen is a brilliant modeler and tremendous thinker. The numerous discussions with him as well as his comments are invaluable to this research. The memory of working with him will never fade.

I am also deeply indebted to my supervisor Professor John D. Sterman for his guidance, inspiration, and support for my work. As a respected leading scholar and engaging teacher, he gives fascinating lectures, encourages systematic thinking, and inspires students in a variety of ways. I enjoyed the experience of assisting him in teaching 15.874, seeing his preparation for lectures, his philosophy of teaching and his approach in organization. The modeling skills I learned from his 15.879 seminar are of much value. I greatly appreciated his efforts to push my work forward while providing me freedom to explore the area of my interest and supporting me when setbacks happened. I indeed feel lucky to have found him as an advisor.

Enormous thanks are also for my parents and beloved one. Without their encouragement and support, I would not have had such a fulfilling and enjoyable time at MIT.

The time at MIT was a journey to me, accompanied by enjoyment, bitterness, wonder and fulfillment. I am glad to have achieved this accomplishment, and look forward to the coming new period of life.

"吾令羲和弭节兮，望崦嵫而勿迫。

路漫漫其修远兮，吾将上下而求索。"