Modeling and accounting for interactions among multiple stakeholders in design optimization

Garrett Waycaster

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Cotutelle internationale avec l'Université de Floride

Présentée et soutenue par :
Garrett Waycaster

le jeudi 16 juillet 2015

Titre :

Stratégies d'optimisation par modélisation explicite des différents acteurs influençant le processus de conception

---

École doctorale et discipline ou spécialité :
ED MEGEP : Génie mécanique, mécanique des matériaux

Unité de recherche :
Institut Clément Ader

Directeur/trice(s) de Thèse :

Christian Bes
Christian Gogu

Jury :

Nicolas Gayton, MCF, Institut Français de Mécanique Avancée (rapporteur)
Rodolphe Le Riche, DR CNRS, Ecole des Mines de Saint Etienne (rapporteur)
Jonathan Hamilton, PR, Université de Floride (président)
Raphael Haftka, PR, Université de Floride
Christian Bes, PR, Université Toulouse III (directeur de thèse)
Christian Gogu, MCF, Université Toulouse III (codirecteur de thèse)

To my wife, Lauren

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to my advisors Dr. Raphael Haftka and Dr. Nam-Ho Kim. I have greatly enjoyed the opportunity to study under your supervision and sincerely appreciate your guidance and insight during my studies.

I would like to thank my committee members at the University of Florida, Dr. John Schueller and Dr. Jonathan Hamilton, for sharing their time and expertise with me. I am grateful to my advisors at Université Paul Sabatier, Dr. Christian Bes and Dr. Christian Gogu for their ongoing collaboration and for hosting my summer research experience in Toulouse. I also acknowledge Dr. Volodymyr Bilotkach and his contributions to this work through our regular interactions during my studies.

My lab mates in the Structural and Multidisciplinary Optimization Group have been an incredible support during my studies and I thank each of you for your collaboration, enlightening discussion, and your friendship.

I am incredibly grateful for the support of my family during my studies, without which none of this would have been possible. Most importantly, I thank my wife Lauren for her love and support which has allowed me to pursue my studies.

Finally, I extend my thanks to the National Science Foundation grant CMMI-0927790 and 1131103 for funding this work, and to the Embassy of France in the United States Chateaubriand Fellowship for supporting my studies in Toulouse.

TABLE OF CONTENTS

LIST OF TABLES

Table page

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODELING AND ACCOUNTING FOR INTERACTIONS BETWEEN MULTIPLE STAKEHOLDERS IN DESIGN OPTIMIZATION

By

Garrett Waycaster

August 2015

Chair: Raphael Haftka, Christian Bes
Cochair: Nam-Ho Kim, Christian Gogu
Major: Mechanical Engineering

The commercial success or failure of engineered systems has always been significantly affected by their interactions with competing designs, end users, and regulatory bodies. Designs which deliver too little performance, have too high a cost, or are deemed unsafe or harmful will inevitably be overcome by competing designs which better meet the needs of customers and society as a whole. Recent efforts to address these issues have led to techniques such as design for customers or design for market systems.

In this dissertation, we seek to utilize a game theory framework in order to directly incorporate the effect of these interactions into a design optimization problem which seeks to maximize designer profitability. This approach allows designers to consider the effects of uncertainty both from traditional design variabilities as well as uncertain future market conditions and the effect of customers and competitors acting as dynamic decision makers. Additionally, we develop techniques for modeling and understanding the nature of these complex interactions from observed data by utilizing causal models. Finally, we examine the complex effects of safety on design by examining the history of federal regulation on the transportation industry.

These efforts lead to several key findings; first, by considering the effect of interactions designers may choose vastly different design concepts than would otherwise be considered. This is demonstrated through several case studies with applications to the design of commercial transport aircraft. Secondly, we develop a novel method for selecting causal models which allows designers to gauge the level of confidence in their understanding of stakeholder interactions, including uncertainty in the impact of potential design changes. Finally, we demonstrate through our review of regulations and other safety improvements that the demand for safety improvement is not simply related to ratio of dollars spent to lives saved; instead the level of personal responsibility and the nature and scale of potential safety concerns are found to have causal influence on the demand for increased safety in the form of new regulations.

STRATEGIES D'OPTIMISATION PAR MODELISATION EXPLICITE DES DIFFERENTS ACTEURS INFLUENÇANT LE PROCESSUS DE CONCEPTION

Garrett Waycaster

Août 2015

Directeurs: Christian Bes, Raphael Haftka

Co-directeurs: Christian Gogu, Nam-Ho Kim

Spécialité: Génie mécanique, mécanique des matériaux

Le succès ou l'échec commercial des systèmes complexes (e.g. avions de transport commercial) ne dépend pas que de la qualité technique intrinsèque du produit mais il est aussi notablement affectée par les différentes interactions avec les autres acteurs du milieu tels la concurrence, les utilisateurs finaux et les organismes de réglementation. Des produits qui manquent de performances, ont un coût trop élevé, ou sont considérées comme dangereux ou nuisibles seront inévitablement surmontés par des produits concurrents qui répondent mieux aux besoins des clients et de la société dans son ensemble.

Dans cette thèse, nous cherchons à utiliser le cadre de la théorie des jeux afin d'intégrer directement l'effet de ces interactions dans un problème d'optimisation de conception qui vise à maximiser la rentabilité du concepteur du système. Cette approche permet aux concepteurs de prendre en considération les effets de l'incertitude venant d'une part des sources traditionnelles de variabilités (propriétés matériaux, tolérances géométriques, etc) ainsi que d'autres incertitudes de nature non technique reflétant par exemple l'incertitude du futur marché ou les interactions avec les autres parties prenantes (nouveaux produits concurrents, nouvelles règlementations, etc). Dans ce cadre nous développons également des techniques de modélisation utilisant des modèles causaux afin de comprendre la nature d'interactions complexes à partir des données observées. Enfin, nous examinons les effets complexes entre sureté de fonctionnement et conception en examinant l'histoire de la réglementation fédérale sur l'industrie du transport.

Ces travaux ont mené à plusieurs résultats clés. D'abord, en considérant l'effet des interactions entre les différents acteurs, les concepteurs peuvent être amenés à faire des choix techniques très différents de ceux qu'ils auraient fait sans considérer ces interactions. Cela est illustré sur plusieurs études de cas avec des

applications à la conception d'avions de transport commercial. Deuxièmement, nous avons développé une nouvelle méthode de construction de modèles causaux qui permet aux concepteurs d'évaluer le niveau de confiance dans leur compréhension des interactions entre les différents acteurs. Enfin, nous avons montré par une étude liens entre la réglementation et les améliorations de la sureté que la demande pour l'amélioration de la sureté ne répond pas toujours à une rationalité économique. En revanche la demande pour plus de sureté est fortement influencée par des facteurs tels que le niveau de responsabilité personnelle et la nature et ampleur des accidents potentiels.

CHAPTER 1
INTRODUCTION

French Chapter Summary

L'objectif de cette thèse est de développer un cadre permettant aux concepteurs de modéliser et de tenir compte au cours du processus de prise de décision des interactions complexes entre les acteurs multiples mis en jeu. Le cadre de la théorie des jeux sera utilisé afin d'apporter des informations sur ces interactions directement dans la fonction objectif d'un concepteur. Les effets de la prise en compte de ces interactions sera analysé sur des exemples numériques simples ainsi que des problèmes réalistes en phase de développement conceptuel d'un système complexe. Ces interactions étant difficiles à modéliser et prédire, des méthodes seront développées pour quantifier les interactions majeures basées sur des données historiques grâce à l'utilisation de modèles causaux. La robustesse, la précision, et l'estimation des incertitude de ces modèles seront analysées afin d'assurer leur utilité pour les prises de décision au cours de la conception. Nous étudierons enfin comment la question complexe de la sureté de fonctionnement est actuellement considéré dans le domaine des transports et explorerons les facteurs jouant le rôle le plus important. Cette analyse permettra de construire de meilleurs modèles pour représenter la valeur de la sureté vis-à-vis des différents types de risques présents dans les systèmes de transport.

Motivation

An important part of many modern design problems is understanding how uncertainties like variability in material properties, unknown operating conditions, or design tolerances will impact the performance and reliability of the final design. An equally important but often overlooked form of uncertainty is the way competing designs, customer preferences, and market changes affect design value or profitability as a function of system cost, reliability, and performance. For example, the value of a new, more fuel efficient airliner will depend on demand for air travel, future fuel prices, and the performance of other available aircraft. The goal of this work is to provide design decision makers with the tools to account for these interactions, to generate models of important interactions, and to predict the effects of complex ideas like safety on design value.

While designers have long been aware of these sorts of economic uncertainties, understanding the complex ways in which these factors affect design has proven difficult. Previous works such as Vincent [1], Rao [2], Badhrinath and Rao [3], and Lewis and Mistree [4] have demonstrated the use of game theory for solving multidisciplinary design problems, but have not addressed the application of game theory to economic uncertainty and interactions. Li and Azarm study the design of a product [5] or product family [6] in the presence of competitive products in the market and uncertain customer preferences, but do not model customers or competitors as dynamic decision makers. Subrahmanyam [7] also considers the idea of market uncertainties as affecting design optimality, but these uncertainties are taken as stochastic values and are not affected by design decisions. Morrison [8] applies game theory to a case study of fuel efficiency innovation among competing airlines, but does not consider additional stakeholders or applications to design optimization. Other important contributions include the foundational ideas of decision based design [9] and value driven design [10] as tools for explaining design value as a function of performance attributes.

Understanding the way these interactions occur in real life systems poses an additional challenge. Little existing research attempts to meaningfully quantify the real world relationship between design decisions and performance with design value. While this relationship is a foundational consideration of value driven design [10] as well as other design methods, the value functions are typically based on simple assumptions about customer preferences and market conditions with no consideration about how design changes might affect customer behavior or competitive products. One tool for measuring these relationships is causal models [11], which have been utilized in economics, health science, and political science as well as engineering applications. Causal models are based on a Bayesian networks framework and have found some applications in engineering cost modeling [12] as well as the area of maintenance prognostics [13]. Methods for fitting causal models generally rely on a score and search criteria and include the Peter and Clark (PC) algorithm [14], the Spirtes-Glymour-Scheines (SGS) algorithm [15], the Inductive Causation (IC) algorithm [16], and the Sparsest Permutation (SP) algorithm [17] for non-temporal data and methods such as the Time Series Causal Model (TSCM) algorithm [18] for temporal data. One factor lacking from many of the existing methods is a quantification of the uncertainty in causal model parameters and predictions, which are critically important for making informed decisions during design.

The importance of safety and risk in design, particularly in the area of aviation and transportation, should not be overlooked. One may gain understanding of the relationship between safety and design by considering the way safety improvements are analyzed with respect to their cost. Several significant prior works have examined the way the cost effectiveness of safety improvements is determined in various fields. Viscusi and Aldy [19] provide a detailed overview of various factors affecting the applied value of statistical life. Morrall [20] and Tengs *et al*. [21] both provide reviews of the cost effectiveness of previously implemented or proposed life-saving measures across many different fields. Cropper and Portney [22] outline some of the difficulties faced by regulators and policy makers in attempting to quantify cost effectiveness for new safety measures. Hammitt and Graham [23] outline the difficulty in assessing survey respondents' willingness to pay for safety, particularly in the case of highly unlikely events. Arrow *et al*. [24] provide a discussion of the ways in which cost-benefit analysis can and should be used to shape public policy.

Objective

The objective of this research is to develop a method which allows designers to model and account for complex interactions between multiple stakeholders during the decision making process. A game theory framework will be utilized in order to bring information about interactions directly into a designer's objective function. The effects of considering these actions on simple numerical examples as well as realistic conceptual design problems will be considered. Since these interactions are difficult to model and predict, methods will be developed for quantifying important interaction effects based on historical data through the use of causal models. The robustness, accuracy, and uncertainty estimates of these models will be tested in order to ensure their usefulness for design decision makers. We will study how the complex issue of safety is currently considered in economic analysis and explore some of the important factors related to the treatment of safety in existing transportation systems. This analysis will allow for better constructs for representing the value of safety as it compares to certain types of risk in transportation systems.

Organization of the Dissertation

This dissertation is divided into five chapters. Chapter 2 details the methodology developed to account for interactions between multiple stakeholders using game theory and some accompanying example problems and case studies. Chapter 3 describes the use of causal models for quantifying models of interactions between stakeholders, procedures for fitting causal models from observed data, and quantification of model confidence and parameter uncertainty including applications to real world data. Chapter 4 discusses the difficulty in assigning monetary value to system safety through a cost-effectiveness review of transport regulations and corrective actions over the past decade. Chapter 5 provides a summary of the work with overall conclusions and suggestions for further investigation.

CHAPTER 2
DESIGN OPTIMIZATION WITH STAKEHOLDER INTERACTIONS

French Chapter Summary

De nombreux systèmes complexes mettent aujourd’hui en jeu différentes parties prenantes, où chaque acteur interagit de manière spécifique avec les autres parties. Dans le cas le plus simple, il n’y a que deux acteurs : un concepteur qui détermine les caractéristiques du système et un client qui détermine la façon d'utiliser le système. Dans le cas des systèmes plus complexes, nous pourrions avoir également des opérateurs du système, des régulateurs, et des fournisseurs. Nous pouvons en outre avoir de multiples acteurs au sein de chacun de ces groupes en concurrence les uns avec les autre, par exemple il y a typiquement plusieurs concepteurs cherchant chacun à satisfaire un besoin du marché avec des produits similaires mais légèrement différents. Chacun de ces acteurs agit comme un décideur dynamique, agissant et réagissant sur la base des décisions prises par d'autres parties prenantes. Ces types d'interactions peuvent être déterminantes sur le succès ou l'échec d'un produit, plus encore que ses qualités techniques intrinsèques.

Il existe actuellement plusieurs méthodes que des concepteurs peuvent utiliser pour tenter de comprendre ces interactions, la plupart du temps en essayant de découvrir les préférences des autres acteurs. Le plus souvent, les concepteurs utilisent des données historiques relatives au succès ou l’échec de différents types de systèmes développés dans le passé. Un concepteur peut également communiquer directement ou indirectement avec les autres acteurs, par exemple à travers une étude de marché, pour tenter de déterminer l'importance relative de différents indicateurs de performance. Cependant, ces méthodes ne sont pas exactes, et la compréhension des préférences des autres parties prenantes auront ainsi erreur. Cela peut être dû à un biais d'échantillonnage de modèles existants, à l’extrapolation dans un nouvel espace de conception, ou en cas de communication directe, une mauvaise communication des préférences, soit par l'ignorance d'une partie prenante de leurs propres préférences ou d'une tentative délibérée d'influencer les décisions des concepteurs. Nous pouvons considérer ces erreurs dans la compréhension des préférences des autres acteurs comme une incertitude de nature économique au sens large, modifiant directement la fonction objectif d'un concepteur et affectant donc les choix fait au travers du processus d'optimisation de la conception.

Afin de comprendre les effets des interactions entre ces différents acteurs, nous proposons d'utiliser la théorie des jeux. La théorie des jeux a été développé en économie comme un moyen de modéliser la prise de décisions stratégiques entre acteurs rationnels, nommés joueurs. Selon la façon dont les joueurs interagissent et l'information partagée entre eux, nous pouvons arriver pour un même problème de base à des résultats très différents. En ce qui concerne le problème d'optimisation de la conception d'un système, la théorie des jeux nous permet de mettre à jour de manière adaptative la fonction objective à optimiser, basé sur notre position de départ dans l'espace de conception, les changements dans le marché, et les actions des autres parties prenantes, considérés comme des joueurs ici. Dans ce chapitre nous reformulons un problème d'optimisation multidisciplinaire en conception préliminaire avion pour tenir compte des interactions dynamiques entre différentes parties prenantes en utilisant un modèle de la théorie des jeux avec des interactions simultanées ou séquentielles. Nous montrons, sur cet exemple aéronautique, l'importance de considérer ces interactions dans cette phase de conception qui permet de mieux rendre compte des différents compris à faire. Nous analysons également une situation dans laquelle le concepteur intègre dans l'optimisation un marché futur incertain ainsi que d'autres incertitudes de nature économique. Les résultats montrent que ces incertitudes ont un impact très important sur les choix effectués, bien plus que les incertitudes traditionnelles, de nature non-économiques qui sont habituellement considérés en conception robuste.

Overview

Many modern engineered systems involve multiple stakeholders, each providing some inputs and receiving some outputs with respect to the system. In the simplest case, this might be a designer who determines system characteristics and a customer who determines how to utilize the system. In more complex systems, we might also have system operators, regulators, or suppliers. We may additionally have multiple stakeholders within each of these groups competing with one another, for example multiple designers each providing similar products to their customers. Each of these stakeholders acts as a dynamic decision maker, acting and reacting based on the decisions made by other stakeholders. These types of interactions can have a dramatic effect on the success or failure of a design.

There are several methods designers currently use to attempt to understand these interactions, mostly by attempting to uncover the preferences of other stakeholders. Most frequently, designers use legacy information based on the types of designs they and their competitors have produced before and the success of those designs. A designer may also use direct communication with other stakeholders, such as via a market study, to attempt to determine the relative importance of different performance metrics. However, these methods are not exact, and the resulting understanding of stakeholder preferences will have some error. This may be due to sampling bias of legacy designs, extrapolation into a new design space, or in cases of direct communication, miscommunication of preferences, either through a stakeholder’s ignorance of their own preferences or a deliberate attempt to sway the designers’ decisions. We can consider these errors in understanding stakeholder preferences as an economic uncertainty, directly changing a designer’s true objective function and therefore affecting the design optimization process.

In order to understand the effects of these stakeholder interactions, we can utilize game theory [25]. Game theory has been developed in economics as a way to model strategic decision making between rational stakeholders, or players. Depending on the way players interact and the information shared between them, we can arrive at different outcomes for the same basic design problem. From the perspective of our optimization problem, game theory allows us to adaptively update our objective function, relating the performance characteristics of our design to designer profits, based on our location in the design space, changes in the market, and actions of other stakeholders. We will introduce this idea in more detail with some simple examples in the next section.

As discussed in the introduction, several prior works have utilized game theory frameworks to look at aspects of design optimization. However, the authors are not aware of any existing attempts to develop a combined framework to consider dynamic decision making of customers and competitors while also considering uncertain future markets. The objective of this chapter is to reformulate a multidisciplinary design optimization problem to account for dynamic interactions between multiple stakeholders and market changes using a game theory model with both simultaneous and sequential interactions considered. We will additionally demonstrate, using simple examples from the aerospace industry, why considering these interactions during design optimization is important, and how it provides a designer with more information about design trade-offs.

Methodology for Formulating Optimization to Include Interactions

Formulation of optimization considering interactions

For the purpose of this dissertation, we will focus on how we can reformulate an optimization problem when considering the effects of the interactions between l stakeholders. Readers interested in the principles of game theory can find more information from introductory game theory text books such as Fudenberg and Tirole [25]. First, let us consider a basic multidisciplinary design optimization problem formulation:

$$\max_{X} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{w_i f_i}(\boldsymbol{X}) \qquad (2\text{-}1)$$
$$s.t.\, g_j(\boldsymbol{X}) \geq 0 \; for \; j = 1, \dots, m$$

where **X** is our vector of design variables, f_i describes the ith performance metric of the design, w_i is the weight of the ith performance metric in the optimization, and g_j describes the jth of m many design constraints.

By varying the vector $\boldsymbol{w}$ in this optimization, we can calculate a set of Pareto optimal designs with a different set of performance attributes, W as specified by the performance equations $f_i(X)$.

$$W = [f_1(X), \dots, f_n(X)] \qquad (2\text{-}2)$$

Now consider that for each design and set of performance values (that is, each weight vector **w**) we can define some profit function for our designer,

$$\boldsymbol{\Pi_1}(\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{E}) \qquad (2\text{-}3)$$

where **Y** describes the decision vector of the other stakeholders in the design and E describes a set of exogenous variables not directly controlled by any stakeholders. This function is used to transform our design performance and other stakeholder decisions directly into the profit for the designer.

The decision vector $\mathbf{Y}$ will be determined by the other stakeholders attempting to maximize their own expected profits, such that

$$\boldsymbol{Y_k} = \underset{Y_k}{\operatorname{argmax}}\, \boldsymbol{\Pi_k}(\boldsymbol{Y_k}, \boldsymbol{W}, \boldsymbol{Y_{\sim k}}, \boldsymbol{E}) \;\; for\ k = 1, \dots, l \tag{2-4}$$

where Π_k describes the profit of the kth stakeholder, $\mathbf{Y_k}$ is the decision vector of the kth out of l many stakeholders, and $\mathbf{Y_{\sim k}}$ is the decision vector of the remaining $l - 1$ stakeholders

We now have $l + 1$ profit functions and $l + 1$ decision sets. This can be thought of as $l+1$ different optimization problems, each dependent on the same decision vector for all players, forming an over determined set of equations. In order to determine a solution, we must apply a set of rules; in our case this is based on a certain game structure that describes the amount of information shared between stakeholders and the order in which decisions are made. Information shared between stakeholders refers to how well each stakeholder is able to approximate the profit functions of the others. For example, a designer may not explicitly know the profit function of their customer, but may make an approximation based on prior designs. We will also show that there situations may arise where one stakeholder may have an incentive to deliberately mislead another stakeholder in order to create a more favorable situation for themselves. This type of behavior need not be detrimental for the stakeholder being misled, and can in some cases be advantageous for both parties.

The order of decisions may be either simultaneous, sequential, or partially both. Sequential decision making means one stakeholder chooses their decision vector first and passes that decision on to the next stakeholder in the sequence. Stakeholders moving first will approximate the reaction of each subsequent stakeholder based on their available information about those stakeholders' profit functions. These approximated reactions are known as a best reply function [25]; that is, given that stakeholder one chooses Y_1, stakeholder 2 will maximize their expected profit by playing Y_2, or simply

$$\boldsymbol{Y_i} = \boldsymbol{\varphi_{ij}}(\boldsymbol{Y_j}, \boldsymbol{\widehat{Y}}) \tag{2-5}$$

where φ_{ij} is the best reply function that relates the given Y_j to the best reply Y_i and $\widehat{Y}$ is the vector of decisions of all the other stakeholders, some of which may be known based on the sequence of the game, and others which require their own best reply function to determine. Each of these can be solved recursively to determine a best reply function for each subsequent decision maker.

We can therefore formulate our profit maximization problem for the designer by combining equations (2-1), (2-3), and (2-5), where the decisions of stakeholder acting in sequence before the designer are given as inputs, and the best reply function for stakeholders acting after the designer act as constraints. This problem will be subject to uncertainty in the exogenous inputs, **E**, as well as uncertainty due to approximations made in determining the best reply function, φ.

$$\begin{gathered} maximize\ \Pi_1(\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{E}) \\ \boldsymbol{X} = \underset{X}{\operatorname{argmax}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{w_i f_i}(\boldsymbol{X}) \\ W = [f_1(X), \dots, f_n(X)] \\ \boldsymbol{s.t.\, g_j(X) \geq 0\ for\ j = 1, \dots, m} \\ \boldsymbol{Y_k} = \varphi_{k1}(\boldsymbol{W}, \boldsymbol{Y_{\sim k}})\ for\ k = 2, \dots, l \end{gathered} \tag{2-6}$$

In the case of simultaneous decisions, we must use the concept of Nash equilibrium [25] to determine a solution. A Nash equilibrium is a point in the decision space where no stakeholder can improve their own profit function by changing their decision vector. This means that a Nash equilibrium acts as a self-enforcing agreement between the players. That is to say, (X, Y) is a Nash equilibrium if and only if

$$\boldsymbol{\Pi_1}(\boldsymbol{W},\boldsymbol{Y},\boldsymbol{E}) > \boldsymbol{\Pi_1}(\boldsymbol{W}^*,\boldsymbol{Y},\boldsymbol{E})\ for\ all\ \boldsymbol{W}^* \neq \boldsymbol{W}, and$$
$$\Pi_k(\boldsymbol{Y_k},\boldsymbol{W},\boldsymbol{Y_{\sim k}},\boldsymbol{E}) > \Pi_k(\boldsymbol{Y_k^*},\boldsymbol{W},\boldsymbol{Y_{\sim k}},\boldsymbol{E})\ for\ all\ \boldsymbol{Y_k^*} \neq \boldsymbol{Y_k}\ , k = 2, \ldots, l \tag{2-7}$$

We can find any pure strategy Nash equilibria by formulating a best reply function for each stakeholder and solving that system of equations to determine where all the best replies intersect. A pure strategy Nash equilibrium means a stakeholder plays a single deterministic decision vector, while a mixed strategy means a stakeholder randomly selects from multiple pure strategies with some predetermined probability of each. It should be noted that there is no guarantee of a single unique Nash equilibrium, and equilibria can exist in both pure and mixed strategies. To solve our problem using simultaneous decision making, we are no longer performing an optimization. Instead, we are looking for the intersection of the surfaces defined by the best reply functions for each of our stakeholders. These intersections represent pure strategy equilibria, of which there may be multiple or none. In cases of multiple Nash equilibria, we can sometimes eliminate some equilibrium through so called refinements. For the purposes of this dissertation, we will present all Nash equilibria as possible outcomes, and we will only deal with simultaneous decision making in the discrete decision context for simplicity.

Numerical Example Problem

Having defined how we may formulate an optimization problem considering interactions with other stakeholders, let us consider a simple example based only on stakeholder profits without considering a design problem. We have two stakeholders, an aircraft manufacturer who specifies the design and their customer the airline. Both are monopolists, meaning they face no competition. We assume that the designer leases aircraft to the airline at a per flight cost that is fixed, regardless of the aircraft design or the number of flights.

The designer's only decision variable is the level of technology to invest in the aircraft, T. This can be thought of as the design effort and material and labor cost associated with producing the aircraft. For our problem, we will consider T to be bounded between 0 and 1. T acts as the only weighting variable w as described in equation (2-1), where a value of 0 is the optimal manufacturing cost, and a value of 1 is the optimal customer value.

The airline's decision variable is the number of flights that they will offer, Q, which will determine the price they charge per ticket based on a fixed linear demand for air travel. The airline has some fixed cost of operation per flight, some cost that is proportional to the price of jet fuel, c_F, and some benefit based on the level of technology invested in the aircraft. We can then formulate the profit functions for both stakeholders as follows

$$\Pi_d(T, Q) = Q(L - c_T T) \tag{2-8}$$

$$\Pi_a(T, Q, c_F) = Q(P(Q)N_p - c_F F - c_L L + v_T T) \tag{2-9}$$

where c_T is the cost to implement new technology for the designer, F is the fuel consumption per flight, L is the lease cost per flight, c_L is some factor greater than 1 describing the total fixed costs for the airline including lease cost, v_T is the value of technology to the airline, N_p is the number of passengers per flight, and $P(Q)$ is the price per ticket based on the linear demand function, given by

$$P(Q) = a - bQN_p \tag{2-10}$$

To create a meaningful example, we first find some reasonable estimates for some of the unknown coefficients in our problem. We select a Boeing 737-700 as the baseline aircraft for our analysis. Considering the standard configuration capacity of 128 passengers [26] and an average load factor of roughly 0.8 [27], we take the number of passengers per flight, N_p, as 100. Given an average flight length of 1000 miles [27], we calculate the fuel consumption per flight, F , as roughly 1500 gallons [28]. Average recent jet fuel prices are around \$3.00 per gallon [29], and we consider a range up to $5.00 to account for possible future changes. Based on the 737-700 list price of $76M [30] and a useful life of 60,000 flights [31] we find a per flight cost of $1,300. Considering additional storage and maintenance costs as roughly doubling this expense, we select the per flight lease cost of the aircraft, L, as $3000. Based on available airfare cost breakdown data [32], we consider that c_L ranges from 10 to 12, meaning that the capital cost of the aircraft ranges from 8% to 10% of the total cost per flight, depending on the airline. In order to determine characteristic numbers for the cost and value of new technology, we consider a new aircraft design project. We consider that this new design will cost an additional $850 per flight, roughly a 25% increase from the initial design, and provides a benefit of $4200 per flight through increased capacity, efficiency, and passenger comfort. Finally, by collecting data on tickets sold and average ticket price over the past 20 years, we fit the linear relationship between quantity and price as shown in Figure 2-1. This approximation assumes that the airline uses this single aircraft design to service all of their routes.

Now let us consider the simplest case of interaction, where the designer first decides on the level of technology investment with full information about the airline profit function, and the airline then determines the quantity of flights in a sequential game. Note that both profit functions, equations (2-8) and (2-9), are concave functions. We can therefore calculate a best reply function for the airline by setting to zero the first derivative of the airline profit function with respect to Q and solving for Q, such that

$$\frac{d\Pi_a}{dQ} = v_T T - c_L L - c_F F + N_p\left(a - bN_p Q\right) - N_p^2 Q b \tag{2-11}$$

$$Q^* = \varphi_{da}(T) = \frac{aN_p + v_T T - c_F F - c_L L}{2bN_p^2} \tag{2-12}$$

We can substitute this best reply function into the designer's profit function to replace Q and solve for the designer's optimal value of T by setting to zero the derivative of the designer's profit function with respect to T and solving for T,

$$\frac{d\Pi_d}{dT} = \frac{v_T(L - c_T T) + c_T(c_F F + c_L L - aN_p - v_T T)}{2N_p^2 b} \quad (2\text{-}13)$$

$$T^* = \frac{v_T L + c_F c_T F + c_L c_T L - ac_T N_p}{2c_T v_T} \quad (2\text{-}14)$$

Using our values for our various coefficients, we can calculate the decision of the designer and airline and the profit for each. Since we have ranges of values for fuel price and the airline cost factor, we perform this analysis at the 4 extreme cases of these coefficients as shown in Table 2-1. Because our problem is linear in these values, we can interpolate between these 4 points to find the decisions and profits at any combination. Note that in the first case, the designer would choose an optimal value of slightly negative technology investment, however we restrict this value to be between 0 and 1. It can be seen that the optimal decisions and resulting profits for both the designer and airline vary greatly with these possible changes in parameters c_F and c_L.

In a realistic design problem, we will likely consider that a designer must make design decisions without knowledge of future fuel prices. These prices will be unknown to the airline as well. A designer will then maximize expected profits based on the possible distribution of future fuel prices. Due to the simple linear nature of our example problem, this will be the same as designing based on the mean value of future fuel prices.

A designer may face additional uncertainty in their understanding of the airlines' profit function, for example in the value of c_L. However, the airline will be able to know this value exactly. This is known in game theory as a game of "incomplete information" [25]. This means the designer will face some error in their prediction of the best reply function of the designer, specifically

$$Q^* = \varphi_{\tilde{da}}(T) = \frac{aN_p + v_T T - c_F F - (c_L + \varepsilon)L}{2bN_p^2} \tag{2-15}$$

where ε describes the error in the designers estimation of airline costs due to insufficient information or changes in c_L over the time lag between the designer and airline decisions.

We can see from our previous example that the designer will invest more in technology if they believe the airlines fixed cost, c_L, is higher. This is because higher fixed costs mean the effect of technology on airline marginal profits is more significant, and therefore more technology investment will have a greater effect on the quantity of flights. This relationship implies that airlines will have an incentive to mislead designers into believing that their costs are higher than in reality, shifting profits away from designers and toward airlines. Without considering the effects of these interactions, designers will be unable to understand the effects of these potential uncertainties.

To explore these interactions in more detail, let us switch from a continuous game to a discrete one. In this case, the designer must either decide to invest in new technology ($T = 1$) or not ($T = 0$). The airline will decide whether to expand their market by offering a higher number of flights ($Q = 2.5M$), or to maintain their current levels ($Q = 1.5M$). We consider that fuel prices will either be \$3 per gallon with probability p_F or \$5 per gallon with probability $1 - p_F$. Finally, the designer assumes the airline is a low cost carrier ($c_L = 10$) with probability p_C or a high cost carrier ($c_L = 12$) with probability $1 - p_C$. We can express this problem using a decision tree, known in game theory as an extensive form game [25].

In the figure, each node represents a decision, and dashed lines between nodes indicate an information set, where the decision maker must act without knowing for certain which node in the information set they are currently in. The solution will therefore depend on the decision maker's beliefs about the values of p_F and p_C. The payoffs for each resulting set of decisions are given at the end of each path, where the top number is the designer's profit, and the bottom number is the airline's profit, both in billions of dollars. We can simplify this game by eliminating dominated strategies for the airline, since we know at the last branch of the decision tree the airline will choose the value that maximizes their own profits; this is known as backwards induction. Figure 2-3 shows these dominated strategies in gray.

We see that, based on this discrete example, the designer can only influence the airline to utilize more flights by increasing technology investment if fuel prices are low and airline costs are high, or fuel prices are high and costs are low. In the remaining two cases, the designer will strictly prefer not to invest in new technology, since they will lease the same number of flights regardless and will have a higher profit margin for each. Airlines will always prefer the case where designers invest in technology, as they always gain higher profits.

From this simple example, we would conclude that if fuel prices are high, airlines will attempt to convince designers that they have low costs, as designers will believe they can then influence flight quantity by investing in technology. If fuel prices are low, airlines will attempt to convince designers that their costs are high, again in an effort to encourage designers to invest in technology.

We may also be interested to know if the possible solutions of this game change if we consider that designers and airline make decision simultaneously. For example, airlines submit orders for new aircraft without knowing future fuel prices or precise aircraft specifications. We can represent this sort of game using strategic form, with 4 payoff matrices representing the 4 possible combinations of fuel price and airline costs as shown in Figure 2-4.

The numbers in each box represent the payoffs for the airline and the designer, respectively. Numbers that are underlined indicate a best reply for that stakeholder. When both numbers are underlined in the same box, meaning the best replies intersect, we have a Nash equilibrium for that individual game, represented by circling that square. We can see that for the simple game we have constructed, it is never advantageous for the designer to invest in technology. This happens because since decisions are made at the same time, the designer's choice cannot influence the quantity selected by the airline. We can also see that when airline costs are high ($c_L = 12$), meaning we are on the two matrices on the right side, the equilibrium solution for this game will be ($T = 0$), ($Q = 1.5M$). When airline costs are low, the equilibrium will depend on the probability of low fuel prices, p_F, as the airline will attempt to maximize their expected profits. If the airline believes p_F is less than 0.11, they will always choose the low quantity ($Q = 1.5M$), and if they believe p_F is greater than 0.11 the airline will choose the high quantity, ($Q = 2.5M$). When p_F is equal to 0.11, the airline is indifferent between these two strategies and may play either one, or play a mixed strategy where they randomly select between both options. It should be noted that the designer would strictly prefer the airline select the higher quantity, but based on this game structure, they have no way to influence that decision.

It should be noted that the solutions we have found for each of these different types of games need not be Pareto optimal in terms of profits for both stakeholders. For example, in Figure 2-4, we can see that both the designer and a high cost airline ($c_L = 12$) would be strictly better off playing the strategy ($T = 1$), ($Q = 2.5M$) as compared to the equilibrium strategy ($T = 0$), ($Q = 1.5M$), regardless of the values of fuel price and airline costs. However, that strategy is not an equilibrium solution because one or both of the stakeholders can improve their profits by modifying their decision. For example, in the case of $[c_F = 5, c_L = 12]$ starting at

$(T = 1)$, $(Q = 2.5M)$, we see that the designer would strictly prefer to select $(T = 0)$ when the airline plays $(Q = 2.5M)$, and similarly the airline prefers $(Q = 1.5M)$ against $(T = 1)$. Because the strategies and payoffs are known, each player will realize the other will try to change their own strategy, and will respond accordingly, resulting in selecting $(T = 0)$, $(Q = 1.5M)$. This is a variation on the classical game theory example known as the prisoner's dilemma [25].

Application to Conceptual Design Optimization of an Aircraft Wing

Design optimization frequently deals with uncertainty due to variations in material properties, operating conditions, and design specifications. One often overlooked source of uncertainty is in the way designers determine tradeoffs between multiple objectives. These tradeoffs affect the value of the design to customers, regulators, and other interested stakeholders in the design, which ultimately determines the profitability of the design. However, traditional multi-objective design optimization rarely considers these dynamic interactions, and when it does it models the preferences of other stakeholders using heuristic methods.

Aircraft design is often viewed as a characteristic multi-objective or multi-disciplinary problem. Aircraft design is also subject to complex relationships between stakeholders; these stakeholders include the airlines who buy the aircraft and the passengers who buy tickets. Additionally, because of the large time gap between design and entire service life, changing market conditions can play a major role in the success or failure of a design; for instance changes in fuel prices or public demand for air travel.

Designers set their objective function based on data regarding the preferences of airlines and the public, either learned from past experience or provided directly by these stakeholders. The interaction between stakeholders in the communication of these preferences is subject to uncertainties because stakeholders do not have perfect knowledge of their own interests now or in the future, but also because it may be advantageous for them to provide misleading information, leading to what is known as information asymmetry. We propose to utilize game theory to model how each of these stakeholders will interact with one another as they make strategic decisions to maximize their own welfare. In this application, we focus on quantifying the importance of considering these economic uncertainties relative to other sources of uncertainty in a characteristic problem.

Design Problem Description

The goal of our simple example problem is to represent a characteristic multidisciplinary design optimization problem using simple analytical formulas. The problem we look at is the design of a commercial transport aircraft wing, which provides a mixture of structural and aerodynamic performance goals. The aircraft designer specifies two configuration design variables: the wing aspect ratio AR and the design safety factor SF beyond what is required by regulations.

The aircraft designer makes a third decision on the number of structural tests N_{test} to perform which will affect the minimum acceptable knockdown factor for certification of the aircraft, similar to the A-basis criteria specified by the FAA [33]. Based on the probability of not meeting this certification criteria, the designer will be assessed some monetary penalty. Additionally, each test performed will have some fixed cost. This means the designer may choose to have higher design cost (more tests) in order to improve performance or reduce certification cost by allowing for a less conservative certification criteria.

The wing is idealized using a trapezoidal shape, where sweep, planform area, and taper ratio are based on the dimensions of a Boeing 737-700 [34] and are constant across all designs, meaning that changing the aspect ratio will scale the span and chord proportionally. The wing box is constrained by the wing cross-section, meaning increasing the aspect ratio will decrease the maximum possible design safety factor due to longer, more slender wings.

Once these design variables are determined, a structural designer optimizes the wing box for minimum weight subject to constraints on stress and deflection. Details on the models used to estimate aerodynamic and structural characteristics of the wing are provided in Appendix A.

This simple design problem introduces some basic tradeoffs similar to those seen in a true multidisciplinary design problem. By increasing the aspect ratio, the aircraft designer can reduce the aircraft fuel consumption, but this change cause penalties structural weight and probability of failure. The aircraft designer can reduce the probability of this design penalty either by increasing safety factor or increasing the number of tests performed. These trade-offs are summarized in Figure 2-5.

Reformulating Optimization

To be able to consider interactions in optimization, we must describe how decisions are made among stakeholders and how information is shared between them. In this case, we will consider that the designer is interacting with an airline, who determines the number of aircraft to purchase based on the number of tickets the airline is able to sell. To deal with these interactions, we use common terminology and techniques utilized in game theory [25].

One key concept is the idea of a *best reply function*, which defines a player's optimal strategy given the strategies of all other players. We can calculate a best reply function by taking the partial derivative of a player's objective function with respect to each of their decision variables, setting the result to zero and solving for the optimal value of that decision variable, as shown in equations (2-16) and (2-17).

$$\Pi_2(X_1, X_2) \tag{2-16}$$

$$\frac{\partial \Pi_2}{\partial X_2} = 0 \Longrightarrow X_2^* = f(X_1) \qquad (2\text{-}17)$$

Second order conditions are guaranteed by the fact that any meaningful profit function for a player should be concave in each player's own decision variables. The resulting expression will provide the optimal value of that decision as a function of the actions of all other players.

For our example problem, we will consider that the stakeholders play a sequential game, where the aircraft designer will act first to determine the nature of the aircraft available. After learning what the aircraft designer does, the airline will determine how many aircraft to purchase. In order for the aircraft designer to act first, they must estimate the best reply function of the airline which can then be inserted into the aircraft designer's own profit function. In doing so, the aircraft designer will incorporate the uncertainties faced by the airline directly into their own optimization problem.

The estimation of this best reply function may itself be subject to some error or uncertainty. We consider that we may have a case of asymmetric information, meaning one player has more available information than another. For instance in our example problem, the airline may know their own profit function exactly, while the aircraft designer may estimate some elements of the airline profit function with some error.

It can then be determined whether or not the airline has an incentive to signal, or to communicate information to the aircraft designer, that would either increase or decrease the error of this estimation by the aircraft designer. Similarly, we can see if the designer has an incentive to screen, or try to gather more information from the airline about their preferences. In some cases, errors in the aircraft designer's estimation of the airline's preferences might be good for both players, bad for both, or might increase one player's profit at the expense of the other.

Understanding the situations that give rise to these cases is an important factor in understanding airline and aircraft designer relations.

Economic Interaction Model

In order to model the interactions between aircraft designers and airlines, we must first develop reasonable ways to express the profits of each group. We attempt to specify objective functions that capture some important trade-offs and interactions for both stakeholders without excessive complexity. The interactions of stakeholders with the design problem and their exchange of information are summarized in Figure 2-6.

For the aircraft designer, revenues are based on the number of aircraft sold to airlines N_{air} and the price the aircraft designer decides to charge P_{air}. The aircraft designer's costs are based on a fixed initial cost of a new project C_{ini}, the number of tests they perform N_{test} which each have a fixed cost C_{test}, and the probability of a certification penalty P_{pen} which we assume will also have a fixed cost associated with making the design safety compliant C_{pen}. The aircraft designer's profit function Π_d is then given as

$$\Pi_d = N_{air}P_{air} - N_{test}C_{test} + P_{pen}C_{pen} - C_{ini} \qquad (2\text{-}18)$$

The airline's revenue is based on the price P_{tix} and quantity N_{tix} of tickets sold. We assume that each aircraft has a useful life of 60,000 flights with an average passenger load of 100 passengers per flight. Additionally, we consider that the number of aircraft is significant enough as to not face scheduling and route constraints. The demand for air travel is defined using a simple linear demand function, such that price is determined for a given number of flights, such that the maximum price is P_{max} and decreases with increasing tickets sold at a rate P_s.

$$P_{tix} = P_{max} - P_s N_{tix} \qquad (2\text{-}19)$$

Airlines have four different sources of costs; the first is based on the fuel consumption over the life of the aircraft given by the product of the number of flights N_{flight}, the cost of fuel C_{fuel}, and the aircraft fuel consumption per flight FC. The second cost is based on the acquisition price of the aircraft they choose to purchase P_{air}. The third cost component is a fixed cost C_{fix} for each aircraft, based on the labor, taxes, fees, and passenger services required for each aircraft. The final cost component is based on the level of safety of the aircraft being utilized, where the cost is equal to the product of the probability of failure PF, the number of passengers N_{pax}, and a penalty per life at risk; for this penalty we use the value of statistical life (VSL) specified by the Department of Transportation [35]. This cost term is intended to reflect the increased safety and maintenance costs related to flying less safe aircraft. Combining these components, the airline's profit function Π_a is

$$\Pi_a = N_{air}\left(N_{flight}N_{pax}P_{tix} - P_{air} - N_{flight}C_{fuel}FC - VSL * N_{pax}PF - C_{fix}\right) \quad (2\text{-}20)$$

We relate the number of tickets sold and the number of aircraft purchased N_{air} by assuming each aircraft is capable of 5 flights per day.

$$N_{air} = \frac{N_{tix}}{5 * 365N_{pax}} \quad (2\text{-}21)$$

Since we have specified a sequential game, the aircraft designer will need to estimate the actions of the airline using a best reply function. We can calculate the airline's best reply function by taking the first derivative of their profit with respect to the number of tickets sold. Combining equations (2-19), (2-20), and (2-21), we can rewrite the profit function as

$$\Pi_a = \frac{N_{tix}}{5 * 365 * N_{pax}}\left(N_{flight}N_{pax}(P_{max} - P_s N_{tix}) - P_{air} - N_{flight}C_{fuel}FC - VSL * N_{pax}PF - C_{fix}\right) \quad (2\text{-}22)$$

Taking the derivative with respect to N_{tix}, setting the result equal to zero, and solving for the optimal value of N_{tix} yields

$$N_{tix}^{*} = \frac{1}{2N_{flight}N_{pax}P_s}(N_{flight}N_{pax}P_{max} - C_{fix} - P_{air} - C_{fuel}N_{flight}FC - VSL * N_{pax}PF) \tag{2-23}$$

We may now use the best reply function to allow the aircraft designer to anticipate N_{air} by using the relation in equation (2-21). In doing so, we have now directly incorporated information that previously only affected the airline, such as fuel cost and demand for air travel, directly into the optimization formulation for the other stakeholders.

Problem Analysis

Now that we have defined the formulation of the design problem and the profit functions for each of our stakeholders, we can calculate the profit maximizing solution for the aircraft designer who is anticipating the reaction of the airline. To do this, we must first find reasonable estimates for some of the coefficients present in the design and profit functions. Table 2-2 provides a summary of these coefficients, their assumed values or a range of values; these values and ranges are estimated based on various sources (Aircraft characteristics [26] [27] [36] [37] Design requirements [38] [39] [40] Fuel prices [29] Ticket prices [41] Manufacturer costs [42] Airline costs [31] [32] [35] Quarterly profits [43] [44] [45] [46]).

Note that for eight of these variables, we have assumed a range of values. This is either due to uncertainty in the true values (e.g. fixed operating cost), or actual randomness in the true values (e.g. yield strength.) To understand the effect of these variations, we perform a case study in which we take each of these uncertainties as an interval variable. We sample from these uncertain ranges in order to understand the effect of changes in these values on the optimal decisions made by the designer and airline. Even though the change in profits may be significant, if the optimal decisions are relatively constant with respect to variation in one of these coefficients, it will be reasonable to neglect it. We search over each of the interval values in order to find the minimum and maximum value of each of the four decisions variables between both stakeholders for the interaction model given in the previous section; these results are shown in Table 2-3.

We see that there is a large variation in the optimal values of each of the decisions, except for number of tests. We also find that for some combinations of variables, the airline will elect not to fly at all meaning that for some combinations of parameters it is impossible for the airline to be profitable. The primary change for the designer comes from the aspect ratio, the optimal value of which varies completely between the upper and lower bounds specified in the optimization problem. In order to understand how much of this variation in optimal decisions is due to economic uncertainty, we can compare to the case where each economic variability is fixed and only variability in material properties remains. The results of this analysis may also be seen in Table 2-3. It can clearly be seen that the addition of economic uncertainties has a significant effect on the optimal decisions, as the range of aspect ratio and safety factor are much larger with economic variability included. This change is due to the fact that the optimal number of tickets sold by the airline varies greatly, as one might expect when market conditions change. The designer then changes the type of aircraft they will construct based on the quantity they expect to sell, constructing cheap, low performance aircraft in poor market conditions and high performance high cost aircraft in favorable markets.

To consider asymmetric information, we look at the relationship between error in the designer's estimate of airline fixed costs and profits for both stakeholders across all cases of our other uncertainties. We find an interesting result; in some cases, both stakeholders benefit from the designer underestimating the airline's fixed costs and the designer always benefits from this underestimation. When the designer believes the airline fixed costs are low, they will build a more efficient and more expensive aircraft, as they believe airlines will have higher profit margins for the same number of tickets and may be willing to pay more for aircraft that can reduce fuel consumption. In the appropriate market conditions, (high fuel cost, high demand, or low fixed costs) the net effect of this change will be a reduction in the cost per flight due to increased fuel efficiency, leading the airline to sell more tickets at lower prices (and in turn buying more aircraft). This can be seen graphically in Figure 2-7 where the relative change in profits is presented for a variety of market conditions represented by each individual line.

The reason that this can happen is that the designer's decision based on the best reply function is not guaranteed to be Pareto optimal for either player. Were the designer allowed to change the design after learning the true number of aircraft purchased by the airline, they would choose a different design which would provide even higher profits for the designer, but lower profits for the airline. This is a phenomenon known as double marginalization, where two firms each add some profit margin to the price of a good, in this case air travel. The net effect of this double marginalization is actually a reduction in profits for both firms. When the designer has error in assessing airline fixed prices, they essentially reduce their own profit margin, and the benefits of this action are passed to the airline. In this case, the designer benefits as well due to the increase in aircraft sales.

Finally, we consider an example that demonstrates some unexpected results of this study. It is commonly known among aircraft designers that fuel prices are an important consideration, and that aspect ratio can provide a trade-off between fuel consumption and increased weight (and therefore increasing aircraft purchase price). Another important consideration that emerges from this work is the effect of consumer demand for air travel.

We consider that the aircraft designer has already designed the optimal aircraft at the current level of demand for air travel where the demand curve intercept is $375; this design is shown in Table 2-4, Case 1. After some time, the demand for air travel drops such that the new intercept is $250. The aircraft designer can now choose to keep the same aircraft and update the price, or design a brand new aircraft; these are shown in case 2A and 2B, respectively.

We observe a significant change in the optimal aspect ratio and safety factor for the redesign case, which also provides more than 60% greater aircraft design profits as compared to using the same design. The airline also more than doubles their profits while selling nearly50 million more tickets each year. This occurs because the lower demand causes the airline to be more sensitive to aircraft prices and to buy fewer aircraft. Taking this into account, the designer uses a lower aspect ratio and safety factor, sacrificing fuel efficiency and reliability for reduced cost. Without incorporating the effects of changes in demand into the optimization framework, an aircraft designer would not be aware of this potential change and might lose profits as a result.

Effect of Forecast Uncertainty

In commercial aviation, as with many design problems, decision makers are often faced with high levels of volatility in future market conditions. In order to understand the impact of these forecast uncertainties, we consider three sources of forecast uncertainty in our aircraft wing design problem: fuel price, demand for air travel (demand curve intercept), and sensitivity of demand to price (demand curve slope). Designers will have an estimated distribution for each of these variables during the conceptual design phase, as compared to the previous example where we considered the effect of changes in these variables which were known to the designer. For this case study, we ignore variability in material properties and the effect of asymmetric information.

For a proposed design concept, the designer will simulate possible future market conditions, airlines' actions in this market, and resulting profits for the designer. After considering all possible futures, the designer will utilize a utility transformation which incorporates the designer's attitude towards risk; in practice this functions as a robust design optimization problem.

This example seeks to quantify the effect of various possible forecasts, as well as the designer's attitude toward risk, and their effect on optimal design choices. If these design choices are highly sensitive to market uncertainty, designers would be well served to make additional effort to reduce this uncertainty by gathering extra information or by implementing a more flexible design process that may adapt as market conditions are revealed.

Forecast uncertainty problem definition

As mentioned above, we will consider three main sources of forecast uncertainty: fuel price C_{fuel}, demand for air travel P_{max}, and sensitivity of demand to price P_s. The distribution types and range of possible parameters for each of these variables is presented in Table 2-5. We will consider various cases of market forecasts based on the defined intervals for distribution parameters. Note that these intervals are not epistemic uncertainty, but are simply used to create different potential forecasts faced by the designer for our case study. Different forecasts might exist due to varying opinions of experts and analysts, or because one designer may have more information than another. By considering various sets of distribution parameters, we will see how changing forecast expectation and volatility result in various design choices.

To begin our study, we select parameter values from the intervals in Table 2-5; the values of these parameters are then known by the designer. For a given design concept (aspect ratio, safety factor, and number of tests), the designer will simulate possible future markets by sampling from the distributions described by those selected parameters and calculate their profits by combining equations (2-18) - (2-23). The designer will balance maximizing expected profits and minimizing risk in all possible futures through use of a utility transformation. This works similarly to a robust optimization problem.

For this example, we utilize an exponential utility function,

$$\begin{aligned} U &= 1 - e^{-\alpha\Pi} \qquad && \alpha \neq 0 \\ U &= \Pi && \alpha = 0 \end{aligned} \tag{2-24}$$

This utility U is a random variable since the profits Π are random. The designer will then maximize the expectation of this utility to arrive at a final design. This is one of the simplest possible utility functions, as it introduces a single parameter to define the designer’s risk aversion, α. This is also the coefficient of absolute risk aversion as defined by the Arrow-Pratt measure [25],

$$A(\Pi) - \frac{U(\Pi)''}{U(\Pi)'} = \alpha \tag{2-25}$$

This implies that the designer’s desire to avoid risk does not change with the relative size of the potential profit or loss. A larger value of α indicates a designer is more conservative, while a negative α implies a designer who is risk seeking. When α is equal to zero, the designer is risk neutral and maximizing expected utility becomes equivalent to maximizing expected profits.

A simple method by which risk aversion can be explained is the idea of a certainty equivalent. Consider a lottery with two equally likely potential outcomes, \$0 or \$200. For a risk neutral individual, the certainty equivalent is equal to the expected payout of the lottery, \$100.

This is also the price the individual would be willing to pay for a ticket to this lottery. A person who is risk averse would only buy a ticket if the price were strictly less than $100, e.g. when $\alpha = 0.01$, the certainty equivalent is roughly $57.

We may therefore consider different designer attitudes towards risk by adjusting the α parameter and maximizing expected utility. We consider a range of values from [0.01, 100], where profits Π are given in billions of dollars.

Forecast uncertainty example

In order to study this problem, we generate 128 samples of parameter values from the 6 intervals listed in Table 2-5, as well as the interval for the designer risk aversion coefficient. This provides us with a 2 degree full factorial design in which we consider a high and low value for each parameter. These values are selected to replicate the levels of uncertainty utilized in the previous example described in Table 2-2 which we consider possible variability for the US airline market. The goal of this study is to see how the type of forecast affects design decisions; that is whether we have a positive or negative expectation and with low or high uncertainty. A flowchart of the case study process is shown in Figure 2-8.

In each case, we carry out the design optimization specified in equation (2-26), where $E(U)$ is the expected value of the utility function. 1000 possible futures are generated by Monte Carlo simulation from the forecast distributions specified and used to calculate the expected value of the utility function. The design decisions of aspect ratio, safety factor, and number of tests are varied and expected utility is maximized using a standard gradient based optimizer. The optimal design decisions for aspect ratio, safety factor, and number of tests are recorded and compared for each case. Table 2-6 provides the ranges of optimal decisions for the various possible forecasts. Due to finite Monte Carlo samples, 90th percentile intervals are used to describe optimal decision values over the range of possible forecasts.

$$\max_{AR,SF,N_{test}} E(U)$$
$$where\ U = 1 - e^{-\alpha\Pi_{\mathrm{d}}}$$
$$\Pi_{\mathrm{d}} = \mathrm{N_{air}P_{air} - N_{test}C_{test} + P_{pen}C_{pen} - C_{ini}}$$
$$N_{tix}^{*} = \frac{1}{2N_{flight}N_{pax}P_s}(N_{flight}N_{pax}P_{max} - C_{fix} - P_{air} - C_{fuel}N_{flight}FC - VSL * N_{pax}PF) \tag{2-26}$$
$$N_{air} = \frac{N_{tix}^{*}}{5 * 365N_{pax}}$$

Clearly, the impact of changing forecasts has a significant effect on the optimal decisions a designer will choose. Additionally, these ranges are even larger than those calculated in our previous work, where markets were considered to be known. Extreme values of aspect ratio relate strongly to mean fuel price and demand, with high aspect ratios when both are low or both are high and low aspect ratios when mean fuel price and demand are one low and one high. Low safety factors are common in many cases, but high safety factors are brought about by high forecast uncertainties and high mean fuel price. Low number of tests occurs in multiple conditions, with high number of tests resulting from low mean forecasts on all variables.

We may additionally determine the effect of the parameters of each distribution on the design choices that are made. For example, if a designer's optimal decisions are highly sensitive to the level of volatility in the fuel price forecast, it might be advantageous for the designer to attempt to either gather more data or attempt to delay their decision in order to reduce this uncertainty as much as possible. Table 2-7 - Table 2-9 present both the mean and standard deviation of optimal design decisions when facing low or high uncertainty as measured by the variance of fuel price, demand, and price sensitivity, respectively where the low and high values represent the extreme cases for variance in Table 2-5. In each case, all other parameters are varied over the ranges described in Table 2-5.

This simple test indicates that the designer's level of forecast uncertainty in market conditions can have a significant impact on their design decisions. For example, our results indicate that high volatility for ticket price sensitivity results in a higher optimal aspect ratio. This is due to the fact that when price sensitivity is high, airlines must charge low ticket prices and need to operate efficient aircraft to reduce their costs. By waiting some time for a better estimate of demand sensitivity, a designer might find that such a high performance, high cost design is not required.

High forecast uncertainty in demand results in a higher aspect ratio, safety factor, and number of tests resulting in an overall more costly aircraft. High uncertainty in future fuel prices results in higher aspect ratios, lower safety factors and increased testing in order to improve performance in anticipation of potentially higher fuel prices. We also note that high uncertainty in fuel price leads to low variability in optimal aspect ratio, suggesting the aspect ratio is driven by the worst case (highest) fuel price considered. Overall, these results indicate design decisions may change significantly based on the level of forecast uncertainty. This type of analysis may be able to uncover situations in which designers can benefit from delaying decisions until they are able to reduce forecast uncertainty.

Discussion of Results

An example problem has been put forward that uses a basic multidisciplinary design problem and a simple model of economic interaction between stakeholders to investigate the relative importance of engineering and economic uncertainty on design decisions and outcomes. We have used a basic model of a wing structure and aerodynamics and simple expressions to describe profit functions for aircraft designers and airlines. Interactions between these stakeholders are modeled using game theory, where we have a sequential game with the aircraft designer moving first and asymmetric information regarding the airline's profit function.

We conduct a case study in which we vary the values of eight important model inputs that are likely to be subject to variability or uncertainty. The range of optimal decision sets across all of these cases is computed and indicates that changes in market conditions can have a large impact in these decision values. We find that variabilities related to market conditions and stakeholder profit functions have a much greater impact on design decisions and outcomes than traditional design variabilities such as material properties. This finding indicates that understanding customer preferences and market variability is as much if not more important than understanding uncertainty in design parameters and operating conditions. Additionally, we show that designers acting with errors or limited information may actually produce a more profitable design for both the airline and the designer.

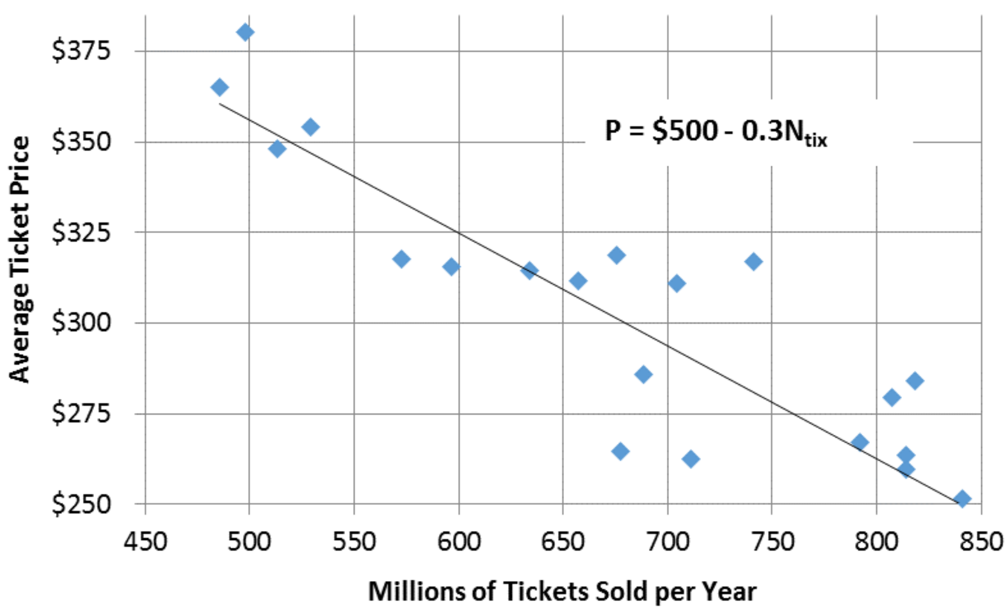


Figure 2-1. Historical ticket price vs quantity sold [41]

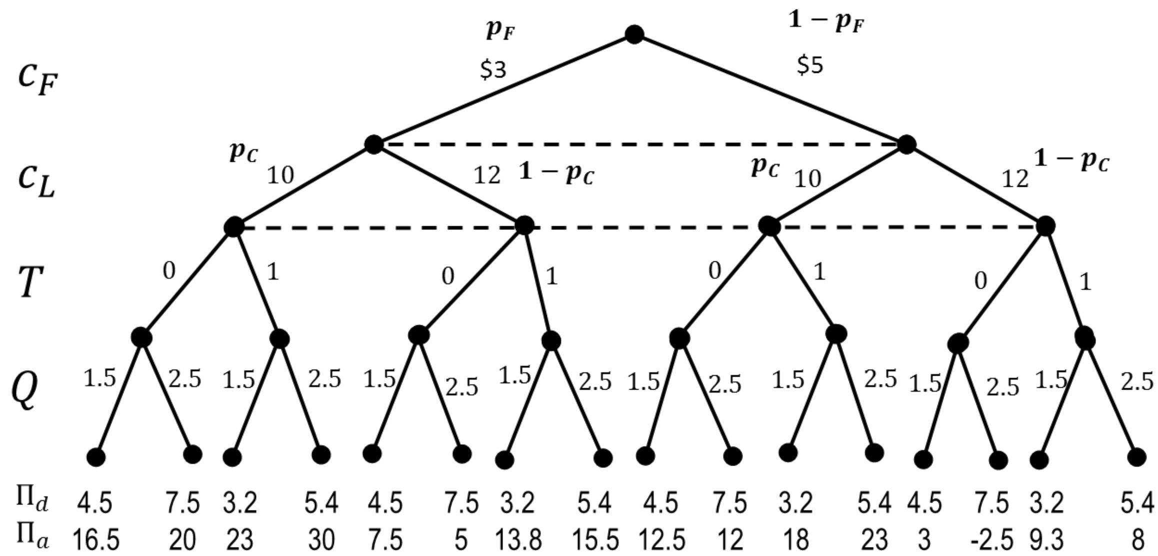


Figure 2-2. Extensive form game with uncertainty in fuel prices specified by probability p_F and in fixed cost specified by probability p_C. Designers choose technology T and airlines choose quantity of flights Q with payoffs for the designer and the airline, respectively

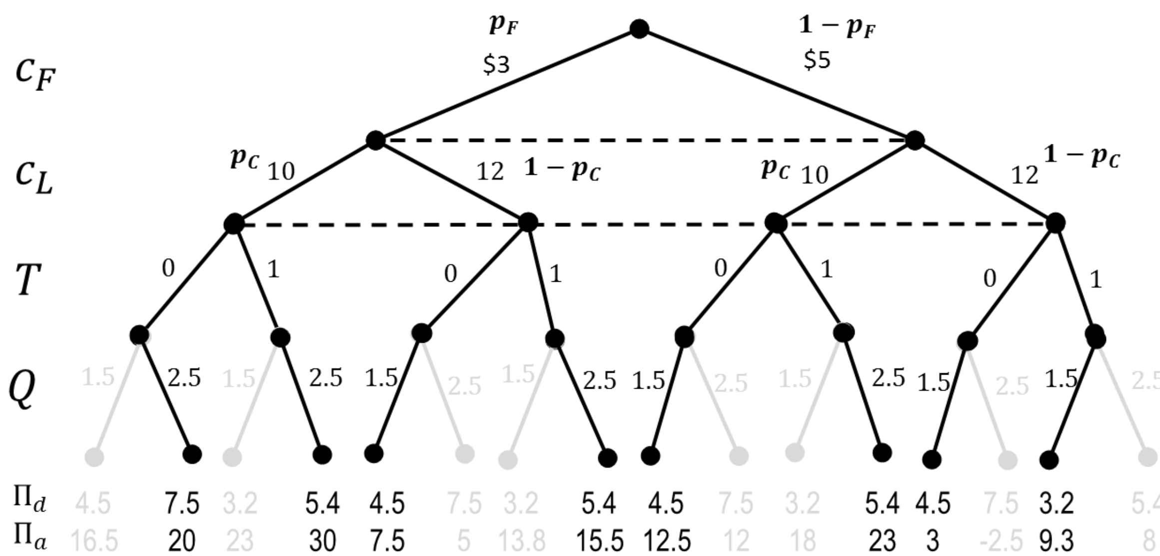


Figure 2-3. Backwards induction indicating strictly dominated choices (gray) for the airline when choosing quantity Q

$(p_F)c_F = 3, (p_C)c_L = 10$

Q \ T	0	1
1.5	16.5, $\underline{4.5}$	23, 3.2
2.5	$\underline{20}$, $\underline{7.5}$	$\underline{30}$, 5.4

$(p_F)c_F = 3, (1 - p_C)c_L = 12$

Q \ T	0	1
1.5	$\underline{7.5}$, $\underline{4.5}$	13.8, 3.2
2.5	5, $\underline{7.5}$	$\underline{15.5}$, 5.4

$(1 - p_F)c_F = 5, (p_C)c_L = 10$

Q \ T	0	1
1.5	$\underline{12.5}$, $\underline{4.5}$	18, 3.2
2.5	12, $\underline{7.5}$	$\underline{23}$, 5.4

$(1 - p_F)c_F = 5, (1 - p_C)c_L = 12$

Q \ T	0	1
1.5	$\underline{3}$, $\underline{4.5}$	$\underline{9.3}$, 3.2
2.5	-2.5, $\underline{7.5}$	8, 5.4

Figure 2-4. Simultaneous game solution

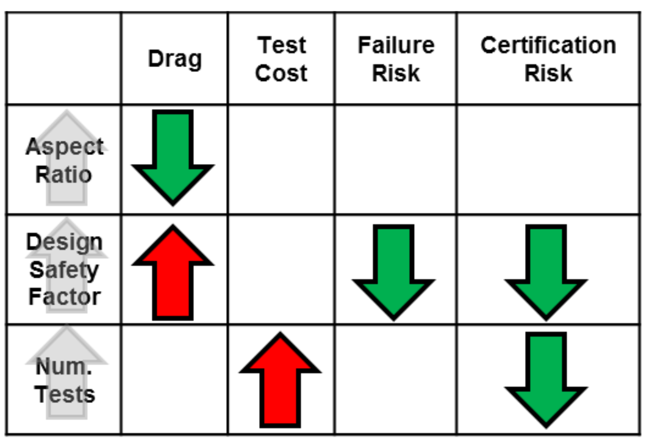


Figure 2-5. Design problem trade-off matrix

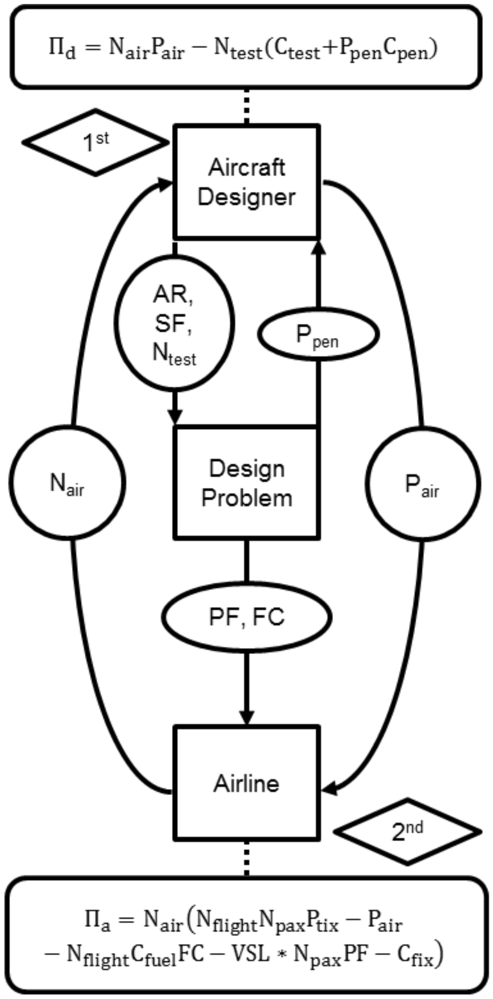


Figure 2-6. Complete interactions between stakeholders and design

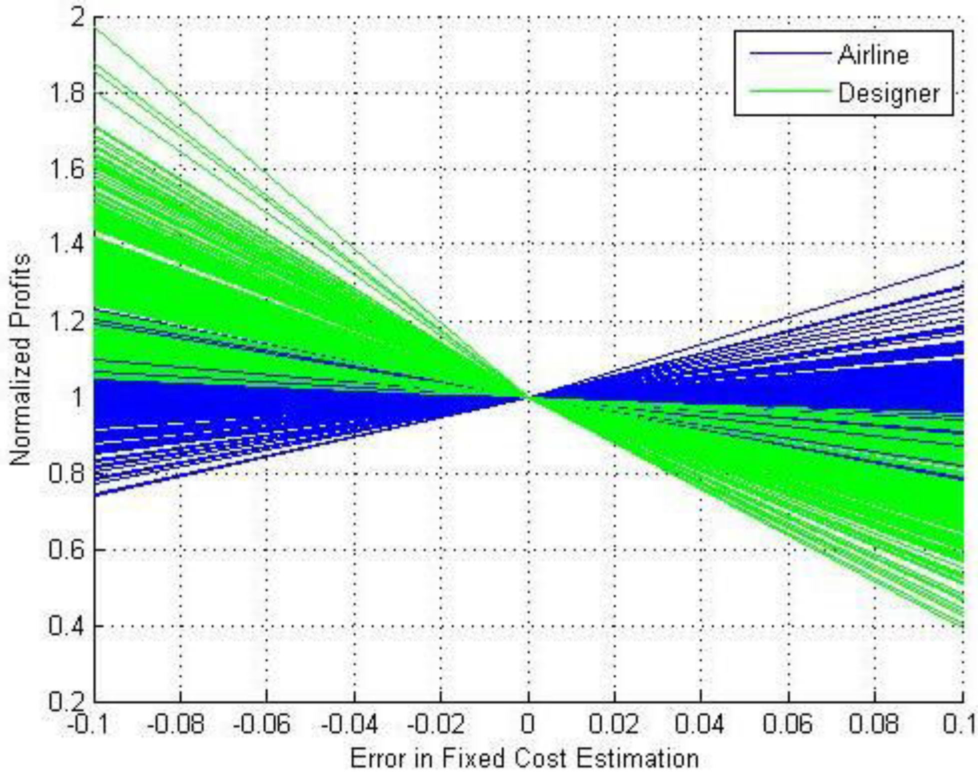


Figure 2-7. Effect of information asymmetry on airline and designer profits

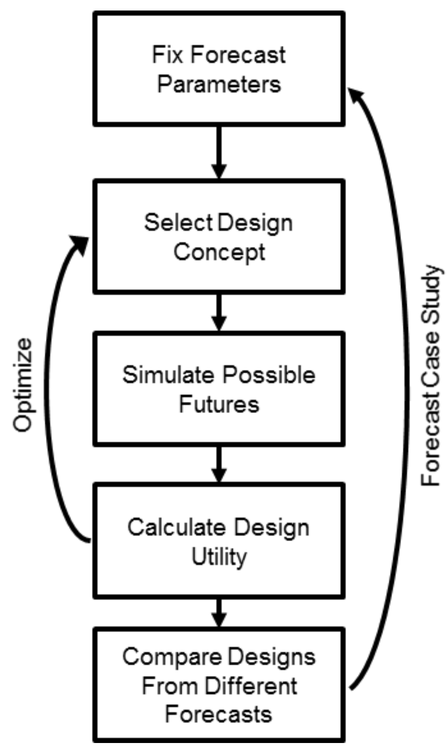


Figure 2-8. Forecast uncertainty case study process

Table 2-1. Solution values for sequential game with no uncertainty

c_F	c_L	Q^*	T^*	Π_d	Π_a
\$3.00	10	2.58M	0 (-0.08)	\$7.75B	\$20.0B
\$3.00	12	2.02M	0.63	\$4.99B	\$12.32B
\$5.00	10	2.28M	0.27	\$6.30B	\$15.55B
\$5.00	12	1.78M	0.99	\$3.83B	\$9.57B

Table 2-2. Fixed coefficient values

Coefficient	Value
Half-span	17.16 m
Taper ratio	0.159
Required lift	80,000 kg
Zero lift drag coefficient	0.01
Planform area	68 m^2
Maximum ticket price, P_{max}	\$400 – \$600
Ticket demand slope, P_s	0.2 – 0.3 \$ per million tickets
Passengers per flight, N_{pax}	100
Test cost, C_{test}	\$1M
Initial design cost, C_{ini}	\$30B
Penalty cost, C_{re}	\$500M – \$5B
Fuel cost, C_{fuel}	\$0.53 – \$1.33 per liter
Lifetime fixed operating cost, C_{fix}	\$780M – \$900M
Value of statistical life, VSL	\$9.1M
Yield stress	450 MPa – 550 MPa
Elastic Modulus	65 GPa – 80 GPa
Thrust specific fuel consumption	0.06 $^{kg}/_{N \cdot hr}$
Nominal wing weight	5,000 kg
Nominal wing volume	2.0 m^3
Test limit load	3 g
Critical deflection/span	0.25
Designer error in C_{fix}	-10% – 10%

Table 2-3. Range of optimal decisions

Decision Variable	Material Properties Only	Including Economic Variables
Aspect Ratio	11.5 – 14.0	7.5 – 14.0
Safety Factor	1.17 – 1.30	1.17 – 2
Number of Tests	10	9 – 11
Number of Tickets	765M – 772M	0 – 3.2B

Table 2-4. Demand shift case study

Parameter	Case 1[1]	Case 2A[2]	Case 2B[3]
Demand Curve Intercept	\$375	\$250	\$250
Aircraft Designer Profits	\$8.08B per year	\$0.8B per year	\$1.36B per year
Airline Profits	\$28.4B per year	\$1.7B per year	\$4.60B per year
Aspect Ratio	12.28	12.28	11.68
Design Safety Factor	1.38	1.38	1.27
Number of Tests	11	11	10
Aircraft Price	\$71.6M	\$49.3M	\$41.7M
Number of Tickets Sold	840M per year	291M per year	339M per year

[1] Initial optimal design with initial demand level
[2] Initial design at new demand with updated price
[3] New optimal design at new demand level

Table 2-5. Definition of forecast uncertainty

Variable	Distribution Type	Parameter Range
Demand for Air Travel P_{max} (\$)	Uniform	μ ~ [400, 600] $\pm$ ~ [25, 100]
Sensitivity of Demand to Price P_s (\$/million tickets)	Uniform	μ ~ [0.2, 0.3] $\pm$ ~ [0.025, 0.05]
Fuel Price C_{fuel} (\$/liter)	Lognormal*	$E(C_{fuel})$ ~ [0.50, 3.50] $V(C_{fuel})$ ~ [0.3, 1.0]

*Fuel prices 10th – 90th percentiles from roughly \$0.50 to \$5.00 per liter

Table 2-6. Optimal decisions with forecast uncertainty (10^{th} – 90^{th} Percentile Values)

Aspect Ratio	Safety Factor	Number of Tests
7.1 – 12.7	1.5 – 6.1	8 – 25

Table 2-7. Effect of Fuel Price Forecast (C_{fuel}) Uncertainty on Design Decisions

Low Uncertainty			High Uncertainty		
Aspect Ratio	Safety Factor	Number of Tests	Aspect Ratio	Safety Factor	Number of Tests
μ: 9.06	μ: 3.36	μ: 11	μ: 9.79	μ: 2.65	μ: 13
σ: 1.66	σ: 2.63	σ: 4.5	σ: 1.56	σ: 1.74	σ: 13.7

Table 2-8. Effect of Demand (P_{max}) Uncertainty on Design Decisions

Low Uncertainty			High Uncertainty		
Aspect Ratio	Safety Factor	Number of Tests	Aspect Ratio	Safety Factor	Number of Tests
μ: 9.40	μ: 2.69	μ: 11	μ: 9.63	μ: 3.13	μ: 14
σ: 1.62	σ: 1.82	σ: 6.6	σ: 1.64	σ: 2.38	σ: 14.2

Table 2-9. Effect of Ticket Price Sensitivity (P_s) Uncertainty on Design Decisions

Low Uncertainty			High Uncertainty		
Aspect Ratio	Safety Factor	Number of Tests	Aspect Ratio	Safety Factor	Number of Tests
μ: 9.34	μ: 2.86	μ: 14	μ: 9.70	μ: 2.97	μ: 11
σ: 1.47	σ: 2.17	σ: 15.1	σ: 1.77	σ: 2.10	σ: 4.7

CHAPTER 3
QUANTIFYING STAKEHOLDER INTERACTIONS FROM OBSERVATIONAL DATA USING CAUSAL MODELS

French Chapter Summary

Comprendre et prévoir les relations entre différentes variables est d'un intérêt majeur dans de nombreux domaines tels que l'économie, la santé ou l'ingénierie. Dans le cadre de cette thèse de tels modèles permettent de rendre compte des interactions entre différentes partiers prenantes. Les modèles causaux sont un outil qui permet d'estimer les relations de cause à effet entre variables allant au délà de simples effets de corrélation tels que ceux dont rends compte la régression. Les modèles causaux sont généralement représentés par les réseaux bayésiens, ou graphiquement comme un graphe acyclique orienté. Nous examinons dans ce chapitre diverses méthodes d'estimation de ces modèles causaux à partir des données observées et nous développons une nouvelle approche pour la construction de modèles causaux sur des séries chronologiques. Une métrique est également développé pour quantifier le niveau de confiance dans un modèle ajusté sur la base du nombre de fois que la même structure de modèle est sélectionné par un algorithme de re-échantillonnage; nous dénommons cette métrique la robustesse du modèle. Nous montrons par des exemples numériques que le niveau de robustesse est un indicateur utile de la précision d'un modèle donné; que le taux de récupération du modèle correct et la précision des paramètres estimés est améliorée lorsque la métrique de robustesse est élevée. En outre, le calcul de la robustesse fournit des estimations de l'erreur standard pour les paramètres des modèles estimation qui est représentative de l'incertitude des paramètres. Ces approches sont d'abord testées sur des exemples simples puis appliquées à deux bases de données réelles : la relation entre salaires et prix dans l'économie américaine d'un part, le marché de l'aviation commerciale aux États-Unis d'autre part. Dans les deux cas le modèle construit a une robustesse relativement faible dû au manque de données plus fréquentes, mais les modèles permettent néanmoins de quantifier des grandes tendances dans ces interactions.

Overview

Understanding relationships between multiple variables and predicting the effects of manipulations of variables are of key interest in many fields such as economics, health sciences, and engineering. Causal models are a tool that allows for estimation of the cause and effect relationships between variables that are not described by correlation based methods such as regression. Pearl [11] provides a thorough overview of the concepts and challenges of causality and identifying causal models. Causal models are generally represented by Bayesian networks, or graphically as a Directed Acyclic Graph (DAG) as shown in Figure 3-1 where f represents a probability density function. In this case, conditional dependencies between variables in the network, represented by directed arrows in the DAG, can be interpreted as causal effects between variables.

Joint Normal Causal Model Definition

We will focus on the case of joint normally distributed data with linear causal effects, such that for multivariate time series data with n variables and T observations as shown in equation (3-1), we may represent a causal model as the vector autoregressive process for any arbitrary time t as defined in equation (3-2).

$$x_t = \begin{bmatrix} x_{1t} \\ \vdots \\ x_{nt} \end{bmatrix} \; for \; t = 1, \dots, T \qquad (3\text{-}1)$$

$$A_0^T x_t + A_1^T x_{t-1} + \cdots + A_p^T x_{t-p} = \epsilon_t \qquad (3\text{-}2)$$

where A_0 represents a matrix of the contemporaneous causal effects, A_k represents a matrix of the temporal causal effects at time lag $k = 1, \dots, p$, respectively, and ϵ_t represents the residual error in the observed data at any arbitrary time t. Note that to avoid cycles in the model, there must exist a permutation of the matrix A_0 such that the matrix becomes triangular with ones along the main diagonal.

We may consider the matrix form of the causal model described in equation (3-2) as shown in equation (3-3).

$$A^T X_{t,p} = \begin{bmatrix} A_0^T & 0 & \cdots & 0 \\ A_1^T & A_0^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ A_p^T & \cdots & A_1^T & A_0^T \end{bmatrix} X_{t,p} = \begin{pmatrix} \epsilon_{t-p} \\ \vdots \\ \epsilon_{t-1} \\ \epsilon_t \end{pmatrix} for\ p+1 \leq t \leq T \qquad (3\text{-}3)$$

where $X_{t,p}$ is defined as a vector of the observed data as described in equation (3-4).

$$X_{t,p} = \begin{bmatrix} x_{t-p} \\ \vdots \\ x_t \end{bmatrix} for\ p+1 \leq t \leq T \qquad (3\text{-}4)$$

Assuming we have a stationary process, meaning that the mean and covariance of the data do not change over time, the causal relationships for any vector $X_{t,p}, p+1 \leq t \leq T$ will be equivalent.

Methods for Fitting Causal Models

Causal effects are often estimated through use of experiments where causal hypotheses can be directly tested [11]. However in many cases, experimentally manipulating variables is not feasible; for these cases many methods have been developed for learning causal models based on observational data. Methods for fitting causal models include the Peter and Clark (PC) algorithm [14], the Spirtes-Glymour-Scheines (SGS) algorithm [15], the Inductive Causation (IC) algorithm [16], and the Sparsest Permutation (SP) algorithm [17] for non-temporal data and methods such as the Time Series Causal Model (TSCM) algorithm [18] for temporal data.

Time Series Causal Model (TSCM) Method

One simple score based criteria for fitting causal models is proposed by Chen and Chihying [18]; we refer to this method as the time series causal model (TSCM) method. The TSCM approach utilizes stepwise process. First, an unconstrained (fully connected) vector autoregression (VAR) model is fit via a multivariate least squares regression to the observed data in order to estimate model residuals, ϵ. These model residuals are used to "learn" the structure of the contemporaneous effect matrix A_0, where the coefficients within the matrix are calculated using a linear recursive simultaneous equation model (SEM),

$$x_j = \sum_{k=1}^{j-1} a_{jk} x_k + \varepsilon_j \tag{3-5}$$

or in equivalent matrix form,

$$A_0 X = \varepsilon \tag{3-6}$$

For a given model structure, i.e. the active coefficients a_{jk} in a model, the matrix A_0 is assigned a score using the Bayesian information criterion (BIC) as shown in equation (3-7).

$$BIC = \sum_i log\big(\varphi(\varepsilon_i)\big) - \frac{1}{2}\log(n_{obs})\, df \tag{3-7}$$

where φ is the standard joint normal probability density function, ε_i is the model residual at the ith observation for the given matrix as calculated in equation (3-6), n_{obs} is the number of observations, and *df* is the degrees of freedom, or the combined number of nonzero elements in each A matrix.

The matrix A_0 is then optimized by starting with a random model structure and searching over all nearest neighbors by adding, removing, or reversing ($a_{jk} \rightarrow a_{kj}$) one connection. The neighbor with the highest BIC score then becomes the updated model, and this process is repeated until no improvement is possible or some iteration limit is reached. This can be done with multiple random starting structures in order to improve the chance of finding the global optimum model structure.

In the next step, we consider the temporal causal effects in matrices A_1 to A_p as described in equation (3-2). The coefficients in these matrices are fit using a vector autoregression (VAR) model to the estimated model residuals updated for the effect of A_0 estimated in the previous step. This fit is carried out by solving a multivariate least squares regression for the expression shown in equation (3-8).

$$A_1^T x_{t-1} + \cdots + A_p^T x_{t-p} = \epsilon_t - A_0^T x_t \qquad (3\text{-}8)$$

As was done with the A_0 matrix, we simultaneously optimize each matrix A_1 to A_p by searching over all nearest neighbors of a random initial model structure and assigning each a score based on the BIC. We may then combine the results of both steps to represent the entire causal model as described in equation (3-2).

Sparsest Permutation Method

The sparsest permutation method, developed by Raskutti and Uhler [17] is another potential method to fit causal models. We find this method to be computationally efficient when applied to time series data with small numbers of variables, which makes it well suited for numerical simulation with many repetitions. The description of this algorithm as presented in Rakutti and Uhler [17] (p. 5) is:

> Let $S(G)$ denote the skeleton of a DAG G and $|G|$ the number of edges in G (or $S(G)$). Then the [sparsest permutation] algorithm is defined as follows:
> (1) For all permutations π of the vertices $\{1,2,\ldots,[n]\}$ construct [described below] G_π and let $s_\pi = |G_\pi|$.
> (2) Choose the set of permutations $\{\pi^*\}$ for which s_{π^*} is minimal amongst all permutations.
> (3) Output G_{π^*} for all π^* such that s_{π^*} is minimal amongst all permutations.

The model G is constructed by checking the conditional correlation between each variable. The edges of each DAG permutation are only retained when they satisfy the Markov condition; that is the conditional correlation of the connected variables given the preceding variables in the permutation is statistically non-zero. Therefore a directed edge between any two variables j and k for $j < k$ is retained if and only if the non-independence condition in equation (3-9) is satisfied at some user specified confidence level α.

$$X_{\pi(j)} \not\perp X_{\pi(k)} | X_S \ where\, S = \{\pi(1), \pi(2), \ldots, \pi(k-1)\} \backslash \{\pi(j)\} \tag{3-9}$$

Raskutti and Uhler [17] show that for the case of joint normally distributed variables for non-time series data, this process can be simplified to the Cholesky decomposition of the inverse covariance matrix of our observations. In order to apply the sparsest permutation algorithm to time series data, we require some additional modifications to the method.

For the case of joint normal variables with linear causal effects for non-time series data, we can represent a DAG structure as a linear recursive simultaneous equation model (SEM) for each ith observation as

$$x_{ji} = \sum_{k=1}^{j-1} a_{jk} x_{ki} + \varepsilon_{ji} \ \ for\, j = 1, \ldots, n \tag{3-10}$$

or in equivalent matrix form,

$$A_0^T x_i = \varepsilon_i \tag{3-11}$$

where A_0 is an upper triangular matrix with ones on the diagonal of the form

$$A_0 = \begin{bmatrix} 1 & -a_{12} & \cdots & -a_{1n} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{(n-1)n} \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{3-12}$$

Because of the assumption of joint normality, we know that ε will be normally distributed and independent, such that

$$A_0^T x_i \sim N(0, D_0) \tag{3-13}$$

where D_0 is the diagonal covariance matrix of ε.

We may solve this expression in equation (3-13) for x_i as

$$x_i \sim N(0, \Sigma) = N(0, (A_0 D_0^{-1} A_0^T)^{-1}) \tag{3-14}$$

From this we may see that the covariance of the observed data, Σ, is enough to determine the matrix of causal model coefficients A_0 by using the Cholesky decomposition. By calculating the Cholesky decomposition of every permutation of Σ and enforcing sparsity using equation (3-9), we may find the model (or set of models) which are the most sparse, meaning the most parsimonious representation of the causal model. Raskutti and Uhler show that this approach is valid under strictly weaker conditions than traditional score based model fitting approaches [17].

For the case of time series data, we are interested in inferring causal effects that may take place across several time periods. Under similar assumptions of normality and linear causal effects, these time series effects may be represented using a vector autoregression (VAR) model with a lag length p, representing the number of prior periods for which causal effects may exist. This idea was described previously by Chen and Chihying [18]. This VAR model may again be expressed as

$$A_0^T x_t + A_1^T x_{t-1} + \cdots + A_p^T x_{t-p} = \epsilon_t \tag{3-15}$$

where x_t is a vector of one observation of each variable at time t. Assuming we have a stable process such that these effects do not vary over time, we can again rewrite this in a matrix form for the complete time series as previously shown in equation (3-3).

$$A^T X_{t,p} = \begin{bmatrix} A_0^T & 0 & \cdots & 0 \\ A_1^T & A_0^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ A_p^T & \cdots & A_1^T & A_0^T \end{bmatrix} X_{t,p} = \begin{pmatrix} \epsilon_{t-p} \\ \vdots \\ \epsilon_{t-1} \\ \epsilon_t \end{pmatrix} \, for\, p + 1 \leq t \leq T \tag{3-3}$$

This large matrix on the left hand side of the equation containing the causal effects can be referred to as A, and similarly the matrix describing the variance of the independent errors ϵ will be a diagonal matrix D where the diagonal elements are the variances of ϵ, D_0, repeated T times. We may therefore describe this time series similarly to the case of time independent observations as

$$X_{t,p} \sim N(0, (AD^{-1}A^T)^{-1}) \tag{3-16}$$

However, when estimating causal models from data, we often only have one observation for each variable at each time step. We are therefore unable to directly estimate the sample covariance matrix, $\hat{\Sigma} = (ADA^T)^{-1}$. However, we can determine the cross covariance, Λ, of a stationary time series up to $T - 1$ lags, and we can use this information to calculate the conditional distribution of the current time step data, x_T, given the data at the previous p lags.

The autocovariance matrix describes the relationships between variables at each time step as well as between different time steps as shown in equation (3-17).

$$\Lambda = \begin{bmatrix} L(1,1) & L(1,2) & \cdots & L(1,k) \\ L(2,1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & L(k-1,k) \\ L(k,1) & \cdots & L(k,k-1) & L(k,k) \end{bmatrix} for\ k = 1, \dots, T \tag{3-17}$$

We may estimate the elements of the autocovariance matrix using equation (3-18) as the covariance of two time shifted sets x_i and x_j of our original T observations, where for n variables, each element L will form an $n \times n$ matrix.

$$L(i,j) = cov\left(x_i, x_j\right) for\ i = 1, .. k, j = 1, \dots, k \tag{3-18}$$

Assuming we have a stationary time series process, we may utilize multiple samples of our time series with the same separation in time to estimate this covariance. Additionally, the autocovariance will not change over time and we may simplify the autocovariance as shown in equation (3-19).

$$L(i,j) = L(i-j) = L(\tau) \equiv L_\tau, \tau = i - j;\ L(i,j) = L(j,i)^T \tag{3-19}$$

Using stationarity, we may therefore simplify the matrix in equation (3-17) as a symmetric block Toeplitz (diagonal-constant) matrix as shown in equation (3-20).

$$\Lambda = \begin{bmatrix} L_0 & L_1 & \cdots & L_k \\ L_1^T & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & L_1 \\ L_k^T & \cdots & L_1^T & L_0 \end{bmatrix} for\ k = 1, \ldots, T \qquad (\ 3\text{-}20\ )$$

Using this autocovariance matrix, we may calculate the distribution of our data as the current time step, x_t conditional on the values of each x at the previous p time steps. The covariance of matrix of this conditional distribution is calculated using the Schur compliment of the autocovariance matrix as shown in equation (3-21).

$$COV(x_t | x_{t-1,\ldots,t-p}) = L_0 - [L_1 \quad \cdots \quad L_{p+1}] \begin{bmatrix} L_0 & L_1 & \cdots & L_p \\ L_1^T & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & L_1 \\ L_p^T & \cdots & L_1^T & L_0 \end{bmatrix}^{-1} \begin{bmatrix} L_1^T \\ \vdots \\ L_{p+1}^T \end{bmatrix} = \Gamma_0 \qquad (\ 3\text{-}21\ )$$

Without loss of generality, we may consider each variable of our stationary time series to have zero mean. The conditional expectation for x_t is then given as

$$E(x_t | x_{t-1,\ldots,t-p}) = [L_1 \quad \cdots \quad L_{p+1}] \begin{bmatrix} L_0 & L_1 & \cdots & L_p \\ L_1^T & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & L_1 \\ L_p^T & \cdots & L_1^T & L_0 \end{bmatrix}^{-1} \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix} = W \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix} \qquad (\ 3\text{-}22\ )$$

Combining equations (3-21) and (3-22), we may represent the distribution of x at the current time as shown in equation (3-23).

$$x_t | x_{t-1,\ldots,t-p} \sim N\left(W \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix}, \Gamma_0 \right) \qquad (\ 3\text{-}23\ )$$

By subtracting the mean term, $W x_{t-1,\ldots,t-p}$, from both sides and performing a Cholesky decomposition of Γ_0, we get

$$x_t | x_{t-1,\dots,t-p} - W \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix} = N(0, (A_0 D_0^{-1} A_0^T)^{-1}) \quad (3\text{-}24)$$

Multiplying both sides by A_0^T results in the final expression,

$$A_0^T x_t | x_{t-1,\dots,t-p} - A_0^T W \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix} = N(0, D_0) = \epsilon_t \quad (3\text{-}25)$$

where the time lagged causal effect matrices A_1 to A_p are given by

$$A_0^T W = -\begin{bmatrix} A_1^T & \cdots & A_p^T \end{bmatrix} \quad (3\text{-}26)$$

This provides a method for obtaining the matrix of causal coefficients, A, directly from the auto-covariance matrix Λ. As with the original sparsest permutation algorithm, we can search over all possible permutations of our n variables, recalculate the auto-covariance, and search for the sparsest coefficient matrix A.

Estimating Model Confidence and Parameter Uncertainty: The Robustness Metric

Using any causal model fitting method, we may always estimate causal relationships between a given dataset and produce one (or sometimes several) causal models as output. However, it then falls to experts and analysts to determine if the resulting model is reasonable and should be trusted. The available data might be insufficient to make accurate estimates of causal effects or important variables may be missing, violating the common causal sufficiency assumption. While significant work has been devoted to selecting models based on minimizing unexplained variance in the observed data [47], these approaches may not necessarily result in correctly identifying model parameters. We therefore propose a metric to be used when fitting causal models which may assist in detecting when to accept a generated causal model. We find this metric also provides useful estimates of standard errors in model parameters in the case that the model is identified correctly. Understanding when to believe a fitted causal model and the uncertainty in its parameters is critically important for decision makers attempting to utilize these models. Though this work is focused on applications with time series data, the metric also may be simplified to non-time series models as well.

Robustness Definition

The robustness metric is based on generating statistically similar datasets and refitting causal models to the newly generated data. By calculating the percentage of cases in which we observe the same model structure, we determine how robust each structure is to the inherent unexplained noise in the data.

The first step of being able to calculate model robustness is the ability to generate statistically similar data. Our ability to generate statistically similar data relies on several assumptions: that we have sufficient number of observations to be representative of the underlying process, that noise in the autoregressive process is uncorrelated, and that our time series is stationary. While these assumptions are significant, they are already required assumptions for accurate estimation of causal models and therefore generating new data does not impose additional restrictions.

For non-time series data, we would commonly use bootstrapping where we generate a new dataset by randomly resampling from our initial data with replacement. Since the ordering of observations is important for time series data, we utilize the autocovariance matrix of our initial dataset to generate new samples. As was described in the previous section on the sparsest permutation algorithm, we may manipulate the autocovariance matrix of observed data to find the conditional distribution of a single time step as given in equation (3-17).

$$x_t | x_{t-1,\dots,t-p} \sim N\left(W \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix}, \Gamma_0\right) \qquad (3\text{-}17)$$

We may use this distribution to generate new statistically similar samples to our original data by sequentially generating samples. This requires a user specified initial condition for $x_{1,\ldots,p}$, we may choose to generate more samples than required and remove some number of initially generated samples such that the effect of this user specified initial condition is negligible on our final data. We also must specify some lag length p in order to generate new data. This lag length may be chosen to be arbitrarily larger than the expected length of significant causal effects in order to ensure the validity of the newly generated data.

We may then utilize any potential time series causal model learning algorithm to estimate a causal model from many repetitions of this generated bootstrap data. We determine the robustness of a given model structure by calculating the percentage of these time series bootstrap samples which produced the same structure. Robustness may be calculated explicitly for N runs with a single structure appearing K times as shown in equation (3-27).

$$R = 100\frac{K}{N} \tag{3-27}$$

We propose that the robustness value of a model structure found using any causal model fitting method is a useful tool to determine whether that structure is equivalent to the true data-generating model structure. We may also consider the value of maximum robustness as a measure of our confidence in the resulting model.

In addition to acting as a model validation criterion, the robustness metric may be used to approximate the uncertainty in the coefficients of the causal model. For every instance of the model structure in the bootstrap sampling, the estimate of model coefficients will be slightly different due bootstrapping producing different time series realization. We may therefore calculate the standard deviation of each model coefficient and use this as a measure of the standard error of the model coefficients, which can be critically important to analyzing the model in order to make decisions.

Evaluation Criteria for Robustness Method

In order to understand the efficacy of the proposed metric, we utilize several scoring criteria for causal model results when working with synthetically generated data with a known true coefficient matrix A. First, we consider whether the model structures suggested by the robustness metric is observationally equivalent to the true data-generating structure. Observational or Markov equivalence is defined by Pearl and Verma [16] as follows:

> "Two DAGs (models) are observationally equivalent if and only if they have the same skeleton and the same set of v-structures, that is two converging arrows whose tails are not connected by an arrow"

Put simply, observational equivalence means the chain of conditional dependencies does not represent a unique DAG, and therefore may not be exactly identified from observed data. Therefore, if a model that is observationally equivalent to the true model is returned by an algorithm, we consider this a success. Note that inspection of the model that is output by an algorithm will allow for identification of whether or not a model is observationally unique. For example, consider the causal model example from Figure 3-1. The relationships between x_1, x_2, and x_3 are not observationally unique as we could also have $x_1 \rightarrow x_2$ or $x_3 \rightarrow x_2$ (but not both) without creating or removing a v-structure, meaning the joint probability distribution described by these 3 possible models is equivalent.

In order to quantify the accuracy of the estimated coefficient values, we propose another scoring criterion utilizing the Frobenius norm of the difference of the true and fitted models normalized by the Frobenius norm of the true model. In order to construct this difference, we consider the full model coefficient matrix A including contemporaneous and temporal effects as was utilized in equation (3-3). We then calculate this accuracy score as shown in equation (3-28).

$$\zeta = \frac{\left\| A_{true} - A_{fit} \right\|_F}{\left\| A_{true} \right\|_F} \qquad (\ 3\text{-}28\ )$$

This results in the normalized root sum square error of the true versus predicted coefficient values. This score provides a quantitative measure of coefficient accuracy without requiring consideration of the model structure. This allows us to account for both when a model may have the correct structure but large error in coefficients, and also when a model might have a single incorrect structural element but be otherwise very close to the true model.

Finally, we propose a scoring criterion to better understand the ability to predict the uncertainty in model coefficient values. To do this, we compute the difference between true and estimated coefficients and divide this difference by the standard deviation of coefficient values obtained for the equivalent model structure by the robustness metric. This results in a matrix of errors in the coefficient estimates normalized by the estimated coefficient uncertainty. Recalling that the coefficient matrix A will have dimension $n(p+1) \times n(p+1)$, where n is the number of variables in X and p is the number of lags, this normalized error matrix Φ is calculated as

$$\Phi = \left[A_{true} - A_{fit}\right] \circ \begin{bmatrix} 1/\sigma_{\alpha_{1,1}} & \cdots & 1/\sigma_{\alpha_{1,n(p+1)}} \\ \vdots & \ddots & \vdots \\ 1/\sigma_{\alpha_{n(p+1),1}} & \cdots & 1/\sigma_{\alpha_{n(p+1),n(p+1)}} \end{bmatrix} \qquad (3\text{-}29)$$

where $\circ$ represents the Hadamard product and $\sigma_{\alpha_{j,k}}$ is the standard deviation calculated for the $[j,k]$ non-zero element of the estimated coefficient matrix A from the robustness metric. If this standard deviation is a reliable estimator of coefficient standard error, we may check that the elements of Φ where model coefficients are non-zero should follow a standard normal distribution.

Testing Robustness with Simulated Data

To test our proposed metric, we randomly generate causal model structures A and noise levels described by the diagonal matrix D_0 as shown in equation.

$$D_0 = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n^2 \end{bmatrix} = cov(\epsilon) \qquad (3\text{-}30)$$

We then synthesize observed data for the generated model, apply the proposed metrics, and compare the resulting fitted model to the known true model by the scoring criteria proposed in the previous section. Once the A and D_0 matrices are created, we may generate any number of synthetic time series observations by drawing from the joint normal distribution specified by equation (3-31). This relationship is found directly by solving equation (3-3) for $X_{t,p}$.

$$X_{t,p} \sim N(0, (AD^{-1}A^T)^{-1}) \qquad (3\text{-}31)$$

Coefficients are selected randomly from the interval $[-1, -0.4] \cup [0.4, 1]$ such that the coefficients are bounded away from zero. In order to avoid fully connected models which will not be observationally distinguishable, we randomly set coefficients in A to zero until meeting a target connectivity ratio r defined as

$$r = \frac{m}{\frac{n(n-1)}{2} + n^2 p} \qquad (3\text{-}32)$$

where m is the number of non-zero coefficients in the generated model A and $\frac{n(n-1)}{2} + n^2 p$ is the maximum number of coefficients in an n variable, p lag model.

Noise values are selected from a lognormal distribution such that the diagonal elements of D_0 are defined as shown in equation (3-33). This distribution results in [5%, 95%] bounds for these variances from [0.31, 0.43]. Note that increasing the level of noise does not have a significant effect on the sparsest permutation algorithm or the robustness metric so long as we have sufficient data to accurately estimate the autocovariance.

$$D_{ii} \sim logN(-1{,}0.1) \qquad (3\text{-}33)$$

The stationarity of the generated time series is ensured by checking that the roots of the characteristic polynomial of the vector autoregressive process as defined in equation (3-34) lie outside the unit circle such that

$$For\ any\ z, |z| \leq 1 \Rightarrow \det\left(I - A_0^{-1}A_1 z - \cdots - A_0^{-1}A_p z^p\right) \neq 0 \qquad (3\text{-}34)$$

Due to the computational expense of generating random models that meet this stationarity condition, we limit our numerical examples to time series of 3 variables with 1 or 2 lag periods. We then perform a case study applying the method of sparsest permutation for time series with the robustness selection metric for 1 or 2 causal lags, connectivity ratios of 0.4 or 0.5, and number of observations, T, set to be 500 or 1000. For each case, 1000 random causal models are generated and tested.

First, we consider the rate of success of the robustness metric to recover the true causal structure. Table 3-1 presents the percentage of the 1000 repetitions where the metric produced successful recovery of the model structure. It can be seen that the most robust model is the correct model in at least 78% of trials, with higher success rates with increasing sample size and smaller models. Across all model size cases, the sparsest permutation algorithm sees an 88% correct identification rate, which agrees well with the results of a similar study for non-time series data from Rakutti and Uhler [17]. Note that based on using 1000 repetitions, uncertainty in these percentages is roughly 1-2% at the 90% confidence level. This may explain the higher recovery rate for $(r = 0.4, p = 2)$ as compared to $(r = 0.4, p = 1)$ with 1000 observations. The recovery rate of the sparsest permutation with the original data, not considering robustness, is included in parentheses. Note that this rate includes cases where the sparsest permutation algorithm outputs multiple equally sparse, non-equivalent results and only one is correct. Conversely, the robustness metric can always provide a single most robust structure.

We may also consider how the calculated level of robustness is related to the chance of a model being identified correctly. If we find that low robustness models are more likely to be incorrect, we may choose only to trust models with a robustness greater than some threshold value. Figure 3-2 and Figure 3-3 show histograms of the robustness level of correctly and incorrectly identified models, respectively. These figures group all lag lengths, connectivity ratios, and sample sizes together as these factors are not found to influence the relationship between structure recovery and robustness level.

We find that correctly identified models typically have a robustness level above 70-80% while incorrectly identified structures are approximately uniform across all robustness levels. This indicates we could use a threshold robustness value below which we may elect not to trust the causal model output. For example, only considering cases with a robustness of 90 or more, the identification rate improves from 88% to 97%, while at a robustness of 55% or lower the chance of correctly identifying the true model falls below 50%. The robustness threshold could be specified by an analyst or expert based on the application and required model confidence.

Next, we consider our coefficient accuracy score based on the normalized Frobenius norm of the difference in fitted and true coefficient matrices. Figure 3-4 and Figure 3-5 present box plots of the accuracy metric ζ for the most robust model across the 1000 random models for each sample size, number of lags, and connectivity ratio for models selected by the robustness metric.

We also consider the difference in accuracy score between the correctly and incorrectly identified models. We find this to be fairly consistent across different model sizes, so these results are all combined in Figure 3-6.

We find that the robustness metric is relatively consistent across all model sizes in terms of accuracy score. Accuracy score is improved slightly with sparser models, lower model lags, and increased sample size but in general will be less than 0.1. This indicates that the parameter values in the model coefficient matrix A will be identified accurately by the most robust model. We may also consider the relationship between robustness level and accuracy score. This relationship for all model sizes and numbers of samples is presented in Figure 3-7.

We observe a correlation between model robustness and accuracy score, with increasing robustness related to more accurate models. As with the structure recovery rate, this indicates that the model robustness level may be utilized as a diagnostic to determine the confidence in the accuracy of the estimated model parameters.

Finally, we calculate the error in coefficient estimation normalized by the standard deviation from the robustness metric. Because these results are consistent across each case considered, we compile all results together. Figure 3-8 presents a probability plot of the observed normalized coefficient errors for all models.

We find that this normalized error does follow a standard normal distribution, suggesting that the coefficient standard deviation is a reasonable estimate for the error in estimation of causal model coefficients. This standard deviation may therefore be used to understand possible causal effects not otherwise available in causal model fitting methods.

Application to Real Data

Wage-Price Dynamics

In order to test our method with realistic data, we attempt to reproduce the findings of Chen and Chihying [18] regarding wage price dynamics using economic data from 1965-2004. We utilize the same dataset which is provided by the authors in their previous paper [48] and is drawn from the Federal Reserve economic database (FRED) [49]. The 6 variables used in the model defined by Chen and Chihying as (w, p, e, u, z, π_m) are respectively wage inflation, price inflation, labor utilization rate, capacity utilization rate, growth of labor productivity, and inflationary climate. These variables are made stationary using the transformations described in Table 3-2.

We attempt to fit models to this data using both the sparsest permutation algorithm used in our numerical study as well as the algorithm proposed by Chen and Chihying which we refer to as the time series causal model (TSCM) method. The TSCM method utilizes a two-step process to first optimize the contemporaneous effects, A_0, using a structural equation model approximation. Temporal effects are then optimized using a vector autoregression approximation, corrected for A_0 from the initial step. In both steps, the model structure is optimized using a greedy search algorithm which searches over all nearest neighbors of a random initial model structure by adding or removing one model connection. This search is repeated with multiple random initial structures. Models are scored for optimization using the Bayesian Information Criterion (BIC). More details of the algorithm may be found in the original paper [18].

The model found by Chen and Chihying is presented in Figure 3-9. We find that we are able to reproduce this model using the TSCM method described in the original paper, while the sparsest permutation produces different results as shown in Figure 3-10. Specifically, the contemporaneous matrix A_0 has the permutation of the structure mostly reversed, while the temporal matrix A_1 more closely agrees with the TSCM result. This may be due in part to the fact that the sparsest permutation fits both A_0 and A_1 simultaneously rather than sequentially. Both methods also utilize very different scoring and search criteria which may lead to divergent results when the available data is insufficient. When adequate data is available in our numerical testing, we do find that both algorithms produce similar results. In order to reproduce model coefficients on the order of 0.02 as shown in the Chen and Chihying model, the sparsest permutation method must also use a very low confidence level when enforcing sparsity as described in equation (3-9); this has a significant effect on the model fit.

Applying the robustness metric for either the TSCM method or the sparsest permutation method with low confidence results in a model robustness of effectively zero; each bootstrap sample results in a different model structure. This implies that the available data is not sufficient to support the very small coefficient values included in the Chen and Chihying model. We can, however, increase the confidence level for enforcing sparsity until the robustness metric produces repeatable models. A similar approach may be taken with the TSCM method by modifying the penalty on model degrees of freedom in the BIC score, though this modification is less intuitive. Using a higher confidence level, the robustness metric applied to the sparsest permutation method produces the model shown in Figure 3-11 with a robustness of 29%.

This robust model suggests that only 4 coefficients from the original model are supported sufficiently by the available data. Considering that even this model has a robustness level of only 29%, one should still be hesitant to assume this model is correct. We may also consider the estimates of standard error in each coefficient obtained by the robustness metric, which are shown in Figure 3-12. Comparing the robust coefficient estimates and their standard error with the original Chen and Chihying model shows reasonable agreement. We may therefore consider that we have higher confidence in the values of these coefficients based on the available data. Overall, the high sparsity of the model compared to the Chen and Chihying model and low robustness suggest that additional data is likely needed in order to quantify some of the more complex effects governing these wage and price dynamics in question, depending on the level of confidence required by analysts. This finding agrees with previous literature such as Thompson [50] who discusses the difficulty of estimating covariance when using small sample sizes with data clustered both across time and across variables such as with the time series data considered here.

Airline Industry Data

In the previous chapter, we presented an example of an aircraft wing design considering interactions between the stakeholders of designers and airlines. However, in that case we developed simple expressions to describe the way that design choices interaction with stakeholder preferences. In reality, these interactions may be very complex and not well understood by the designer. We therefore consider relationships between the main stakeholders in commercial aviation; aircraft designers and manufacturers, airlines, and the public. We will attempt to understand how welfare for these three groups is affected by various design choices and market conditions.

Data collection

In order to model the relationships between technology and welfare, we collected 20 years of quarterly historical data on aircraft characteristics, stakeholder welfare, and other potentially influential market characteristics. To avoid the impact of seasonality in the data, we considered all of our measurements on a four quarter rolling average. For aircraft characteristics, we utilized the Bureau of Transportation TranStats database [27]. This database provides quarterly data on the type of aircraft and number of passenger miles flown dating back to 1990. Using this information, we constructed a representative aircraft as a weighted combination of all the different aircraft models utilized, where weights are proportional to the passenger miles flown by each model. We considered 11 physical characteristics of the aircraft, which are shown in Table 3-3 and are collected directly from the manufacturer aircraft specification documents [26] [51]. Note that vehicle miles represent the total number of miles travelled by the aircraft, while passenger miles are multiple by the occupancy of the aircraft during travel.

The TranStats database additionally provides quarterly profits as reported by each US airline [45] [46]. We simply summed the profits of each airline in each quarter as a measure of the overall welfare of the airline industry as a whole. Similarly, we collected quarterly reports published by the two main commercial aircraft manufacturers, Boeing [43] and Airbus [44], and used their reported profits to determine the welfare of aircraft manufacturers. For the public, we define welfare as the number of tickets purchased per capita as a measure of the accessibility and utilization of air transportation which is also obtained from the TranStats data.

Finally, we collected data on several additional factors which may be influential to one or more of the stakeholders' welfares. These include the date, price of jet fuel, average airline ticket price, listed cost of the representative aircraft, and 6 indicators of US economic health. Jet fuel prices and average ticket prices are collected from Airlines for America [41], and the economic indicators are collected from Bureau of Labor Statistics [52] [53], Institute for Supply Management [54], University of Michigan [55], Federal Reserve Bank of St. Louis [56], and the Bureau of Economic Analysts [57]. A list of the economic indicators used can be found in Table 3-3.

Factor analysis

To better understand the data we collected, we utilized factor analysis to reduce our data. This serves two purposes; first, reducing the dimensionality of our input space will reduce the computational expense of fitting a model and make the results much easier to interpret. Secondly, it is possible that multiple inputs from our collected data are all results of a common unobserved force. For instance, the multiple weights of the aircraft, the fuel and passenger capacity, and the range are all strongly correlated, and changes in all of these observed values may simply represent a shift in airline behavior from utilizing smaller aircraft flying point to point routes to large aircraft utilizing a hub and spoke schedule.

The goal of factor analysis is to construct a set of orthogonal factors as weighted combinations of our original inputs. To perform factor analysis, we must first specify the number of factors to use. One common method for determining an appropriate number of factors is based on principal component analysis, which transforms a set of inputs into an equal number of orthogonal components. The total variance of the original data will be equal to the sum of the variances of each of these components; this means we can find a reduced number of components that still explains some prescribed threshold of the variance in our original data.

In the case of our collected data, we find that four principal components are able to explain nearly 95% of the variance of the original data. Using this information, we choose to construct a four factor model using factor analysis. Factor analysis also performs rotation of the factors such that the magnitude of the weights, or factor loadings, for each input is as close to 1 or 0 as possible. This makes it easier to interpret the resulting factors, as inputs will either be included or excluded in each factor. Inputs whose factor loading magnitude are below some threshold value can be thought to have no real impact on that factor. The four factors constructed along with their factor loadings are shown in Table 3-4.

We note that each factor can be shown to represent a distinct set of inputs. Factor 1 is a collection of all of the physical characteristics of the aircraft. Based on the sign of the factor loadings, an increase the factor 1 is related to increased aircraft weight and reduced fuel efficiency. We could therefore expect that improvements in technology would contribute to decrease the value of factor 1. This factor might be explained as a representation of the fleet mixture being utilized, such as the percentage of wide-body versus narrow-body aircraft, where a higher magnitude represents more large aircraft in the fleet.

Factor 2 contains all of the inputs related to the price of fuel, where an increase in the factor value relates to an increase in fuel cost. Factor 3 contains all of the considered economic indicators, where an increase in the factor value is generally related to improved economic conditions. Finally, factor 4 is mostly related to the listed price of the aircraft utilized, along with the maximum range. An increase in factor 4 indicates an increase in aircraft cost and maximum range.

Figure 3-13 and Figure 3-14 show the value of the three stakeholder welfare metrics and the four factors, respectively. All measurements are normalized to have zero mean and unit standard deviation. We observe factor 1 has shown a general downward trend over this date range, indicating aircraft generally becoming lighter and more fuel efficient, or equivalently that airlines have shifted to more narrow-body aircraft. Factor 2 has steadily increased, meaning fuel prices have risen as well as the fuel cost per seat mile flown. Factor 3 has been relatively steady, with large drops in both 2001 and 2008; possibly reflecting the financial crises in both periods. Finally, factor 4 has increased slightly, meaning the aircraft being utilized by airlines have become more expensive and with increased range.

Model results

The representative directed acyclic graph for the resulting model fit with the TSCM method described earlier in this chapter is shown in Figure 3-15. The circles at the center of the graph represent the variable at the current time, where M, A, and P are the manufacturer, airline, and public welfare, respectively, and each F term is one of the four factors. The arrows between each of these circles indicate a contemporaneous causal effect in the direction of the arrow, and the number next to the arrow is the magnitude of the causal effect. Note that the arrow from M to A is a dashed line, which indicates that this causal effect is not observationally distinguishable; the arrow could be turned in the other direction and this model would explain the observed data equally well.

The squares around the outside of the figure represent all of the significant temporal effects, where the number in parentheses represents the lag of the effect. For example, the term "M(-1)" in the box connected to P indicates that the manufacturer welfare in the previous period is found to have a causal effect on the public welfare in the current period. The number to the left of each term represents the magnitude of the causal effect.

In order to determine the complete magnitude of causal relationships between inputs we calculated the impulse response of a unit change in one variable on each of the others. This allows us to express the contemporaneous and temporal effects together to see the net effect over time of a change in one variable on the others. Figure 3-16 - Figure 3-22 present the impulse responses of all seven variables to a change in each variable individually. Each variable is initially at a steady state value of zero at period 1, and the impulse is applied at period 2, after which the value is fixed at zero. The effects of this impulse are then seen over the subsequent four periods, returning to a steady state of zero in the final period. The number to the right of each response plot shows the cumulative effect of the impulse on that variable as measured by the sum of the changes in the variable over the four impacted time periods. Large peaks and valleys in the impulse response indicate that the variable may be more sensitive to the change in the input over time (the derivative) than the absolute magnitude of the input. Note that a variable may have an impact on itself through temporal effects.

Figure 3-16 shows that an increase in manufacturer welfare will have very little net effect on the aircraft value, public welfare, or fuel costs. Airline welfare is negatively impacted by an increase in manufacturer welfare. It can also be seen that factor 1, fleet characteristics, is positively impacted by an increase in manufacturer welfare. Finally, manufacturer welfare is shown to affect factor 3, economic health. The fact that the effect is first negative and then positive indicates the change in manufacturer welfare, as opposed to the magnitude, is influential on economic health. While it seems unreasonable that aircraft manufacturers alone are influential on the entire US economy, these finding may indicate that aircraft manufacturers are a leading indicator of economic health, while our factor 3 is a lagging indicator.

Figure 3-17 shows the impact of an impulse to airline welfare. Airline welfare is found to have no impact on the manufacturer. Fleet characteristics, fuel costs, and aircraft value see negligible effects from a change in airline welfare. Like with manufacturer welfare, airline welfare is found to have a significant impact on economic health, though the net impact is small.

Again, this may indicate that airline profits are a leading indicator of economic health. Public welfare is negatively impacted by an increase in airline welfare.

Figure 3-18 shows the impacts of an impulse in factor 3, economic health. The effects of economic health on manufacturers, aircraft value, and fuel prices are found to be negligible. The cumulative effects of an impulse in economic health on all variables are quite small. Airline and public welfare are found to be affected by changes in economic health from the previous period, as shown by the positive and negative peaks in both impulse responses. Fleet characteristics are slightly impacted in the long term by an impulse in economic health.

The impulse response for factor 4, aircraft value, is shown in Figure 3-19. Manufacturer and airlines see no effect from aircraft value, and the change in economic health is negligible. Increasing aircraft value has a negative impact on factor 1, fleet characteristics. Public welfare is significantly impacted by the change in aircraft value from the previous period, though the cumulative effects of a change are nearly zero. Fuel costs are decreasing as aircraft value increases.

Effects on an impulse in public welfare are shown in Figure 3-20. We find that the manufacturer welfare is uninfluenced by public welfare, and aircraft value, fleet characteristics, and fuel costs are weakly impacted. Public welfare is shown to have a positive impact on economic health, suggesting the tickets sold per capita may also be a leading indicator of economic health. Airline welfare is significantly positively impacted by an increase in public welfare.

Figure 3-21 displays the impulse responses for factor 1, fleet characteristics. We see that economic health, public welfare, fuel costs, and aircraft value all have minor effects from changes in fleet characteristics. Manufacturers and airlines are both shown to experience a net increase in profits from an impulse in factor 1.

Finally, Figure 3-22 shows the results of an impulse in factor 2, fuel costs. Fuel costs are found no have no effect on manufacturer welfare and negligible effect on aircraft value. Airline welfare, public welfare, and economic health all experience little net impact from an impulse, but are highly sensitive to changes in fuel prices from period to period as illustrated by the large peaks and valleys in their impulse responses. Fleet characteristics experience a net negative impact from increases in fuel costs.

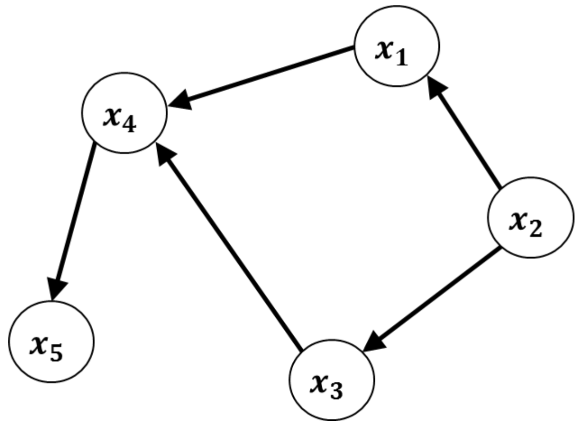
We can also look at the residual error of our causal model fit to the true values of each input in order to understand how well our model is explaining the data. Figure 3-23 shows the residuals for all 7 model inputs in their standard normal transformed values. We see that in general, the prediction is within two standard deviations of the true value for every input. Over the periods 1998-2002 and 2008-2010, we see large spikes in the residuals, indicating our model does not provide accurate values in these regions. This may indicate that our measure of economic health, factor 3, does not capture the full impact of the financial crises that took place in these two periods, or that other compounding events have a causal impact over these dates. Additionally, we can see that the residual values for each input are highly correlated, and there is a upward linear trend in the residuals over time. This suggests that there are likely one or more additional important causal effects not included in our available data.

We may check our confidence in this model by applying the robustness metric. Similarly to then data from Chen and Chihying, we find that very few coefficients are considered significant, and even then model robustness is low suggesting that the model may not be reliable due to having insufficient data available. However, we may still consider that the results of the model may be useful to a decision maker if they are aware of these concerns.

Model discussion

Based on these results, we might infer several relationships between stakeholder welfare, economic conditions, fuel prices, and aircraft characteristics. First, we observe that manufacturer welfare is only influenced slightly by changing fleet characteristics, and that manufacturer welfare at previous time periods is capable of approximating their current welfare. This result seems reasonable, as decisions about development and sales of new aircraft are typically made well in advance of their actual implementation. Were we able to extend the lag time of our model back multiple years, we might begin to see more interaction effects based on economic conditions, other stakeholder decisions, and new technology. Even still, we find that the manufacturer's past performance in the short time lag considered does a reasonably good job of predicting their welfare in the current period.

We also note that changing fleet characteristics are only a part of changes to the welfare of manufacturers and airlines. At the same, changes in aircraft characteristics are shown to be reactive to changing market conditions. This may be strongly influenced by the short time frame we are able to model, where changes in aircraft characteristics may be dominated by airlines choosing to modify the mixture of different types of aircraft in their utilized fleet as opposed to the introduction of new technology. Still, this model suggests that the success of new research and design activities will depend largely on changing market conditions and how airlines are available to change their fleet, so efforts should be made to consider future conditions when developing new technologies.



$$f(x_1, x_2, x, x_4, x_5) = f(x_5|x_4)f(x_4|x_1, x_3)f(x_1|x_2)f(x_3|x_2)f(x_2)$$

Figure 3-1. Directed acyclic graph for a Bayesian network. (From Chen and Chihying [18])

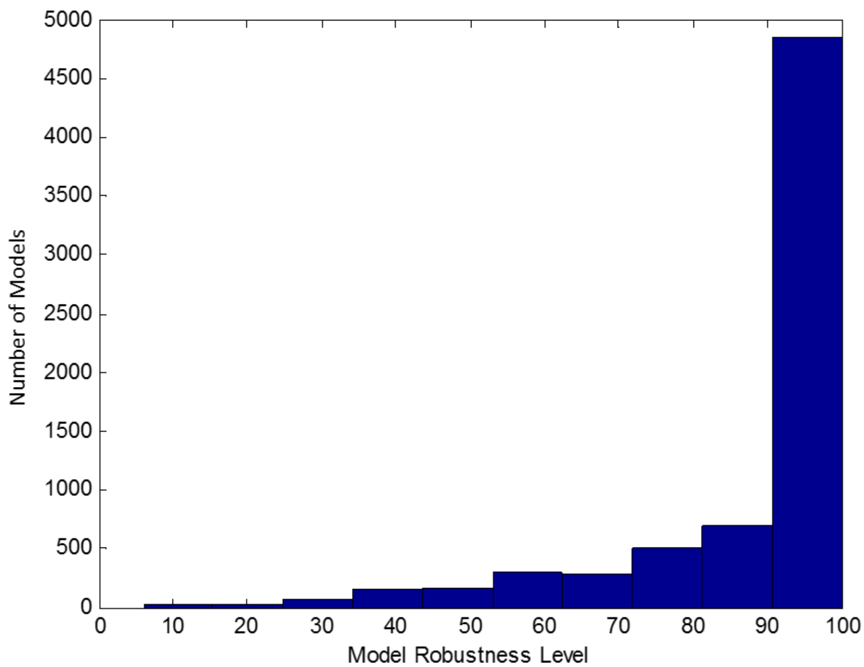


Figure 3-2. Number of correctly identified structures by robustness level (n = 7073)

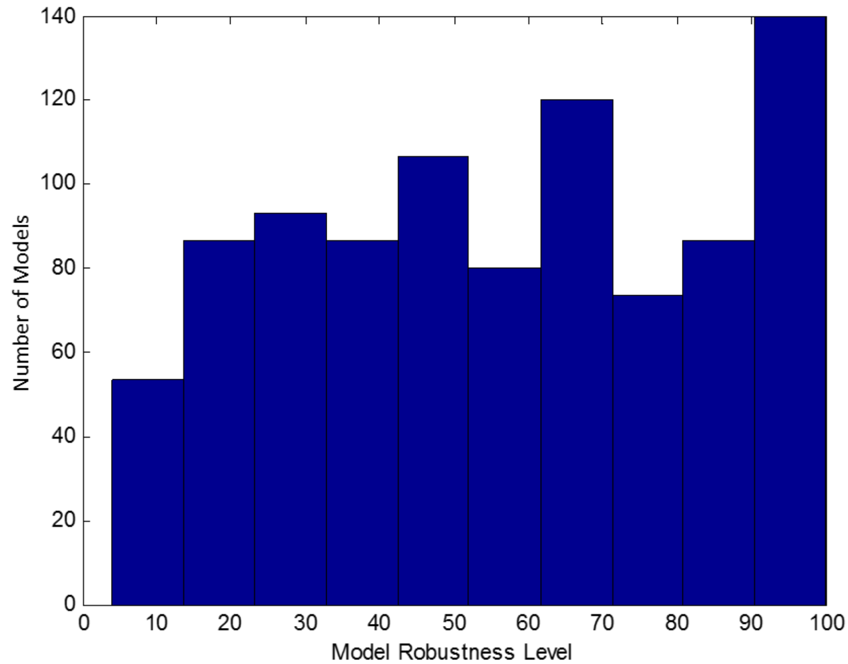


Figure 3-3. Number of incorrectly identified structures by robustness level (n = 927)

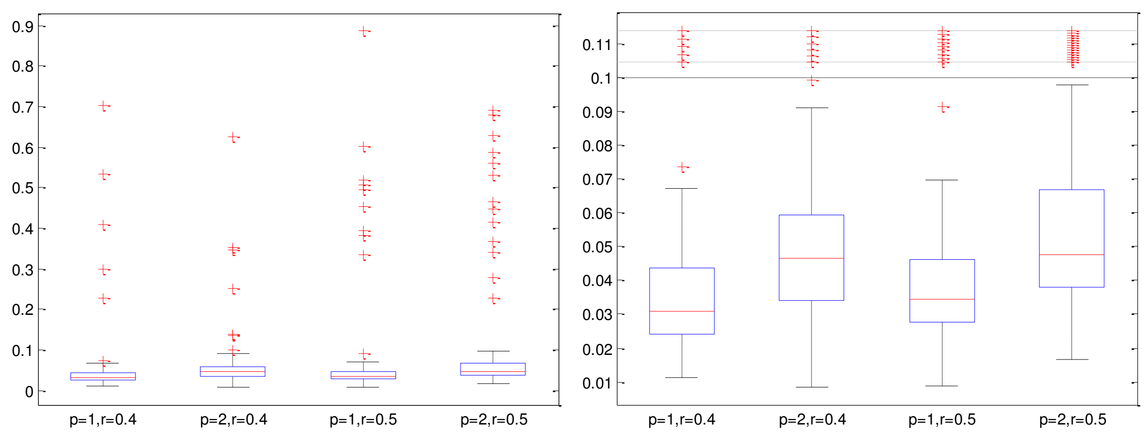


Figure 3-4. Accuracy score for each model size with 500 observations in (a) full and (b) zoomed plots

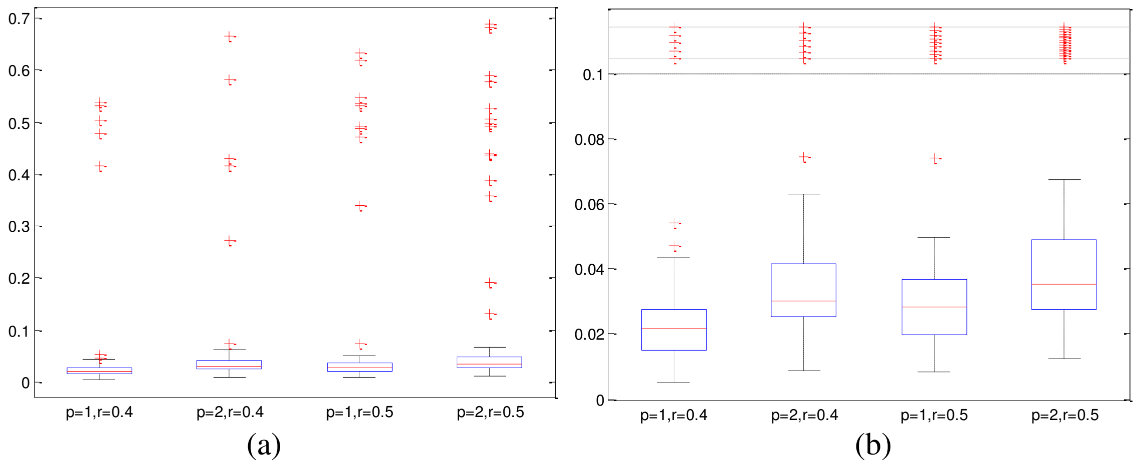


Figure 3-5. Accuracy score for each model size with 1000 observations in (a) full and (b) zoomed plots

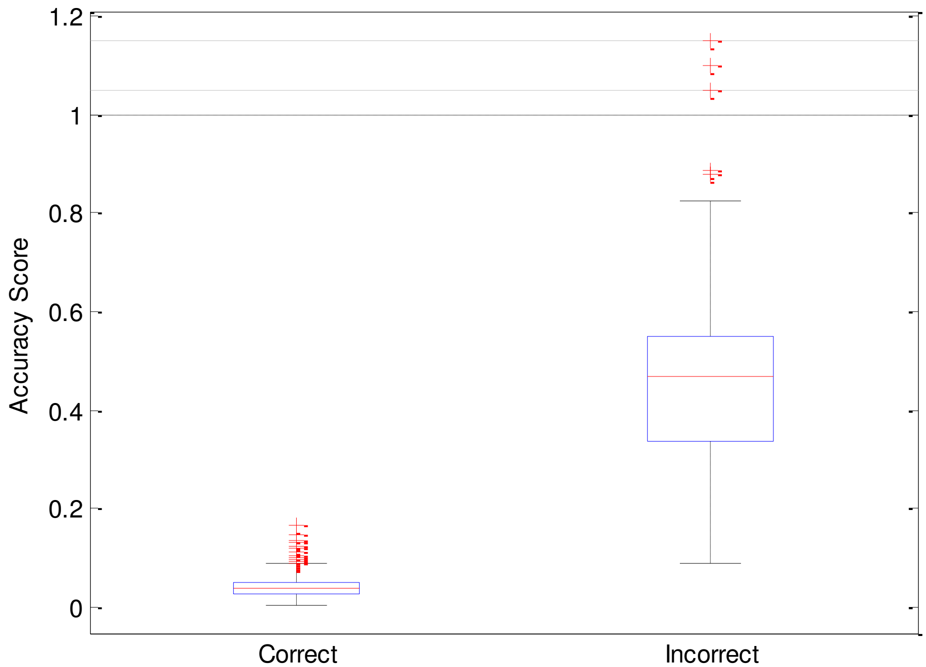


Figure 3-6. Accuracy score for correctly and incorrectly identified models

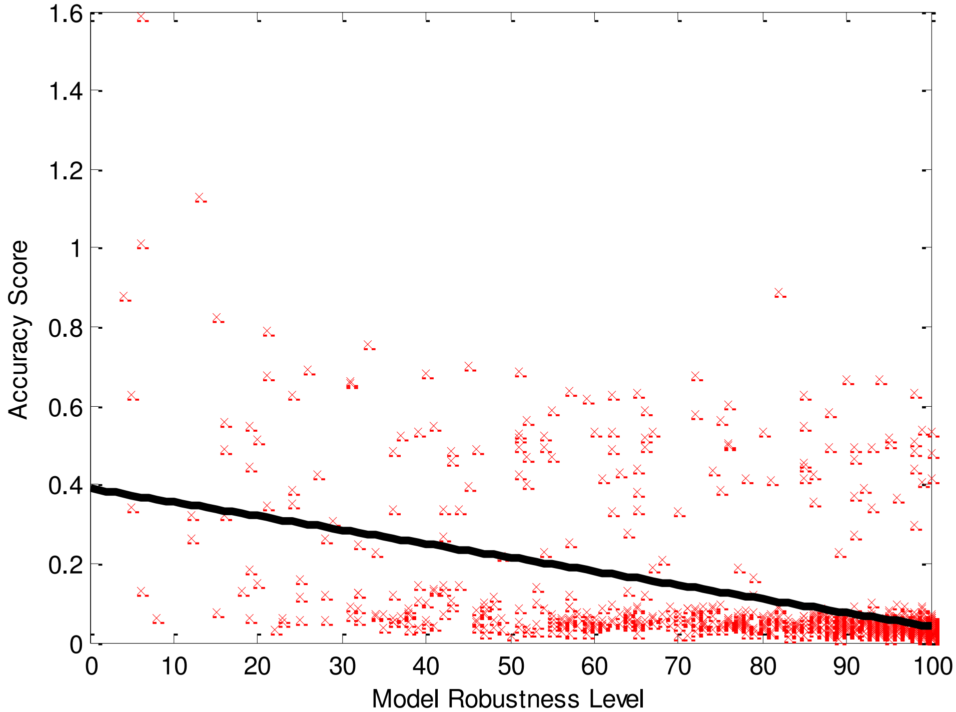


Figure 3-7. Comparison of model robustness and accuracy score for all cases

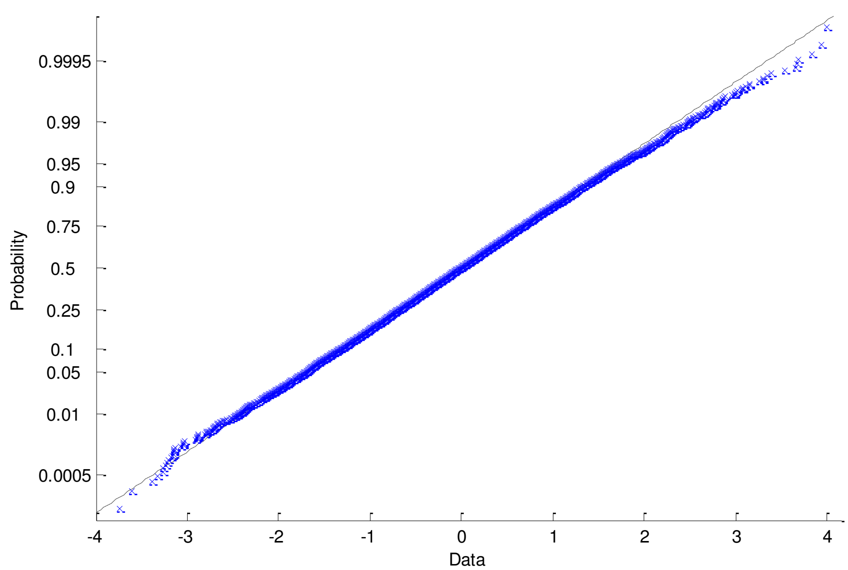


Figure 3-8. Probability plot of normalized error for all model sizes

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -2.73 & 0.38 & -3.18 & 1 & 0 & 0 \\ -0.36 & 0 & 0 & 0.05 & 1 & 0 \\ 1.56 & -0.47 & 0 & -0.50 & -1.93 & 1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} -1.03 & 0 & 0 & 0 & 0.28 & 0 \\ -0.19 & -0.52 & -0.40 & 0.07 & 0 & 0 \\ 0.02 & -0.12 & -0.90 & 0 & -0.08 & 0 \\ 0.64 & 0 & 0 & 0 & 0 & -0.21 \\ 0.30 & 0.02 & 0 & 0 & -0.91 & 0 \\ -2.01 & 0 & -0.54 & 0 & 1.94 & 0 \end{pmatrix}$$

Figure 3-9. Chen and Chihying causal model using TSCM method

$$A_0 = \begin{pmatrix} 1 & -0.04 & -0.16 & -0.21 & -0.91 & 0.05 \\ 0 & 1 & 0 & 0 & 0 & -0.13 \\ 0 & -0.06 & 1 & 0 & 0.29 & 0 \\ 0 & 0.55 & 0 & 1 & -1.29 & -0.38 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} -1.00 & 0 & 0.15 & 0 & 0.95 & 0 \\ -0.14 & -0.50 & -0.29 & 0.09 & 0 & -0.08 \\ 0 & -0.07 & -0.88 & 0 & -0.28 & 0 \\ 0.89 & -0.20 & 0.13 & 0.07 & 0.66 & -0.20 \\ -0.19 & 0.07 & 0.04 & 0 & -0.77 & 0 \\ -0.22 & -0.18 & -0.61 & 0 & -0.29 & 0 \end{pmatrix}$$

Figure 3-10. Sparsest permutation model results for wage-price dynamics

$$A_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} -1.19 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.87 & 0 & 0 & 0 \\ 0.80 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.90 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 3-11. Robustness metric suggested model for wage-price dynamics

$$\sigma_{A_1} = \begin{pmatrix} 0.05 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.03 & 0 & 0 & 0 \\ 0.13 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.03 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 3-12. Robust model temporal coefficient standard error

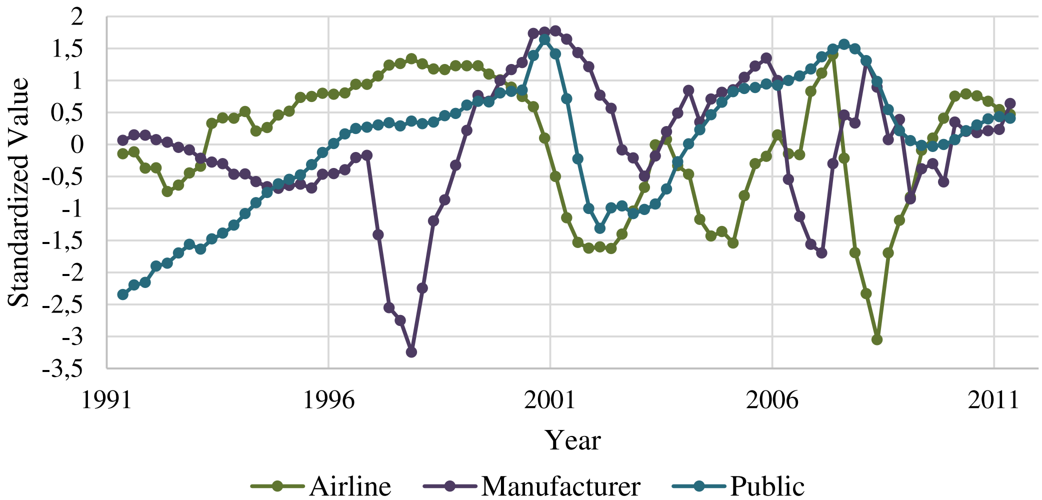


Figure 3-13. Normalized stakeholder welfare

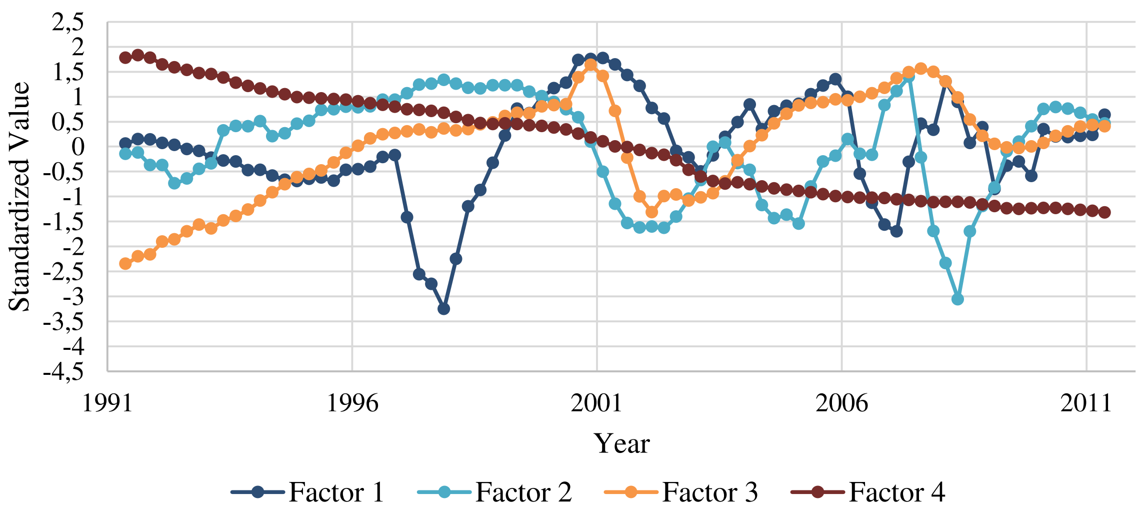


Figure 3-14. Normalized factor magnitude

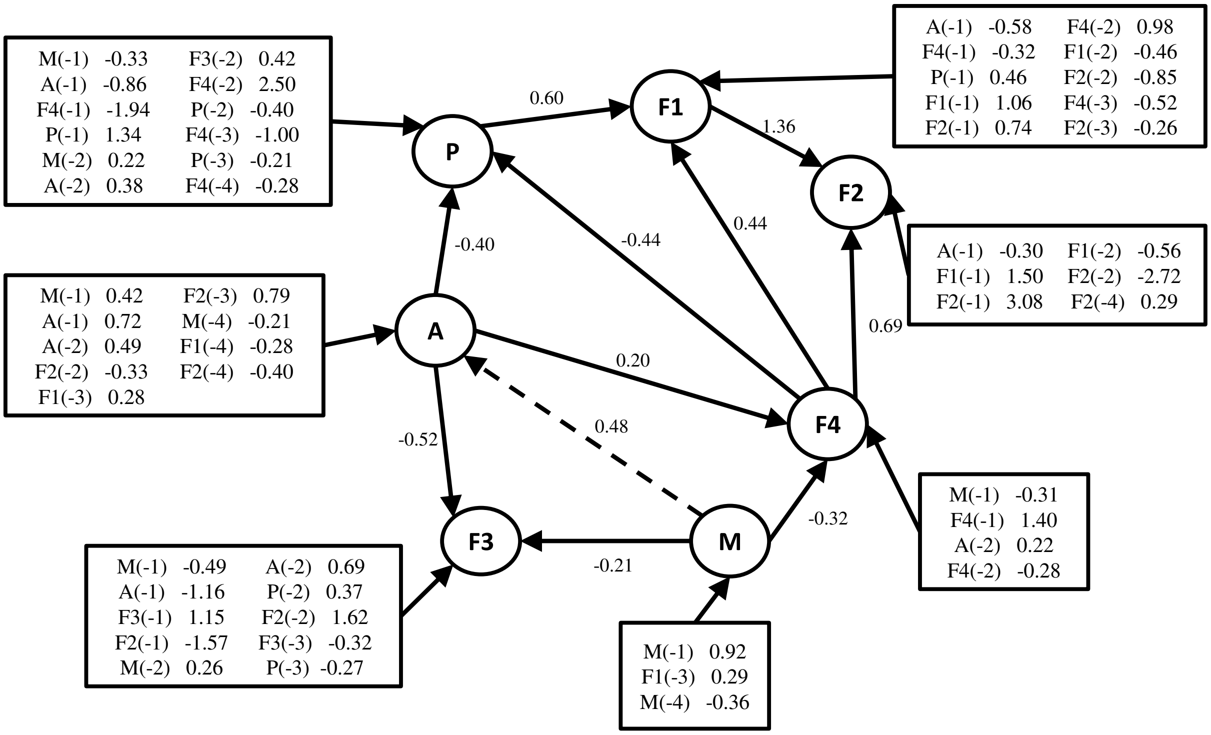


Figure 3-15. Causal model DAG structure

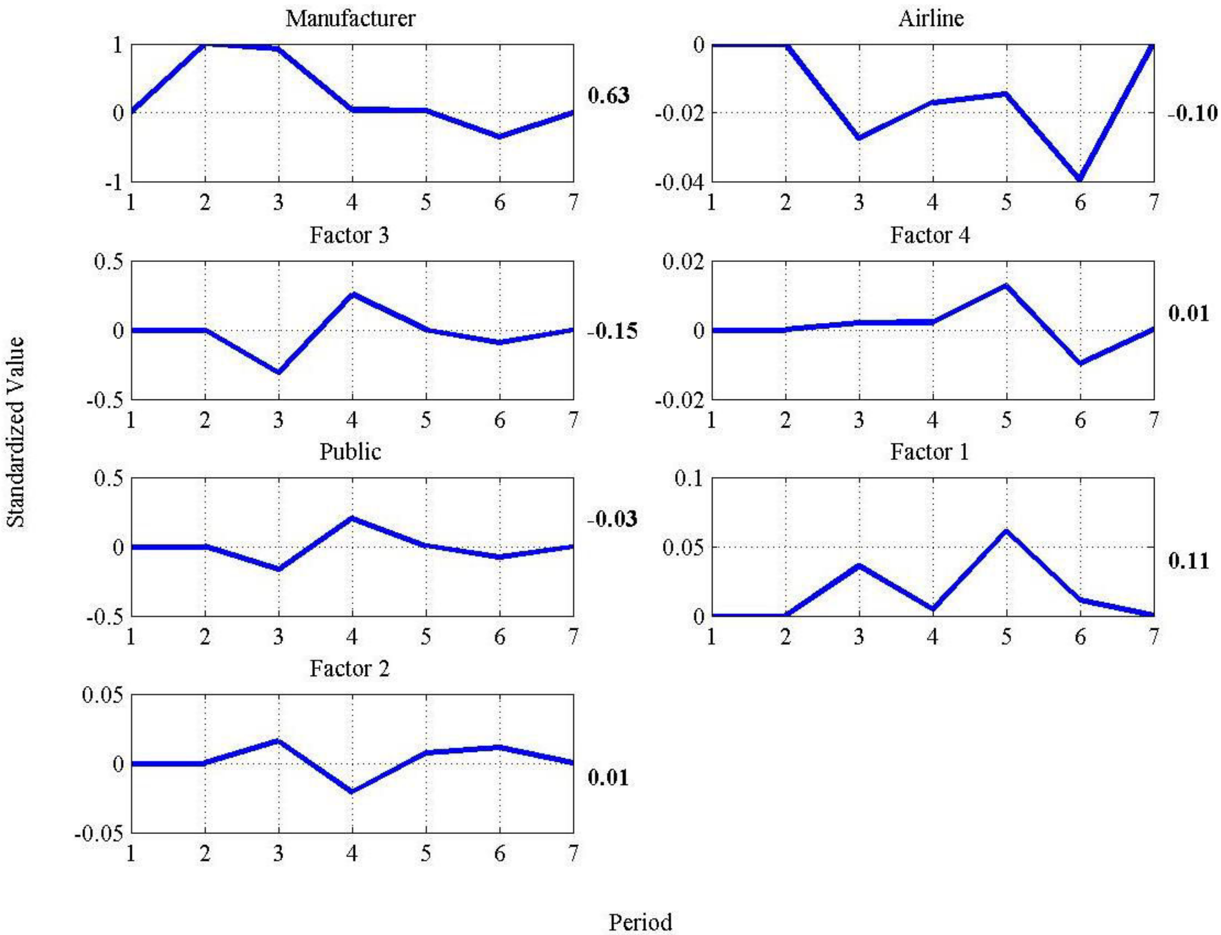


Figure 3-16. Manufacturer welfare impulse response

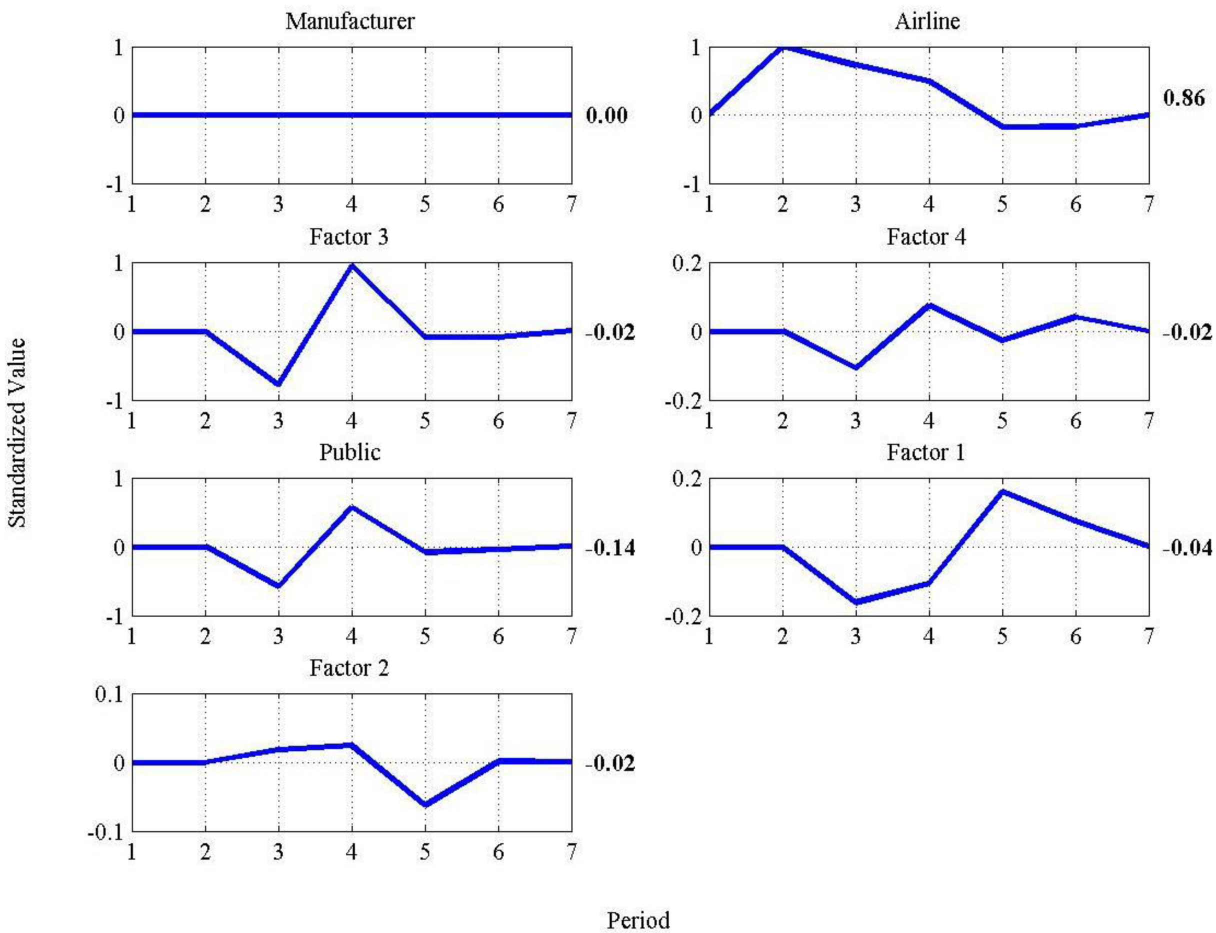


Figure 3-17. Airline welfare impulse response

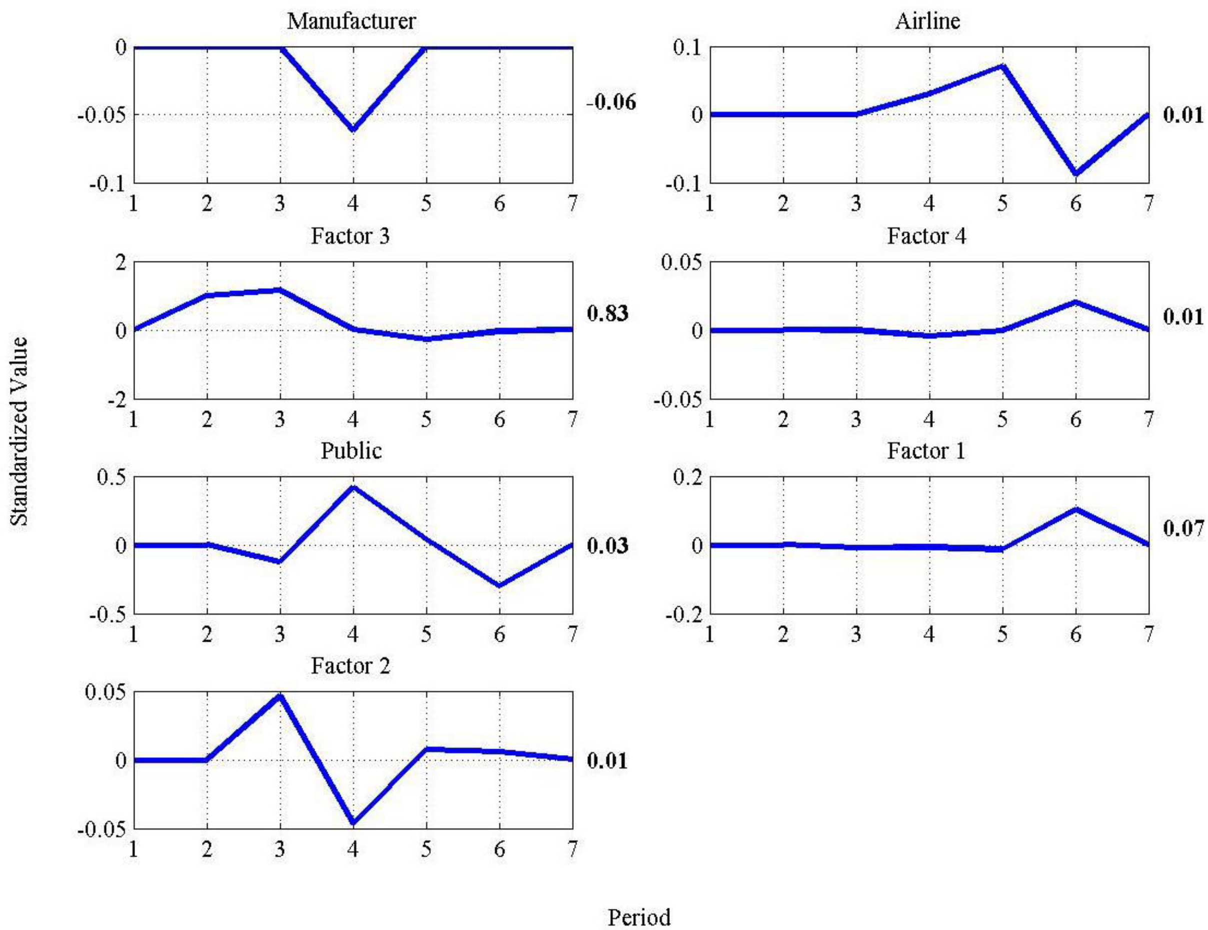


Figure 3-18. Factor 3 (economic health) impulse response

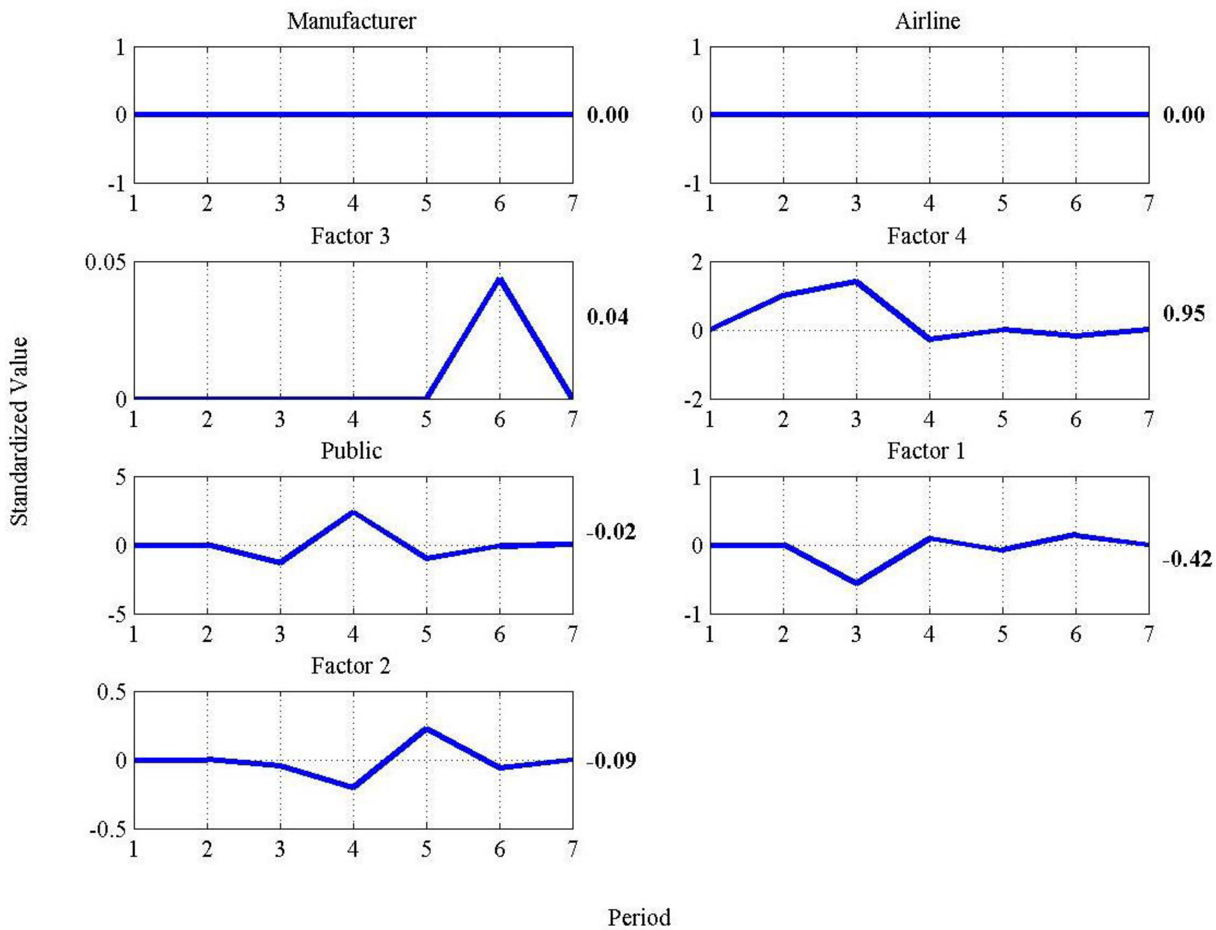


Figure 3-19. Factor 4 (aircraft value) impulse response

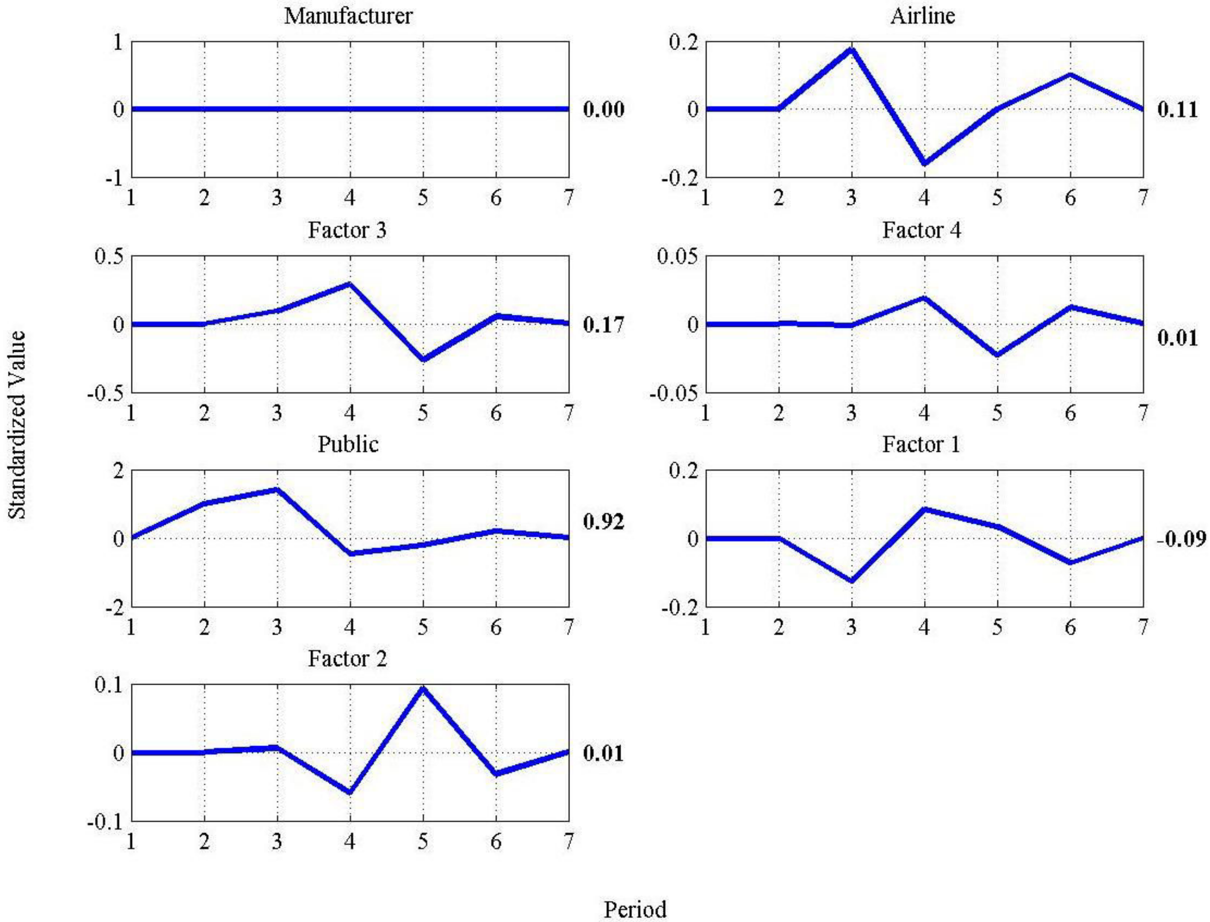


Figure 3-20. Public welfare impulse response

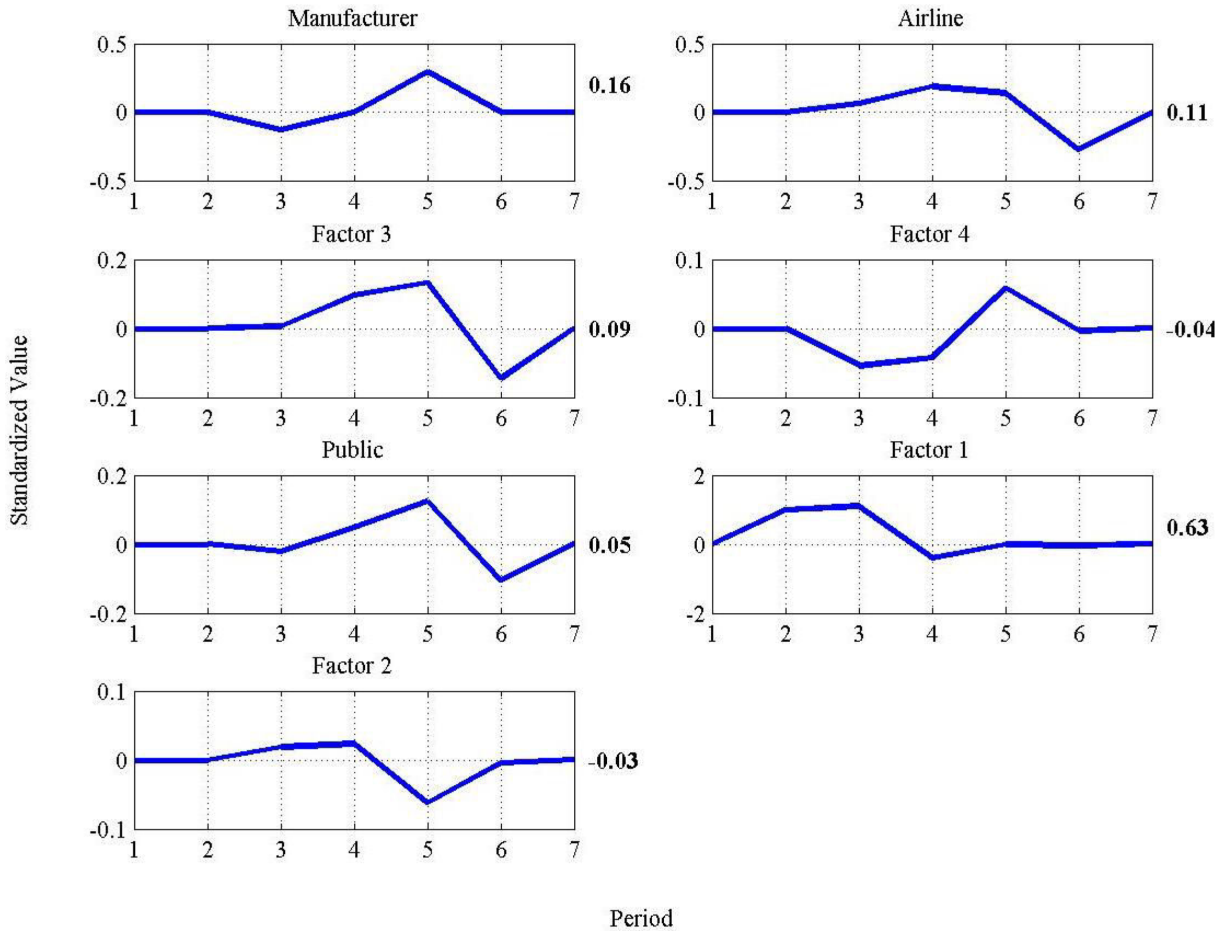


Figure 3-21. Factor 1 (fleet characteristics) impulse response

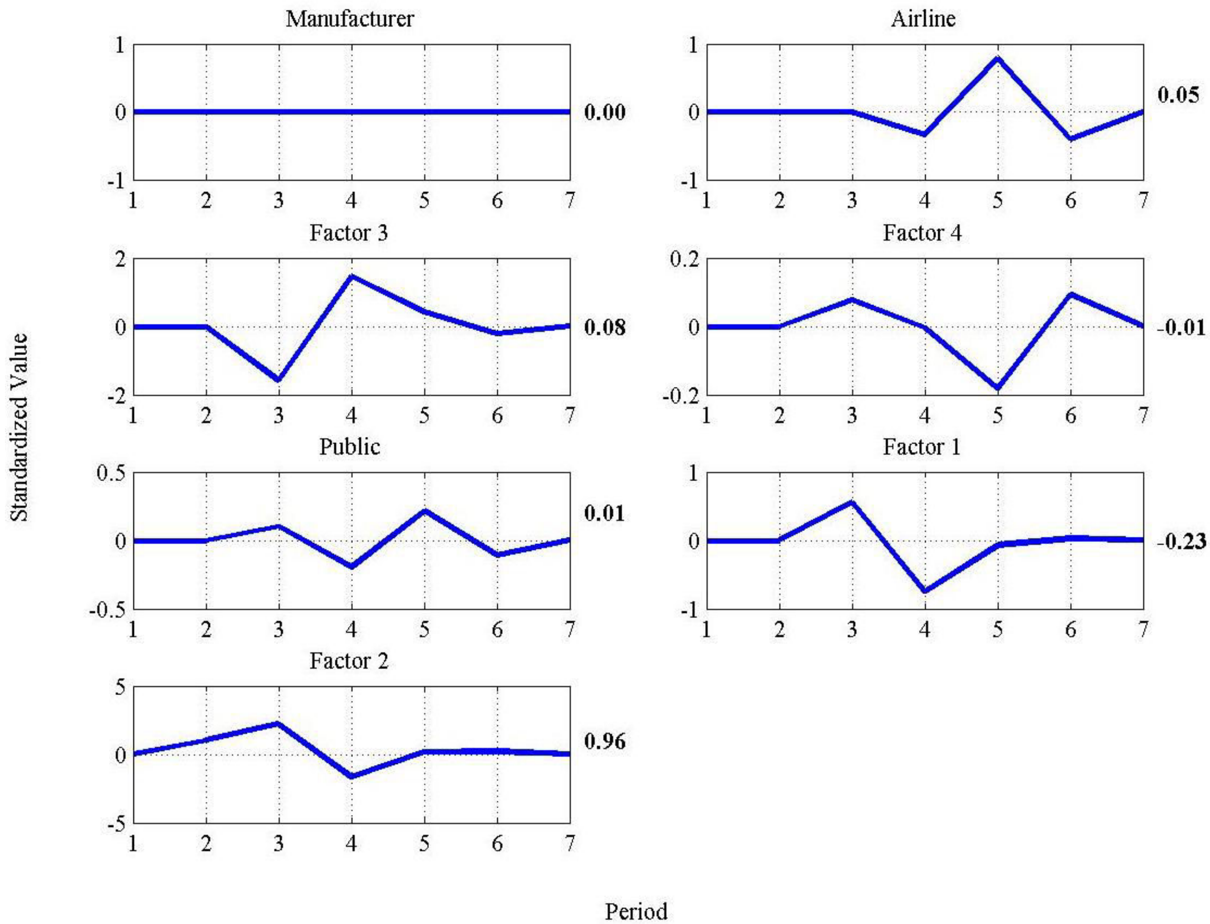


Figure 3-22. Factor 2 (fuel costs) impulse response

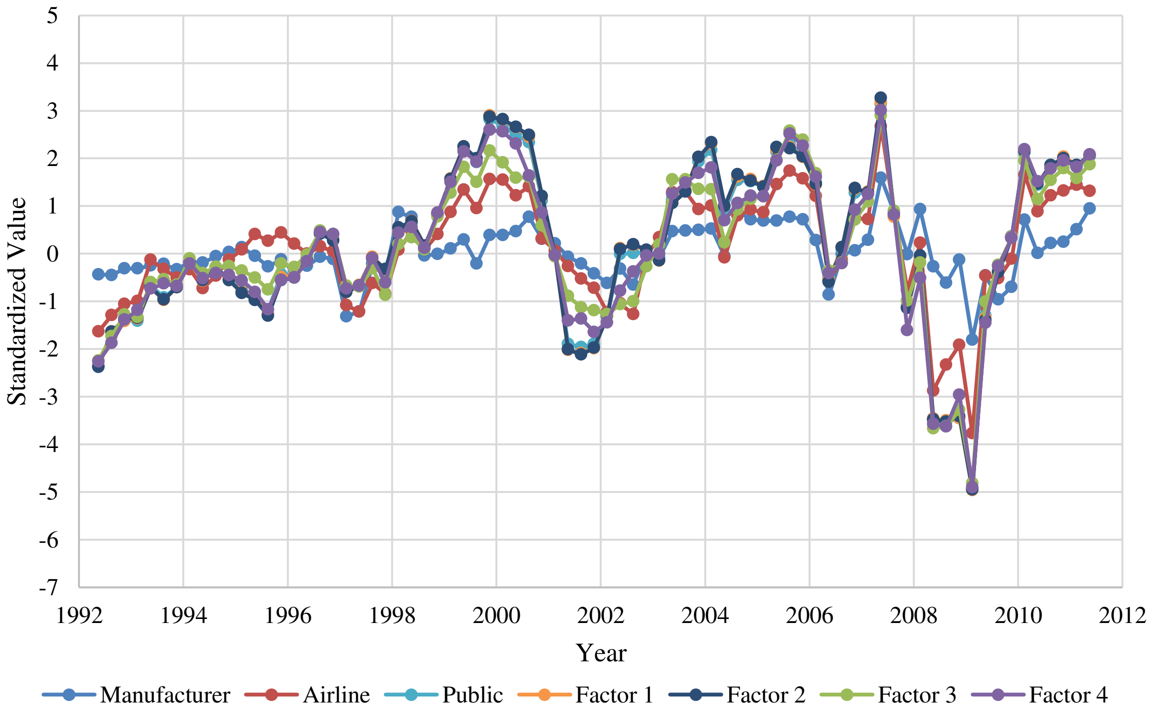


Figure 3-23. Residual values of causal model fit

Table 3-1. Robustness method structure recovery rate for p lags and connectivity ratio r (recovery rate of sparsest permutation without considering robustness)

	500 observations	
	$r = 0.4$	$r = 0.5$
$p = 1$	95% (93%)	88% (92%)
$p = 2$	87% (92%)	78% (78%)
	1000 observations	
	$r = 0.4$	$r = 0.5$
$p = 1$	96% (95%)	90% (90%)
$p = 2$	97% (91%)	86% (84%)

Table 3-2. Raw data used for empirical investigation of the model [18]

Variable	Transformation	Mnemonic	Description of the untransformed series
e	log(1-UNRATE/100)	UNRATE	Unemployment Rate (%)
u	log(GDPC1/GDPPOT)	GDPC1, GDPPOT	GDPC1: Real Gross Domestic Product of Billions of Chained 2000 Dollars, GDPPOT: Real Potential Gross Domestic Product of Billions of Chained 2000 Dollars, u: Capacity Utilization: Business Sector (%)
w	log(HCOMPBS)	HCOMPBS	Business Sector: Compensation Per Hour, Index 1992=100
p	log(IDPBS)	IDPBS	Business Sector: Implicit Price Deflator, Index 1992=100
z	log(OPHPBS)	OPHPBS	Business Sector: Output Per Hour of All Persons, Index 1992=100
π_m	MA(dp)		Inflationary climate measured by the moving average of price inflation in the last 12 periods

Table 3-3. Collected input data

Physical Characteristics	Welfare Measures	Economic Indicators	Other
Maximum Range	Manufacturer Profit†	Consumer Sentiment	Fuel Price†
Maximum Speed	Airline Profit†	Expected vs Actual Inflation	Fuel Cost per Available Seat Mile†
Useful Payload	Tickets Sold per Capita	Production Manager's Index	Date
Fuel Capacity		Change in Gross Domestic Product†	Average Ticket Price†
Vehicle Miles per Gallon		Change in Consumer Price Index	Listed Aircraft Price†
Passenger Capacity		Change in Unemployment	
Zero Fuel Weight			
Operating Empty Weight			
Required Flight Crew			
Maximum Takeoff Weight			
Available Seat Miles per Gallon			

†All dollar values corrected for inflation

Table 3-4. Constructed factors and corresponding factor loadings

Factor 1: Physical Characteristics		Factor 2: Fuel Cost		Factor 3: Economic Conditions		Factor 4: Aircraft Value	
Maximum Range	-0.74	Fuel Cost	0.67	Consumer Sentiment	0.26	Maximum Range	0.44
Maximum Speed	0.92	Fuel Cost per Available Seat Mile	0.70	Expected vs Actual Inflation	-0.28	Listed Aircraft Price	0.84
Useful Payload	1.00			Production Manager's Index	0.85		
Fuel Capacity	0.99			Change in Gross Domestic Product	0.25		
Vehicle Miles per Gallon	-0.88			Change in Consumer Price Index	0.62		
Passenger Capacity	0.98			Change in Unemployment	-0.91		
Zero Fuel Weight	0.99						
Operating Empty Weight	1.02						
Required Flight Crew	0.94						
Maximum Takeoff Weight	0.99						
Available Seat Miles per Gallon	-0.81						

CHAPTER 4
ASSIGNING VALUE TO SAFETY: A STUDY OF COST EFFETIVENESS OF TRANSPORTATION SAFETY MEASURES

French Chapter Summary

Dans les chapitres précédents, nous avons généralement considéré les préoccupations envers la sureté du transport aérien comme étant simplement modélisé par un indicateur économique définissant la valeur statistique de la vie. Toutefois, une analyse préliminaire des pratiques en matière de sureté des transports révèle que ces approximations simples ne sont souvent pas respectées. La sureté a des effets importants sur la perception qu'ont les consommateurs de concepteurs ou des fournisseurs de services. De plus les types et les probabilités de différents problèmes de sureté jouent un rôle important dans l'attention qu'ils reçoivent. Pour tenter de mieux comprendre ces préoccupations, nous procédons à une analyse du rapport coût-efficacité des mesures de sécurité adoptées dans le domaine du transport aux États-Unis sur la période 2002 - 2009. Grâce à cette analyse, nous pouvons acquérir une meilleure compréhension de l'importance des facteurs qui influent sur la relation entre valeur et sureté dans les problèmes de conception de systèmes complexes. Nous constatons que malgré une fatalité très inférieure, les compagnies aériennes et le transport par autobus sont soumises à un surcoût réglementaire bien plus important en comparaison avec l'aviation de loisir et le transport en voiture personnelle. Nous menons une étude sur deux grandes enquêtes sur les accidents de l'aviation commerciale et sur deux grandes campagnes de rappels de sureté automobile pour lesquels nous calculons le rapport coût-efficacité. Cela montre que les mesures prises pour améliorer la sureté suite à des accidents ont tendance à être très rentables pour le transport aérien, tandis que pour le transport automobile elles peuvent ne pas être rentables. Basé sur l'analyse de ces études, nous trouvons que la demande du public pour plus de sureté semble être grandement affecté par le niveau de la responsabilité personnelle des victimes impliquées.

En outre, nous trouvons que l'efficacité de différents types de mesures de sureté (réglementation ou enquête / rappel) est affecté par le nombre relatif de véhicules concernés, le nombre de vies en danger, et le niveau de sureté relative du mode. La question de savoir si une réponse réglementaire reflète la véritable demande du public pour la sureté ou simplement leur éventuelle perception erronée de la sureté est une considération importante afin de maximiser l'impact des ressources limitées affectées à l'amélioration de la sureté.

Overview

Thus far, we have generally considered the effect of safety concerns as simply modeled using a common economic construct known as the value of statistical life. However, even a cursory analysis of practices in transportation safety reveals that such simple approximations are often violated. Safety has important effects on consumer perception of designers and service providers, and relationships between the types and probabilities of various safety issues play an important role in the amount of attention they receive. To attempt to better understand these concerns, we conduct an analysis of the cost-effectiveness of safety measures enacted in the transportation industry in the United States in recent history. Through this analysis, we may gain a better understanding of the important factors affecting the relationship between value and safety in engineering design problems. The work in this chapter represents a joint effort with Dr. Taiki Matsumura who developed the cost effectiveness calculation for accident investigations [58].

Background

Travel related fatalities continue to be a leading cause of accidental death in the United States [59], despite significant improvements in recent decades [60]. In attempting to improve safety performance, one important constraint of new safety measures is cost, for which the US Department of Transportation uses the guideline of the value of statistical life, most recently set at $9.1 million [35]. Such constraints avoid undue financial burden on individual travelers and commercial transportation while ensuring that funds are not over-allocated to a single issue but spread across multiple risk sources.

A survey is conducted of the cost of US federal government regulations for safety enhancement in various modes of transportation, including commercial air carriers, commuter and air taxi, general aviation, private automobiles, and buses. This survey is intended to reveal how resources have been allocated for different modes of transportation. We will seek to uncover whether public demand for safety varies across different modes of transportation and if so, to determine why this might happen. We also investigate four significant transportation safety cases from the past decade, two from commercial aviation and two from automobiles, to attempt to determine if their remedies are cost effective. In each of these cases, we consider a system design flaw related to a miscalculation, manufacturing or maintenance error, or unexpected operating condition which led to a risk of fatal accidents. We determine the break-even point of the investment by calculating the probability of a fatality prior to the safety improvement necessary to justify the cost of the improvement. Through this demonstration, we discuss the cost effectiveness of the current safety improvement cycle in the civil transportation sector.

Several significant prior works have examined the way the cost effectiveness of safety improvements has been analyzed. Viscusi and Aldy [19] provide a detailed overview of various factors affecting the applied value of statistical life. Morrall [20] and Tengs *et al*. [21] both provide reviews of the cost effectiveness of previously implemented or proposed life-saving measures across many different fields. Cropper and Portney [22] outline some of the difficulties faced by regulators and policy makers in attempting to quantify cost effectiveness for new safety measures. Hammitt and Graham [23] outline the difficulty in assessing survey respondents' willingness to pay for safety, particularly in the case of highly unlikely events. Arrow *et al*. [24] provide a discussion of the ways in which cost-benefit analysis can and should be used to shape public policy.

Air travel has enjoyed many advances in safety technology since its inception. Safety enhancement in aviation is achieved not only by the evolution of technology, but also by incremental design improvements triggered by accidents. There were several epoch-making accidents that facilitated the evolution of the safety system [61] [62], such as repeated accidents of the De Havilland Comet in the 1950's leading to the recognition of metal fatigue. Aviation accidents have high public profiles due to a large number of potential fatalities. However, past research on the economics of aviation safety [63] [64] [65] [66], mainly triggered by the public concerns of airline deregulation in 1978 in the U.S., showed that market forces do not provide sufficient incentives for additional safety improvement. Thus, one important incentive for safety enhancement for air travel is safety measures mandated by laws and government regulations.

Private automobiles have also seen a great improvement in safety over recent decades. Safety features such as seat belts and air bags have become standard on all new vehicles, and campaigns for and in some cases laws requiring the use of seat belts have reduced the morbidity of accidents [67]. The advent of crash testing performed by the National Highway Traffic Safety Association (NHTSA) starting in 1979 has allowed for objective measurement of safety and facilitated competition between auto manufacturers on their safety performance [68]. A major issue for improving highway safety has been related to driver behavior, such as speeding, impaired driving, distracted driving, or the use of safety devices such as seat belts [68]. There is substantial investment by local governments in enforcing laws that promote safe driver behavior. Enforcement of pilot training and responsible behavior is stricter in commercial air travel, where a small group of highly trained pilots receive much more oversight by the FAA, relative to the standards for private automobile drivers.

Transportation Safety and Regulation

Safety Statistics

In order to examine the emphasis of the US government on transportation safety, we surveyed the accident related statistics and economic impact of new safety regulations enacted between 2002 and 2009.

To understand resource allocation for transportation safety, we must first quantify the level of safety in each transport mode. We utilize the conventional metric of fatalities per billion passenger miles[1] travelled. Data for passenger miles traveled and fatalities for various modes of transportation for years 2002 to 2009 were obtained from the 2010 National Transportation Statistics report . We categorized the mode of aviation according to the Code of Federal Regulations (CFR). CFR Part 121 is the regulation governing scheduled commercial airliners (we call it ‘commercial air’ in this work); CFR Part 135 governs on-demand air taxis and scheduled commuter carriers, such as business jets and regional airlines (‘commuter and air taxi’); and CFR Part 91 governs general aviation.

Additionally, the report contains safety statistics for highway transport, from which we distinguish private automobiles (cars, SUVs, light trucks, and motorcycles) and buses. Since the number of passengers involved in private transport is not explicitly known, we rely on survey estimations. The 2009 National Household Travel Survey [69] offers estimates of number of passengers and average distance traveled, from which we can determine passenger miles and total departures. Similarly, the annual FAA General Aviation and Part 135 Activity Survey [70]

[1] Passenger miles represent the total vehicle miles travelled multiplied by the average passenger load for a given mode

provides an estimation of the total number of departures for general aviation and commuter and air taxi operations separately. However, there is no data regarding the average number of passengers on these trips, and therefore we assume ranges of load factors to compute passenger miles traveled for these modes.

Table 4-1 shows the number of fatalities per billion passenger miles by mode of transport. We see that air carrier would be judged to be the safest mode, and roughly 250 times safer than private automobiles. General aviation, even with our highest estimates for passenger loads, is the least safe mode by a significant margin, while commuter and air taxi safety is comparable to private automobiles'. Buses rank as the second safest mode, but still are nearly an order of magnitude less safe than air carriers.

Regulation Review

Next, we consider the economic impact of Federal Government regulations that were enacted over the same time period. Transport safety regulation system in the United States is complex. A review we have conducted on http:www.regulations.gov yielded over 3,500 relevant regulations published from 2000-2009 by the Federal Aviation Administration (FAA), Federal Motor Carrier Safety Administration (FMCSA), National Highway Traffic Safety Administration (NHTSA), and the Department of Transportation (DOT). The reason for including safety regulations that pre-date the numbers for transport safety indicators we report above is to include any regulations whose effects might be delayed several years.

A summary review of the number and cost of regulations reviewed by agency is provided in Table 4-2. While there are too many regulations to list each individually, the highest cost regulations issued by each agency are explained in more detail in the following section with more listed in Appendix B. If regulations are deemed to involve a significant cost[2], the Office of Management and Budget (OMB) requires a cost-benefit analysis, and we use this analysis to determine the cost of each regulation to the US economy and transportation industry.

Notable Regulations Reviewed

While NHTSA is responsible for the most significant cost with a total of $73B, the FAA issued the highest number of regulations with 3297, 2140 of which were Airworthiness Directives (AD). Only 330 of the 3578 regulations reviewed were deemed to have a significant cost and therefore reviewed for cost effectiveness by the OMB. While regulations not deemed significant may have some non-zero cost, we feel this may be neglected due to the low cost of some significant regulations, as low as hundreds of dollars over 10 years. We additionally note that a small number of these regulations constitute a majority of the costs, so it is therefore worthwhile to consider some of these individually.

The single most expensive regulatory action of this time period is NHTSA's Corporate Average Fuel Economy standards which cost an estimated $47B; as this regulation has a minimal effect on transportation safety, we remove this regulation from our analysis and this cost is not included in Table 4-2 or subsequent analyses. Other significant regulations include: two regulations for $14B and $10B from NHTSA to improve rollover and roof crush risks, $12B from NHTSA for the Transportation Recall Enhancement, Accountability, and Documentation Act, $13B from the FMCSA to update commercial driver rest requirements, and FAA regulations of $1.3B, $1.1B, and $1.0B which related to maintenance on late life aircraft, aircraft material flammability standards, and catastrophic fuel tank explosions, respectively. Regulations

2 A regulatory action is considered "economically significant" under Executive Order 12866 § 3(f)(1) if it is likely to result in a rule that may have: "an annual effect on the economy of $100 million or more or adversely affect in a material way the economy, a sector of the economy, productivity, competition, jobs, the environment, public health or safety, or State, local, or tribal governments or communities."

which are not considered significant are often FAA ADs which re-designate airspace, or any agency's updated testing requirements which do not require significant changes to existing practice; we therefore consider potential costs of such regulations as negligible.

Regulatory Attention by Transport Mode

We aggregate the above safety regulations by mode based on the description provided in the regulation documentation, calculate the dollars spent on each mode over the entire period, and compare this to the number of observed fatalities, as shown in Table 4-3. Note that some regulations may be counted multiple times in this table if they are estimated to affect multiple transport modes. It can be seen that air carriers and buses receive much more regulatory attention per fatality than other modes. The regulation cost per fatality of air carriers is about 200 times as large as that of private automobiles.

While some of the regulations considered may have direct economic value beyond improving safety, we consider that improving safety is the primary benefit of each regulation. Based on our review of the most cost significant regulations as shown in Appendix B, we feel this is a reasonable assumption. We must recognize that a regulation would be deemed cost effective by the number of fatalities prevented, which we have no way of estimating. Instead, we consider the regulation cost per fatality observed, which gives us an idea of the total fatalities that could have been prevented, assuming that the number of fatalities would not have increased significantly absent the regulations. For comparison, we may consider the DOT specified value of statistical life (VSL), which is currently set at $9.1 million per fatality prevented.

Using this cost per observed fatality as a metric, we find that the cost of regulations for public modes of transportation, air carriers and buses, is significantly higher than their private counterparts, general aviation and private automobiles. As the rate of regulatory spending per fatality in these two modes is higher than the VSL, it may be argued that regulators are responding to a higher public demand for transportation safety in these modes. Commuter and air taxis receive moderate regulatory attention as compared to commercial aviation and general aviation, while general aviation and private automobiles, the two least safe modes, received the least regulatory spending per fatality by the US federal government.

Cost Effectiveness Study of Specific Fatal Accident Responses

The cost of regulations cited in the previous section is based on government estimates justifying the regulations. It is instructive to look at the actual costs incurred to correct some well publicized safety defects. Such an examination reveals that there are additional costs, including the cost of investigations to determine what safety defect caused fatalities, and the cost to recall vehicles.

Cost Effectiveness Measures of Safety Investigations and the Resulting Remedies

Accident investigations have been playing a central role in improving aviation safety. Elaborate investigations identify the probable causes of accidents and lead to safety recommendations to prevent similar accidents from occurring in the future. Independence of investigators from other authorities and separation from blame guarantee the quality of investigations, and aviation pioneered in this regard among other civil transportation modes [71] [72]. More recently, it has been proposed that the approaches and methods of aviation accident investigation be extended to wider context of social concerns, such as natural disaster, or economic fraud [73] [74]. Aviation is also a mode of civil transportation for which accident investigation is mandated in the U.S, and the National Transportation Safety Board (NTSB) is responsible for it.

Similarly, the NTSB carries out accident investigations for private automobiles, though due to the sheer number of accidents, not all will receive NTSB attention. When a safety issue is found requiring attention, typically after one or more accidents take place, the Department of

Transportation may mandate a recall of a certain vehicle or family of vehicles. As with aviation accidents, a series of accident and safety investigations may be undertaken; however unlike some aviation cases this is a relatively negligible cost for automobile recalls. It then falls to the responsibility of the manufacturer to provide an appropriate safety remedy for the affected vehicles and to cover the costs of this repair. Though automotive recalls and aircraft accident investigations are not identical, we view the results of both actions in terms of reacting to a safety issue with new measures as similar enough for comparison.

For a cost effectiveness study, we deploy a simple break-even calculation of the investment in an accident investigation or recall and focus on fatal accidents. The expense, C_{inv}, is the cost of the investigation and the following safety remedies, if needed. The payoff is the expected monetary value of lives to be saved, V_{saved}, in the future as a result of the investigation and remedies. Potential future fatalities related to an accident are calculated by the product of the expected number of fatalities N_f that would result from a similar accident, the number of airplanes or automobiles N_a that have the same failure potential, and the probability of reoccurrence of the accident in the remaining lifetime. For estimating N_a, one may take into account not only existing vehicles but also not-yet-built ones that will potentially benefit in the future from the improved design and safety regulations. Accident investigation has the potential to change the probability of accident reoccurrence, through implementation of the recommended safety measures. On this basis, the expected monetary value of lives to be saved (V_{saved}) can be calculated as

$$V_{saved} = V_{1life} N_f \, N_a (P_{before} - P_{after}) \tag{4-1}$$

where V_{1life} is the value of a single life, P_{before} is the probability of a fatal accident occurring per remaining lifetime of one vehicle before safety improvement is applied, and P_{after} is the probability of an accident after the improvement is applied. The break-even point happens when the invested cost in the investigation and remedies, C_{inv}, is equal to V_{saved}.

The dollar value of a fatality, V_{1life}, is defined as the amount we are willing to give up in exchange for a small decrease in the probability of one less fatality, called the value of a statistical life [75]. This is a common approach in economics, used to evaluate effectiveness of policies in medicine, environment and other areas. How much a society should invest in preventing fatalities is controversial, as seen in many ongoing discussions in different communities, e.g., health care, transportation, environment, etc. Viscusi [76] analyzed data on worker deaths across different industries, and suggested that the value of a life lies in the range of \$4.7 to \$8.7 million. In aviation, economic values used in investment and regulatory decisions of the U.S. Department of Transportation (DOT) were analyzed and determined. The guidance led to the value of \$6.2 million per fatality adopted in 2011 [77] and most recently updated it to \$9.1 million in 2013 [35]. Similarly in Europe, an aviation fatality avoided is valued at € 4.05 million by the European Transport Safety Council in 2003 [78].

For a given investigation and remedy cost C_{inv}, it is possible to calculate how much we spend to prevent the loss of one life in the future as

$$C_{1life} = \frac{C_{inv}}{N_f\, N_a (P_{before} - P_{after})} \tag{4-2}$$

This measure would be compared to the DOT guideline to determine whether accident investigation is cost effective or not. On the other hand, the value of lives to be saved can be used as the cost effective threshold of the invested cost $C_{inv,th}$ or the accident probability $P_{before,th}$ assuming that P_{after} is zero as in equations (4-1) and (4-2) respectively.

$$C_{inv,th} = V_{1life} N_f\, N_a P_{before} \tag{4-3}$$

$$P_{before,th} = \frac{C_{inv}}{V_{1life} N_f N_a} \tag{4-4}$$

Case Studies for Investigations into Fatal Accidents

American Airlines flight 587

The first example is the fatal accident of American Airlines Flight 587, which occurred on November 12, 2001. The airplane, an Airbus A300-605R, crashed into a neighborhood in Belle Harbor, New York, after taking off from the John F. Kennedy International Airport. All 260 people aboard and five people on the ground were killed [79] [80].

NTSB determined that the probable cause was "the in-flight separation of the vertical stabilizer as a result of the loads beyond ultimate strength that were created by the first officer's unnecessary and excessive rudder pedal inputs (when the pilot reacted to wake turbulence)." The NTSB report concluded that "The American Airlines Advanced Aircraft Maneuvering Program excessive bank angle simulator exercise could have caused the first officer to have an unrealistic and exaggerated view of the effects of wake turbulence." The report also discussed a widespread misunderstanding among pilots about performance of the rudder limiter system; pilots believed that a limiter would prevent structural damage no matter how they moved the control. However, the limiter did not take into account structural damage caused by repetitive opposite direction rudder inputs which resulted in the excessive load.

FAA issued an airworthiness directive (AD) in 2011 [81] requiring the modification to the rudder control system, called the pedal travel limiter unit (PTLU). The AD estimates the implementation cost of PTLU for 215 airplanes in the fleet at $42,677,500. For the cost effectiveness study, the number of potential fatalities was estimated at 213, based on the typical passenger capacity of the model (266 passengers) and a load factor of about 80% , and nine crewmembers. Adding the costs of accident investigation and other safety remedies (e.g., pilot training), which are not publicly available, the total invested cost is roughly estimated as $52 million US 2013 dollars.

Using above data, we calculate the cost effectiveness threshold of the accident probability $P_{before,th}$ defined in equation (4-4). Based on $9.1 million for V_{1life}, $P_{before,th}$ is estimated at 1.2×10^{-4} in the remaining lifetime of a single airplane. This probability corresponds to 6.0×10^{-9} per flight assuming that the remaining life time is roughly half the design service goal of the airplane (40,000 flight cycle [51]). It is remarkable that this probability is substantially smaller than the actual rate of fatal accident per departure from 2002-2009, 1.8×10^{-7} . Therefore, it can be said that this accident investigation is cost effective unless the probability of the accident is extremely small.

Alaska Airlines flight 261

The next example is the crash of Alaska Airlines Flight 261, which occurred on January 31, 2000. Fatalities included two pilots, three cabin crewmembers, and 83 passengers. The airplane, MD-83, was destroyed by impact forces [82]. The NTSB concluded that the probable cause was “a loss of airplane pitch control resulting from the in-flight failure of the horizontal stabilizer trim system jackscrew assembly’s acme nut threads. The thread failure was caused by excessive wear resulting from Alaska Airlines’ insufficient lubrication of the jackscrew assembly.”

According to the NTSB report, several factors contributed to the accident. First, lubrication of the nut threads was not adequately performed. Second, there were inappropriately wide lubrication and inspection intervals for the wear condition; because of this, wear exceeding its critical condition could not be discovered before the following lubrication or inspection point. The FAA issued airworthiness directives (ADs) [83] [84] [85] [86] [87] [88] [89] requiring repetitive inspections and lubrication. These improvements were applicable not only to MD series but also to Boeing airplanes. Table 4-4 shows the fleet sizes of airplane models, to which the ADs were applied, and the passenger sizes of those airplanes obtained from the company's website. We roughly assume that five inspections and lubrications would be needed in the rest of the lifecycle of each airplane, and the overhaul of nut and screw, which was applied only to Boeing-737, is a one-time item. Based on the work hours and labor rates provided by the ADs, as well as the estimated cost of inspection, we calculate the total cost for the safety improvements as roughly $18.5 million US 2013 dollars.

In the same manner as the previous example, we calculate the cost effective threshold of the accident probability as 2.5×10^{-6} per lifetime of single airplane. Here, we used $9.1 million for V_{1life}, and the number of potential fatalities to be saved is calculated by summing up $N_f N_a$ of each airplane model shown in Table 4-4 with a load factor of 80%. This probability can be converted into 1.2×10^{-10} per flight, which is also much lower than the actual fatal accident rate ($1.8\text{x}10^{-7}$). Note that we assume that the remaining flight cycle of the airplane is 20,000, half the design service goal of a typical airplane.

Despite being the two most fatal aviation accidents in the US over the time period we cover in this study, their total safety regulation cost (about $67 million) only represents about 1 percent of the total regulation cost of $7.5B as described in Table 4-4.

Ford/Firestone tread separation and rollover recall, 2000

The first automotive recall considered is the Ford/Firestone tire recall in 2000. NHTSA found that Firestone Wilderness AT and ATX tires produced at the Firestone plant in Decatur, IL were subject to tread separation during operation in certain conditions, particularly low pressure, high speed, hot weather operations, which could lead to increased risk of vehicle rollover when the tires fail. In particular, these tires were installed on Ford sport utility vehicles, where it was thought this issue would lead to a higher risk of roll over [90]. Firestone issued a recall of a total of 6.5 million tires, affecting roughly 1.3 million vehicles [91]. Though estimates vary, somewhere between 150 and 300 fatalities are thought to have been caused by this tread separation issue [92]. The total direct cost of this recall, shared between Firestone and Ford, is estimated at between approximately $1.3-1.7B in 2013 dollars [93]; the cost of investigation is assumed to be negligible as NHTSA's entire annual highway safety program budget is roughly $120M.

Using $9.1M as the value of statistical life, and considering an average passenger load for private automobiles as 1.5 based on the 2009 National Household Travel Survey [69], we calculate a threshold probability of a fatal accident per vehicle between 7.4×10^{-5} and 9.7×10^{-5}. In the case of automobile accidents, we can also estimate the actual probability of failure before the recall since we have a large number of vehicles and accidents. Based on the 1.3M vehicles affected and NHTSA's most likely estimate of tread separation related fatalities of 197 [92], we find the probability of a fatal accident as 1.5×10^{-4}, with 5% and 95% confidence bounds of 1.3×10^{-4} and 1.7×10^{-4}. We may additionally correct this probability to account for the fact that the 6.5M tires recalled were not all at end of life, and some may have failed later had they not been recalled. However, since this would depend on individual operating conditions, driving habits, tire age, and probability of tire failure over tire life, it is difficult to make any meaningful assessment, though we conservatively estimate a range from a factor of 1.5 to 3 increase in fatal accident probability based on the average age of vehicles included at the time of the recall and their expected life.

We find that the estimated range of probability of a fatal accident is only slightly higher than the range of threshold probability of failure based on the data collected. This indicates that the safety increase due to the recall likely at least broke even with the costs, and could be as much as a factor of 3 more. However, we note that this is a much narrower margin that seen with the previous two cases of accident investigations in air travel.

It should be noted that this investigation led to one of the most significant regulations for private automobiles during the time of our survey. This is the Transportation Recall Enhancement, Accountability, and Documentation (TREAD) Act, which cost a total of $24B, 38% of all private automotive regulation costs during our survey period from 2002-2009. However, this regulation deals exclusively with the way manufacturers report recalls and safety concerns to NHTSA, and does not specifically address the issue of tread separation or rollover and therefore this cost is not included in our calculations for this investigation.

Toyota unintended acceleration recall, 2009/2010

The second auto case considered is actually two related recalls which occurred at roughly the same time which dealt with the unintended acceleration accidents involving Toyota vehicles in 2009 and 2010. The first recall replaced floor mats in some Toyota vehicles which were believed to potentially cause the accelerator pedal to stick. The second dealt with wear in the accelerator pedal that could cause sticking unrelated to the floor mats. These recalls affected 2.23M and 4.44M vehicles, respectively. Due to overlap in the recalls, a total of nearly 5M vehicles were recalled. Toyota vehicles with fatal accidents attributed to unintended acceleration account for as many as 48 deaths [94], though DOT investigations concluded that many such accidents may actually be related to driver error [95]. The direct cost of the recalls, as reported by Toyota, was $1.12B [96] and the cost of the investigation is again assumed to be negligible in comparison.

Again using the value of statistical life of $9.1M and an average passenger load of 1.5, we find that the threshold probability of a fatal accident per vehicle for cost effectiveness is

1.65×10^{-5}. We again estimate the actual probability of failure based on the number of fatal accidents observed, and find a nominal value of 4.96×10^{-6} with confidence bounds of 3.00×10^{-6} and 6.91×10^{-6}. As with our previous auto recall example, we recognize that these estimates are based on some vehicles which are not at end of life. However, in this case we consider that it may be reasonable to assume that the probability of this specific accident is constant over the lifetime of the vehicle, and we may try to estimate the average age of vehicles in the recall. Based on available data, we consider that the average vehicle recalled was 5 years old, with a useful life of 15 years. Since we assume the probability of an accident is constant over time, we may simply correct our calculated probability of failure by a factor of 3, such that we find the estimated probability before the recall as 1.49×10^{-5} with confidence bounds of 9.00×10^{-6} to 2.07×10^{-5}. Even with this correction, we see that it is very likely that the probability of a fatal accident prior to the recall was below the cost effectiveness threshold.

Since this recall happened outside the date range of our regulation study, we have no direct information about any regulations resulting from this recall and investigation. However, the authors are unaware of any current or proposed regulations related to these Toyota recalls.

Discussion of Cost Effectiveness

Based on the studies presented in the previous sections, we may draw some conclusions about the cost effectiveness of safety measures for various modes and the allocation of resources within the US transportation sector. We may also attempt to understand some broader implications for safety measures based on differences between different transport modes.

First, we find that commercial aviation receives much more regulatory attention per fatality than general aviation. At the levels of regulatory spending seen during the years considered (2002-2009), commercial aviation received regulatory spending per fatality at a rate of roughly three times the DOT VSL. At the same time, general aviation received the second lowest regulatory attention per fatality despite being the least safe mode in our study based on fatalities per passenger mile. General aviation additionally received the lowest regulatory emphasis in terms of absolute dollars spent over the time period considered at $2.82 billion as compared to $6.43 billion for commercial aviation. A recent study of general aviation safety by Thomas Frank at USA Today [97] found that 86% of general aviation accidents are attributed to pilot error, including cases where subsequent investigations reveal defective parts contributed to an accident. These findings suggest that public demand for safety is lower when an individual is perceived to be responsible for their own safety, even though this may not always truly be the case.

We consider several reasons why this disparity between modes might exist. First, while fatal commercial airline accidents tend to be catastrophic events involving a large number of fatalities, general aviation accidents typically few people and occur more frequently and with less national coverage. Additionally, a different level of individual responsibility exists in general aviation accidents, where those involved are at least perceived to have some control over ensuring their own safety. This personal responsibility does not exist in commercial air travel, where travelers must place their trust in the pilot to ensure their safety. These factors together may lead to a higher perceived risk from commercial air travel by the general public, which is then reflected in their demand for new regulations. Even if risks are well understood, individuals may feel that a higher level of safety is appropriate in modes where a third party is providing transportation.

This idea is reinforced by the allocation of new regulations in private automobiles and bus travel, where buses receive much more regulatory attention, possibly due to the fact that, as with the aviation case, bus accidents involve a larger number of people who are dependent on a single driver to ensure safety. Finally, in the case of private automobile, much of the regulation and enforcement of responsible driver behavior is executed by state and local government. In that realm we see new regulations such as requiring seat-belt use and prohibitions on using cell phones. However, the increased regulatory cost of such measures is not expected to be able to offset greatly the factor of 450 between regulatory spending per fatality between buses and private automobiles.

These findings have important implications for future developments in transportation such as partially or fully automated transport systems. Many automobile companies and research groups have made significant efforts to develop autonomous or self-driving cars. As of 2015, four states (California, Nevada, Florida, and Michigan as well as the District of Columbia) have legalized the use of autonomous vehicles on public roads. A 2014 University of Michigan study of public perception of autonomous cars [98] found that 88% if survey respondents expressed concern about the prospect of riding in a fully autonomous vehicle, while over 60% also responding that the use of fully autonomous vehicles would be expected to reduce both the number and severity of accidents. These results reflect our suggestion that entrusting one's safety to an external actor increases an individual's demand for safety. Designers of autonomous systems as well as policy makers should therefore consider the public's increased demand for safety in these new modes as they develop.

We also reviewed two of the largest airline accident investigations and two of the largest automobile recalls from roughly the same time period. While both airline accident investigations were found to be easily cost effective by at least an order of magnitude based on the expected number of fatalities prevented, the automotive recalls were not so clear, with one being slightly cost effective and one likely below the cost effectiveness threshold. This implies that these more reactive safety measures are more likely to be cost effective for aviation compared to automobiles.

We propose that this difference is due to two factors: the population size of the vehicles affected, and the relative reliability of both systems. Though safety investigations for private automobiles are relatively inexpensive, the cost of performing a recall is often quite high even when the cost of a remedy is low due to the large numbers of vehicles involved, often millions. Conversely, commercial aviation accident investigations may be much more costly, but the resulting safety recommendations are only implemented on several hundred aircraft or less. Compounding this issue is the relative value at risk in terms of the number of people affected per accident, which is much greater for air travel than for automobiles. This means that fixing any one issue for an airline will result in a large proportional increase in safety, while for automobiles this safety increase may be only marginal.

This reveals important considerations for designers as well as regulators, as the types of safety issues faced in all transport modes typically have very low probabilities of occurrence, on the order of one in one million or less. This means that uncovering and preventing safety issues during the design process requires extensive testing and simulation, which is costly, and these costs are ultimately passed on to consumers. These costs are only exacerbated when the method of failure is unknown or difficult to predict, such as those related to operator error. We see that for airliners, safety concerns are easier to detect due to the rarity of accidents and in-depth investigation into each accident, while with private automobiles individual issues may be more difficult to find and are clearly more expensive to remedy due to the large cost of recalls.

Furthermore, the financial incentives for designers in commercial aviation and private automobiles are quite different. The relative low cost of addressing safety concerns in commercial aviation does not provide significant financial incentive for designers to attempt to avoid them, and public opinion after accidents generally affects airlines with little impact on designers [63] [64] [65] [66]. Conversely, we have shown that automobile recalls pose a significant financial burden applied directly to automobile designers. Automobile designers face additional losses related to public perception of their brand as they compete on safety records [99]. This may suggest that the automobile market is more efficient at creating improved safety, while commercial aviation safety requires more regulatory involvement. This might help to justify the differences in regulatory attention between modes. Understanding these differences between each mode may assist in an effective balance between preventative design and testing improvements versus oversight and improvements to existing products.

Summary of Findings

We have shown that despite their significantly better safety records, airlines and buses receive much more regulatory emphasis as measured by dollars per observed fatality as compared to general aviation and private automobiles, respectively. This could be due in part to non-regulatory safety actions in modes like private automobiles such as traffic enforcement and seat belt requirements, as well as an avoidance to the higher cost of safety remedies in these modes as seen in our review of recalls. We conducted a study of two major commercial aviation accident investigations and two major automobile safety recalls and calculated the cost effectiveness of each, finding that these tend to be much more cost effective for air travel, and may not be cost effective for automobiles. Based on analysis of these studies, we find public demand for increased safety appears to be greatly affected by the level of personal responsibility for ensuring safety. Additionally, the effectiveness of various types of safety measures (regulation or investigation/recall) is shown to be affected by the relative number of affected vehicles, the number of lives at risk, and the relative safety level of the mode. The question of whether this application of regulatory emphasis reflects the true public demand for safety or merely their possible misperception of safety is an important consideration for policy makers attempting to maximize the safety impact of limited resources. Designers and policy makers should be aware of these affects as they work on improving existing transportation as well as the development of novel modes. New forms of transportation such as self-driving cars may not only need to be safer than traditional cars, but may have even more stringent safety requirements.

Table 4-1. Annual fatalities per billion passenger miles (Year 2002-2009). The number in parentheses is the average number of fatalities per year during the period.

Air Carrier	Commuter and Air Taxi	General Aviation	Private Auto	Bus
0.038 (21)	4-11[a] (42)	30-160[a] (560)	9.09 (41,000)	0.26 (45)

[a]Passenger loads for commuter and general aviation are estimated at 5-10 and 1-3 passengers, respectively

Table 4-2. Breakdown of reviewed federal-government regulations by cost and agency

	FAA	FMCSA	NHTSA	DOT	Total
Number of Regulations	3297	34	187	60	3578
Number of Regulations with Significant Cost[b]	282	13	31	4	330
Total Significant Regulation Cost	\$7.5B	\$18B	\$73B	\$375M	\$99B
Cost of Top 3 Regulations	\$1.3B, \$1.1B, \$1.0B	\$13B, \$0.48B, \$0.38B	\$14B, \$12B, \$10B	\$280M, \$89M, \$5M	\$14B, \$13B, \$12B

[b]Based on regulations with cost estimates reviewed by US OMB

Table 4-3. Total federal regulation cost per fatality in millions for various transport modes (Year 2002-2009). The number in parentheses is the total cost in billions during the period.

Air Carrier	Commuter and Air Taxi	General Aviation	Private Auto	Bus
\$31 (\$6.4)	\$11 (\$4.8)	\$0.50 (\$2.8)	\$0.15 (\$63)	\$69 (\$31)

Table 4-4. Parameters estimated for Alaska Airlines case

Aircraft Family	Fleet size*	Passenger size (model)
MD-80	1218	155 (MD-83)
Boeing-767	411	218 (767-300ER)
Boeing-737	1641	162 (737-800)
Boeing-747	236	416 (747-400)
Boeing-757	730	280 (757-300)
Boeing-777	203	365 (777-300ER)

* Fleet size registered in the US

CHAPTER 5
CONCLUSIONS

In this work, we utilized a game theory framework in order to directly incorporate the effect of stakeholder interactions into a design optimization problem in order to maximize designer profitability. This approach allows for consideration of the effects of uncertainty both from traditional design variabilities as well as uncertain future market conditions and stakeholder interactions. Additionally, we developed techniques for modeling and understanding the nature of these complex interactions from observed data by utilizing causal models. Finally, we examined the complex effects of safety on design by examining the history of federal regulation on the transportation industry.

The key contributions of this research are as follows:

- The development of a novel optimization framework utilizing game theory methodology to account for the effect of interactions between multiple stakeholders, each of whom are performing their own profit maximization. Both simultaneous and sequential interactions were considered. This method allows for designers to maximize their expected profits while considering customer demands, competitors' actions, and changes in exogenous market variables.

- Optimization considering interactions was employed for several example problems, including the conceptual design of a commercial transport aircraft wing. Our findings from this work indicate that reasonable changes in market conditions are much more likely to have a significant effect on design choices than changes in traditional design uncertainty sources such as mechanical properties. Additionally, we showed that the level of market forecast uncertainty faced by a designer also has a significant effect on the optimal concept selected. These findings point to the importance of considering interactions and market uncertainty during design, as well as providing a framework through which designers may generate designs which are robust to various possible future conditions.

- A basic overview of learning causal models from observed data was provided. In particular, a novel method was developed for fitting causal models to time series data utilizing the existing sparsest permutation algorithm. We found that this algorithm showed high accuracy and low computational cost for the general size of model that a designer might generally be faced with.

- A novel metric for causal model confidence was proposed, based on the robustness of a given causal model with resampled data. We demonstrate through simulated data that this metric provides a measure of confidence in the accuracy of the resulting fitted causal model. In addition, this metric is able to produce accurate estimates of model coefficient uncertainty, which can be critically important information for designers who need to estimate the impact of new design concepts.

- A review of transportation safety statistics and improvements was conducted for the period of 2002 – 2009 in the United States. This revealed that regulatory attention for improving safety is not allocated among different transport modes proportional to the safety of each mode. Instead, we find that modes which involve a high level of personal responsibility for safety, such as general aviation or personal automobiles, receive significantly less regulatory attention per fatality than their public counterparts. We propose that this reflects an increased demand for safety from the public when entrusting their safety to a third party. This finding has important implications for both policy makers as well as designers working on novel transportation systems such as autonomous vehicles.

Dans de ce travail, nous avons proposé d'utiliser le cadre de la théorie des jeux afin d'intégrer directement l'effet des interactions entre les parties prenantes dans un problème d'optimisation de la conception visant à maximiser la rentabilité du concepteur. Cette approche permet de tenir compte des effets de l'incertitude liés à la fois aux variabilités traditionnelles sur les paramètres de conception ainsi que ceux liés au futur état du marché et aux interactions avec les autres acteurs du milieu. En outre, nous avons développé des techniques de modélisation visant à comprendre la nature d'interactions complexes à partir des données observées en utilisant des modèles causaux. Enfin, nous avons examiné les effets de la sureté sur la conception en examinant l'histoire de la réglementation sur l'industrie du transport.

Les contributions clés de cette recherché sont les suivantes:

- *Le développement d'un nouveau cadre pour la conception optimale utilisant une méthodologie issue de la théorie des jeux pour tenir compte de l'effet des interactions entre des parties prenantes multiples, où chacune effectue sa propre maximisation du profit. Des interactions simultanées et séquentielles ont été considérées. Cette méthode permet aux concepteurs de maximiser leurs profits attendus tout en tenant compte des demandes des clients, les actions des concurrents, et les changements dans les variables exogènes de marché.*

- *Ce nouveau cadre d'optimisation a été employé sur plusieurs problèmes d'application, notamment la conception d'une aile d'avion de transport commercial. Les résultats indiquent que des variations raisonnables dans les conditions du marché sont beaucoup plus susceptibles d'avoir un effet significatif sur les choix de conception que les variations liés aux sources d'incertitude traditionnels tels que les propriétés matériaux. En outre, nous avons montré que l'incertitude sur la prévision du marché futur a également un effet significatif sur le concept optimal sélectionné. Ces* *résultats soulignent l'importance de considérer les interactions et l'incertitude du marché lors de la conception, ainsi que la nécessité d'un cadre permettant la conception de systèmes robustes vis-à-vis de différents états futurs de l'environnement socio-économique.*

- *Un aperçu de l'apprentissage de modèles causaux à partir des données observées a été fourni. Une nouvelle méthode a été développée pour la construction des modèles causaux à partir des données de séries chronologiques utilisant l'algorithme de la permutation creuse. Nous avons trouvé que cet algorithme a une bonne précision et un faible coût de calcul pour des tailles de modèle raisonables.*

- *Une nouvelle métrique a été proposé pour quantifier le niveau de confiance dans le modèle causal construit. Cette métrique est basée sur la robustesse d'un modèle causal obtenu avec des données rééchantillonnées. Nous démontrons sur des données simulées que cette mesure fournit une bonne mesure de confiance dans l'exactitude du modèle de causalité construit. En outre, cette mesure est en mesure de produire des estimations de l'incertitude des paramètres du modèle. Cette mesure d'incertitude peut être d'une grande utilité pour les concepteurs qui ont besoin d'estimer de manière fiable l'impact de nouveaux concepts.*

- *Une étude des statistiques et des mesures pour améliorations de la sureté des transports a été réalisée pour la période de 2002 - 2009 aux États-Unis. Cette étude a révélé que l'attention des organismes de régulation n'a pas été répartie entre les différents modes de transport proportionnellement à la sécurité de chaque mode. A la place, nous constatons que les modes qui impliquent un haut niveau de responsabilité personnelle, comme l'aviation de loisir ou le transport en voiture personnelle, reçoivent nettement moins de contraintes réglementaires que les modes transport en commun (aviation commerciale, autobus). Cela reflète sans doute une demande accrue pour la sureté lorsque le public confie sa sécurité à un tiers.*

APPENDIX A
AIRCRAFT DESIGN PROBLEM DEFINITION

The total induced drag and lift distribution are approximated using lifting line theory [100], and can be calculated as

$$\mathrm{L(y)} = \rho \mathrm{V}_\infty \Gamma(\mathrm{y}) \tag{A-1}$$

$$\mathrm{D} = \rho \mathrm{V}_\infty \int_{-s}^{s} \Gamma \sin\left(\sum_n \frac{n A_n \sin(n * s * \cos y)}{\sin(s \cos y)}\right) dy \tag{A-2}$$

Where ρ is the air density, V_∞ is the free stream velocity, s is the wing half-span, and Γ is the circulation given as

$$\Gamma = 4 s V_\infty \sum_n A_n \sin(n * s * \cos y) \tag{A-3}$$

And the terms A_n can be determined by taking some finite n and solving the system of equations given by

$$\sum_n A_n \sin(n * s \cos y)\left(\sin(s \cos y) + \frac{n C_{l\alpha} c}{8s}\right) = \frac{C_{l\alpha} c}{8s} \sin(s \cos y)(\alpha - \alpha_0) \tag{A-4}$$

Where c is the local chord length, $\boldsymbol{\alpha}$ is the angle of attack, α_0 is the zero-lift angle of attack, and $\boldsymbol{C_{l\alpha}}$ is the slope of the lift coefficient, approximated from thin airfoil theory as $2\pi/rad$ [100].

Based on the drag determined above, we calculate the fuel burn for the aircraft using the thrust specific fuel consumption for the engines. Fuel burn serves as our primary aerodynamic discipline performance measure.

We then consider the structure of the wing as a tapered box beam subjected to the distributed load described by the lift distribution in equation (A-1). Since the wing structure must fit inside the wing, we constrain the outer dimensions of our box beam based the dimensions of a NACA 22112 airfoil, similar to those used on commercial transport aircraft. Based on this airfoil, we restrict the width and height of the box beam to 60% and 10% of the chord length, respectively, as shown in Figure A-1.

The box beam has two design variables: the horizontal member thicknesses at the wing root and wing tip. Because the wing considered in our example is only subjected to pure bending, the dimensions of the vertical members are not significant. Figure A-2 shows the dimensions of the box beam.

We optimize this structure in order to minimize weight (volume) subject to maximum stress and deflection constraints at some limit load, with the aircraft designer defined safety factor, SF. The weight of the wing structure is then calculated based on the volume of the design using the ratio of the nominal volume of a Boeing 737-700 wing structure and wing weight. The angle of attack is then updated based on equation (A-1) such that the total weight of the aircraft with the new structure is equal to the total lift at cruise.

The design is then subjected to a certification test subject to material property uncertainty, where the design must meet a specified knockdown factor against constraint violation determined as a function of the number of tests performed; the probability of not meeting this criterion will cause the aircraft designer to face a certification penalty. This penalty knockdown factor is given as

$$\mathrm{K_{pen}} = \mathrm{K_{min}} + \frac{\mathrm{K_{rate}}}{\mathrm{N_{test}}} \qquad \text{(A-5)}$$

This is intended to represent design requirements such as A-basis and B-basis used by the FAA, where increased number of tests would reduce the 95% confidence bounds, thereby reducing the required knockdown factor.

To calculate the probability of failure and probability of certification penalty, we calculate the design stress and design critical elastic modulus where the maximum deflection/span is achieved and compare to a prescribed variation in yield stress and elastic modulus based on a 5% COV in both properties. Assuming that both yield stress and elastic modulus follow a normal distribution, we calculate the probability of failure and certification penalty directly from these properties' cumulative distribution functions.

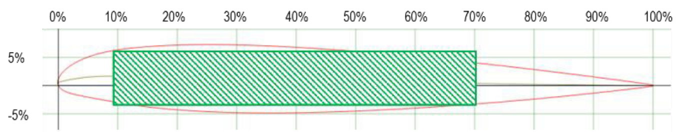


Figure A-1. NACA 22112 airfoil and approximate box beam dimensions [101]

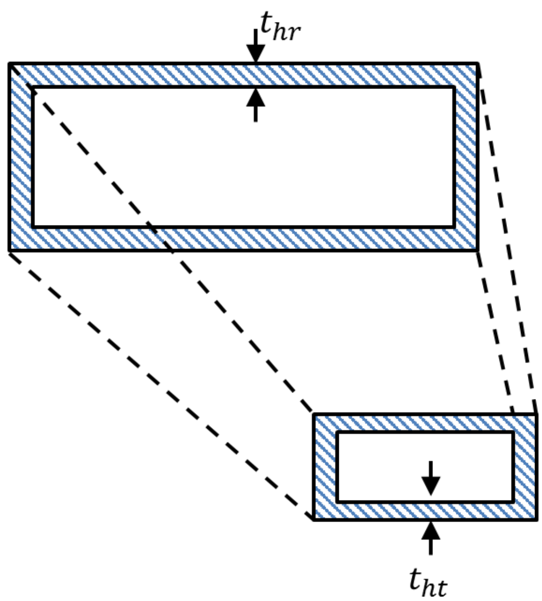


Figure A-2. Box beam dimension definitions

APPENDIX B
SIGNIFICANT REGULATIONS

Table B-1. Significant FAA Regulations

Document ID	Summary	Mode(s) Affected[1]	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
FAA-1999-5401-0145	Amends inspection and records keeping for aircraft of greater than 14 years of use.	C,A	3/4/2005	$ 1,350,000,000	18%
FAA-2000-7909-0043	The FAA is adopting upgraded flammability standards for thermal and acoustic insulation materials used in transport category airplanes.	C,A,G	8/2/2003	$ 1,084,000,000	14%
FAA-2005-22997-0154	Amends FAA regulations that require operators and manufacturers of transport category airplanes to take steps that, in combination with other required actions, should greatly reduce the chances of a catastrophic fuel tank explosion.	C	9/19/2009	$ 1,012,000,000	13%
FAA-2002-12261-009	This final rule permits the initiation of Reduced Vertical Separation Minimum (RVSM) flights in the airspace over the contiguous 48 States of the United States, the District of Columbia, Alaska, that portion of the Gulf of Mexico where the Federal Aviation Administration (FAA) provides air traffic services, the San Juan Flight Information Region (FIR), and the airspace between Florida and the San Juan FIR.	C,A,G	10/27/2003	$ 579,466,667	8%
FAA-2007-28058-0008	Adopts a new airworthiness directive (AD) for various IAE turbofan engines. This AD requires removing certain No. 4 bearing oil system components from service at the next shop visit or by an end date determined by the engine model.	C	8/20/2008	$ 450,371,650	6%

[1]Affected modes are defined as: C = Commercial, A = Air Taxi, G = General

Table B-1. Continued

Document ID	Summary	Mode(s) Affected[1]	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
FAA-2000-8490-0010	This final rule amends the list of airspace locations where Reduced Vertical Separation Minimum (RVSM) may be applied to include the New York Flight Information Region (FIR) portion of West Atlantic Route System (WATRS) airspace.	G	12/10/2001	$ 262,000,000	3%
FAA-2005-20836-0046	Adopts a new airworthiness directive (AD) for certain Boeing transport category airplanes related to the flammability of installed insulation blankets	C	12/15/2008	$ 177,700,000	2%
FAA-2004-18379	Fuel tank safety requirements related to electrical wiring, including updated inspection requirements for wiring systems.	C	12/10/2007	$ 166,400,000	2%
FAA-2001-11133-2709	The FAA is creating a new rule for the manufacture, certification, operation, and maintenance of light-sport aircraft. Represents an overall update to manufacture of aircraft and certification of pilots.	G	9/1/2004	$ 158,400,000	2%
FAA-2005-20245-0075	Amends cockpit voice recorder (CVR) and digital flight data recorder (DFDR) regulations affecting certain air carriers, operators, and aircraft manufacturers in order to improve the availability of CVR and DFDR information.	C,A	4/7/2008	$ 153,636,364	2%
FAA-2001-11032-0007	This amendment implements two security design requirements governing transport category airplanes related to the security of commercial aircraft cockpit doors to unauthorized intrusion.	C	1/15/2002	$ 131,000,000	2%
FAA-2003-15085-0075	The Federal Aviation Administration (FAA) is amending its hazardous materials (hazmat) training requirements for certain air carriers and commercial operators.	C,A	11/7/2005	$ 107,500,000	1%
FAA-2005-23500	Airworthiness Directive for International Aero Engines V2500 Turbofan Engines	C	11/15/2007	$ 99,338,400	1%

[1]Affected modes are defined as: C = Commercial, A = Air Taxi, G = General

Table B-1. Continued

Document ID	Summary	Mode(s) Affected[1]	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
FAA-2000-7018-0120	The interim final rule established fees for FAA air traffic and related services for certain aircraft that transit U.S.-controlled airspace but neither take off from, nor land in, the United States.	C	8/20/2001	$ 97,000,000	1%
FAA-2001-8724-0002	This final rule amends the existing airport security rules. It revises certain applicability provisions, definitions, and terms; reorganizes these rules into subparts containing related requirements; and incorporates some requirements already implemented in security programs.	C,A	11/14/2001	$ 92,200,000	1%
FAA-2007-0411-0001	Revises an existing airworthiness directive (AD) that applies to all Boeing Model 747 series airplanes. That AD currently requires that the FAA-approved maintenance inspection program be revised to include inspections that will give no less than the required damage tolerance rating for each structural significant item, and repair of cracked structure.	C	1/22/2008	$ 90,090,000	1%
FAA-2002-12504-0001	This final rule requires improved flightdeck security and operational and procedures changes to prevent unauthorized access to the flightdeck on passenger-carrying aircraft and some cargo aircraft operated by foreign carriers under the provisions of part 129.	C,A	6/21/2002	$ 83,200,000	1%
FAA-2004-18775	Brings US European Airworthiness standards closer in regards to flight guidance systems	C,A	5/11/2006	$ 69,636,364	1%
FAA-2007-0412-0001	Revising an existing airworthiness directive (AD) that applies to all Boeing Model 747 series airplanes related to corrosion and cracking certification.	C	1/22/2008	$ 62,304,000	1%

[1]Affected modes are defined as: C = Commercial, A = Air Taxi, G = General

Table B-1. Continued

Document ID	Summary	Mode(s) Affected[1]	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
FAA-2006-26722-0039	Adopts several standards of the International Civil Aviation Organization (ICAO) and requires manufacturers to incorporate certain security features in the design of new transport category airplanes related to unauthorized access to the cockpit of commercial category aircraft.	C	11/28/2008	$ 60,500,000	1%
FAA-2001-10910-0484	The FAA is revising the applicability of certain collision avoidance system requirements for airplanes.	C	5/1/2003	$ 59,000,000	1%
FAA-2004-18019-0006	The FAA is adopting a new airworthiness directive (AD) for Honeywell International Inc. related to stage 1 disk inspection and service	A	4/18/2005	$ 58,151,000	1%
FAA-2004-18038	Airworthiness directive for Honeywell T53 turboshaft engines life limit reduction for certain engine components	G	2/16/2006	$ 58,000,000	1%
FAA-2001-10047-0232	The Federal Aviation Administration (FAA) is updating and revising the regulations governing operations of aircraft in fractional ownership programs.	G	11/17/2003	$ 57,200,000	1%
FAA-2007-28283-0013	Adopts a new airworthiness directive (AD) for certain Boeing Model 737–600, –700, –700C, –800 and –900 series airplanes. This AD requires a one-time general visual inspection of frames between body station (BS) 360 and BS 907 to determine if certain support brackets of the air conditioning (A/C) outlet extrusions are installed; medium- and high-frequency eddy current inspections for cracking of the frames around the attachment holes of the subject brackets; and repair if necessary.	C	2/27/2009	$ 46,216,954	1%

[1]Affected modes are defined as: C = Commercial, A = Air Taxi, G = General

Table B-1. Continued

Document ID	Summary	Mode(s) Affected[1]	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
FAA-2001-8725-0003	This final rule amends the existing airplane operator security rule to include security requirement for additional types of operators related to terrorism and hazardous material threats.	C,A	11/14/2001	$ 40,000,000	1%
FAA-2001-10428-0020	This action amends the flight data recorder regulations by expanding the recording specifications of certain data parameters for specified airplanes, and by adding aircraft models to the lists of aircraft excepted from the 1997 regulations.	C	7/18/2003	$ 38,000,000	1%
All Other Regulations				$ 856,399,342	11%
Total				$ 7,553,710,740	

[1]Affected modes are defined as: C = Commercial, A = Air Taxi, G = General

Table B-2. Significant NHTSA Regulations

Document ID	Description	Mode Affected	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
NHTSA-2009-0093-0001	As part of a comprehensive plan for reducing the risk of rollover crashes and the risk of death and serious injury in those crashes, this final rule upgrades the agency's safety standard on roof crush resistance in several ways.	P	7/13/2009	$ 14,883,700,000	25%
NHTSA-2005-20586-0001	This final rule establishes a new Federal motor vehicle safety standard (FMVSS) requiring installation of a tire pressure monitoring system (TPMS) capable of detecting when one or more of a vehicle's tires is significantly under-inflated.	P,T,B	4/8/2005	$ 13,899,600,000	23%
NHTSA-2007-27662-0001-0001	As part of a comprehensive plan for reducing the serious risk of rollover crashes and the risk of death and serious injury in those crashes, this document establishes a new Federal motor vehicle safety standard (FMVSS) No. 126 to require electronic stability control (ESC) systems on passenger cars, multipurpose passenger vehicles, trucks, and buses with a gross vehicle weight rating of 4,536 Kg (10,000 pounds) or less.	P,T,B	6/5/2007	$ 10,835,000,000	18%
NHTSA-2000-8572-0219-0001	The first response to the TREAD act. It establishes a new Federal Motor Vehicle Safety Standard that requires the installation of tire pressure monitoring systems (TPMSs) that warn the driver when a tire is significantly under-inflated. The standard applies to passenger cars, trucks, multipurpose passenger vehicles, and buses with a gross vehicle weight rating of 10,000 pounds or less, except those vehicles with dual wheels on an axle	P,T,B	8/5/2002	$ 8,966,200,000	15%

[2] Affected modes are defined as: P = Private Auto, T = Commercial Truck, B = Bus

Table B-2. Continued

Document ID	Description	Mode Affected	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
NHTSA-2007-29134-0005	This final rule incorporates a dynamic pole test into Federal Motor Vehicle Safety Standard (FMVSS) No. 214, Side impact protection.' To meet the test, vehicle manufacturers will need to assure head and improved chest protection in side crashes.	P	11/13/2007	$ 6,160,000,000	10%
NHTSA-2005-23439-0001	In 6/2003, NHTSA published a final rule establishing upgraded tire performance requirements for new tires for use on vehicles with a gross vehicle weight rating of 10,000 pounds or less. We are amending the performance requirements for snow tires used on light vehicles.	P	6/1/2007	$ 1,199,000,000	2%
NHTSA-2004-19807-0002	This final rule upgrades NHTSA's head restraint standard in order to reduce whiplash injuries in rear collisions. For front seats, the rule establishes a higher minimum height requirement, a requirement limiting the distance between the back of an occupant's head and the occupant's head restraint, as well as a limit on the size of gaps and openings within head restraints.	P	3/14/2005	$ 985,140,000	2%
All Other Regulations				$ 3,160,905,325	5%
Total				$ 60,089,545,325	

[2]Affected modes are defined as: P = Private Auto, T = Commercial Truck, B = Bus

Table B-3. Significant FMCSA Regulations

Document ID	Description	Mode Affected[3]	Effective Date	Estimated Cost Over 10 Years	Percentage of Total Agency Cost
FMCSA-1997-2350-23305	Increased requirements for commercial vehicle driver rest and drive time limits	T	6/27/2003	$ 16,250,000,000	89%
FMCSA-2001-11061-0055	Improves requirements for safety audit of new commercial vehicle carriers	T,B	2/17/2009	$ 510,390,000	3%
FMCSA-2001-9709-0786	Updated requirements and penalties for commercial driver license holders related to non-commercial vehicle offences or convictions	T,B	1/29/2003	$ 466,250,000	3%
FMCSA-2000-7017-0028	Subjects commercial passenger transport (9-15 passengers) to same safety requirements as motorcoaches	B	9/11/2003	$ 213,000,000	1%
FMCSA-1997-2210-0209	Updates medical certification requirements for obtaining a commercial driver's license	T,B	1/30/2009	$ 199,020,000	1%
FMCSA-1997-2199-0218	Updates training requirements for obtaining a commercial driver's license	T,B	7/20/2004	$ 146,410,000	1%
FMCSA-1997-2277-0093	Updated rules and requirements for obtaining prior safety records of prospective commercial driver's license holders	T,B	4/29/2004	$ 136,730,000	1%
FMCSA-2002-13015-0023	Allows for better enforcement of existing rules to prevent motor carriers from operating outside their prescribed authority	T,B	9/27/2006	$ 108,300,000	1%
All Other Regulations				$ 206,738,750	1%
Total				$ 18,236,838,750	

[3]Affected modes are defined as: T = Commercial Truck, B = Bus

LIST OF REFERENCES

[1] T. Vincent, "Game Theory as a design tool," *Journal of Mechanisms, Transmissions and Automation in Design,* vol. 105, no. 2, pp. 165-170, 1983.

[2] S. Rao, "Game theory approach for multiobjective structural optimization," *Computers & Structures,* vol. 25, no. 1, pp. 119-127, 1987.

[3] K. Badhrinath and J. Rao, "Modeling for Concurrent Design Using Game Theory Formulations," *Concurrent Engineering,* vol. 4, no. 4, pp. 389-399, 1996.

[4] K. Lewis and F. Mistree, "Modeling subsystem interactions: a game theoretic approach," *Journal of Design and Manufacturing Automation,* vol. 112, no. 4, pp. 17-35, 2001.

[5] H. Li and S. Azarm, "Product design selection under uncertainty and with competitive advantage," *Journal of Mechanical Design,* vol. 122, no. 4, pp. 411-418, 2000.

[6] H. Li and S. Azarm, "An approach for product line design selection under uncertainty and competition," *Journal of Mechanical Design,* vol. 124, no. 3, pp. 385-392, 2002.

[7] S. Subrahmanyam *et al*., "Design of batch chemical plants under market uncertainty," *Industrial & Engineering Chemistry Research,* vol. 33, no. 11, pp. 2688-2701, 1994.

[8] D. Morrison *et al*., "Game Theory Analysis of the Impact of Single-Aisle Aircraft Competition on Emissions," *Journal of Aircraft,* vol. 49, no. 2, pp. 483-494, 2012.

[9] G. Hazelrigg, "A framework for decision-based engineering design," *Journal of Mechanical Design,* vol. 120, no. 4, pp. 653-658, 1998.

[10] P. Collopy and P. Hollingsworth, "Value-driven design," *Journal of Aircraft,* vol. 48, no. 3, pp. 749-759, 2011.

[11] J. Pearl, Causality: models, reasoning, and inference, Cambridge, MA: MIT Press, 2000.

[12] R. Curran *et al*., "Review of aerosapce engineering cost modelling: The genetic causal approach," *Progress in aerosapce sciences,* vol. 40, no. 8, pp. 487-534, 2004.

[13] A. Muller *et al*., "Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system," *Reliability Engineering & System Safety,* vol. 93, no. 2, pp. 234-253, 2008.

[14] C. Glymour *et al*., Discovering Causal Structure: Artificial intelligence, philosophy of science, and statistical modeling., Academic Press, 1987.

[15] P. Sprites *et al.*, Causation, Prediction and Search, Cambridge, MA: MIT Press, 2000.

[16] J. Pearl and T. Verma, A theory of inferred causation, San Mateo, CA: Morgan Kaufmann, 1991.

[17] G. Raskutti and C. Uhler, "Learning directed acyclic graphs based on sparsest permutations," *arXiv preprint,* vol. arXiv: 1307.0366, 2013.

[18] P. Chen and H. Chihying, "Learning Causal Relations in Multivariate Time Series Data," *Economics,* vol. 1, no. 11, pp. 1-43, 2007.

[19] W. Viscusi and J. Aldy, "The value of a statistical life: a critical review of market estimates throughout the world," *Journal of Risk and Uncertainty,* vol. 27, no. 1, pp. 5-76, 2003.

[20] J. Morall III, "Saving lives: A review of the record," *Journal of Risk and Uncertainty,* vol. 27, no. 3, pp. 221-237, 2003.

[21] T.O. Tengs *et al.*, "Five-hundred life-saving interventions and their cost-effectiveness," *Risk Analysis,* vol. 15, no. 3, pp. 369-390, 1995.

[22] M. Cropper and P. Portney, "Discounting and the evaluation of lifesaving programs," *Journal of Risk and Uncertainty,* vol. 3, no. 4, pp. 369-379, 1990.

[23] J. Hammitt and J. Graham, "Willingness to pay for health protection: inadequate sensitivity to probability?," *Journal of Risk and Uncertainty,* vol. 18, no. 1, pp. 33-62, 1999.

[24] K.J. Arrow *et al.*, "Benefit-cost analysis in envirmonmental, health, and safety regulation," American Enterprise Institute, Washington, DC, 1996.

[25] D. Fudenberg and J. Tirole, Game Theory, Cambridge, MA: MIT Press, 1991.

[26] Boeing, "Boeing 737-700 Technical Characteristics," [Online]. Available: http://www.boeing.com/boeing/commercial/737family/pf/pf_700tech.page. [Accessed 7 April 2014].

[27] Bureau of Transportation Statistics, "TranStats Database: Air Carrier Summary: T2: U.S. Air Carrier TRAFFIC And Capacity Statistics by Aircraft Type," [Online]. Available: http://www.transtats.bts.gov/Fields.asp?Table_ID=254. [Accessed 7 April 2014].

[28] Puget Sound Business Journal, "Mindful of rivals, Boeing keeps tinkering with its 737," 11 August 2008. [Online]. Available: http://www.bizjournals.com/seattle/stories/2008/08/11/story13.html?page=all. [Accessed 7 April 2014].

[29] Airlines for America, "Annual Crude Oil and Jet Fuel Prices," [Online]. Available: http://www.airlines.org/Pages/Annual-Crude-Oil-and-Jet-Fuel-Prices.aspx. [Accessed 7 April 2014].

[30] Boeing, "Boeing: Jet Prices," [Online]. Available: http://www.boeing.com/boeing.commercial/prices/. [Accessed 7 April 2014].

[31] H. Jiang, "Key Findings on Airline Economic Life," Boeing Technical Report, 2013.

[32] S. McCartney, "How Airlines Spend Your Airfare," The Wall Street Journal, 6 June 2012. [Online]. Available: http://online.wsj.com/news/articles/SB100001424052702303296604577450581396602106. [Accessed 7 April 2014].

[33] Federal Aviation Administration, "Material Strength Properties and Material Design Values (Advisory Circular 25.613-1)," 2003.

[34] Boeing, "The Boeing 737 Technical Site: Boeing 737 Detailed Technical Data," [Online]. Available: http://www.b737.org.uk/techspecsdetailed.htm. [Accessed 7 April 2014].

[35] U.S. Department of Transportation, "Guidance on Treatment of the Economic Value of a Statistical Life in U.S. Department of Transportation Analyses," Memorandum to: Secretarial Officers Modal Administrators, 2013.

[36] G. van Es, "Rapid Estimation of the Zero-Lift Drag Coefficient of Transport Aircraft," *Journal of Aircraft,* vol. 39, no. 4, pp. 597-599, 2002.

[37] Stanford University, "Data on Large Turbofan Enginges," [Online]. Available: http://adg.stanford.edu/aa241/propulsion/largefan.html. [Accessed 7 April 2014].

[38] B. Snyder, "Bend Not Break: Boeing 787 Passes Wing Load Test," CBS News, 31 March 2010. [Online]. Available: http://cbsnews.com/news/bend-not-break-boeing-787-passes-wing-load-test/. [Accessed 7 April 2014].

[39] M. D. Ardema *et al.*, "Analytical fuselage and wing weight estimation of transport aircraft," *NASA Technical Memorandum 110392,* 1996.

[40] J. Rustenburg *et al.*, "Statistical Loads Data for Boeing 737-400 Aircraft in Commercial Operations," Technical Report DOT/FAA/AR-92/28, 1998.

[41] Airlines for America, "Average round-trip fares and fees," [Online]. Available: http://www.airlines.org/Pages/Annual-Round-TripFares-and-Fees-Domestic.aspx. [Accessed 7 April 2014].

[42] D. Gates, "Boeing celebrates 787 delivery as development costs top $32B," Seattle Times, 24 September 2011. [Online]. Available: http://seattletimes.com/html/businesstechnology/2016310102_boeing25.html. [Accessed 7 April 2014].

[43] Boeing, "Boeing Archived Quarterly Earnings Releases," [Online]. Available: http://www.boeing.com/boeing/companyoffices/financial/quarterly.page. [Accessed 7 April 2014].

[44] EADS, "EADS Archived Financial Results," [Online]. Available: http://www.eads.com/eads/int/en/investor-relations/key-financial-information/Financial-Statements-and-Presentations/archived-results.html. [Accessed 2 June 2013].

[45] Bureau of Transportation Statistics, "TranStats Database: Air Carrier Summary: Schedule P-1.1," [Online]. Available: http://www.transtats.bts.gov/Fields.asp?Table_ID=290. [Accessed 7 April 2014].

[46] Bureau of Transportation Statistics, "TranStats Database: Air Carrier Summary: Schedule P-1.2," [Online]. Available: http://www.transtats.bts.gov/Fields.asp?Table_ID=295. [Accessed 7 April 2014].

[47] D. Madigan and A. Raftery, "Model selection and accounting for model uncertainty in graphical models using Occam's window," *Journal of the American Statistical Association,* vol. 89, no. 428, pp. 1535-1546, 1994.

[48] P. Chen and P. Flaschel, "Measring the Interaction of Wage and Price Phillips Curves for the U.S. Economy," *Studies in Nonlinear Dynamics & Econometrics,* vol. 10, no. 4, pp. 1-33, 2006.

[49] US Bureau of Economic Analysis, "Federal Reserve Economic Database," [Online]. Available: https://research.stlouisfed.org/fred2/. [Accessed 15 April 2015].

[50] S. Thompson, "Simple formulas for standard errors that cluster by both firm and time," *Journal of Financial Economics,* vol. 99, no. 1, pp. 1-10, 2011.

[51] Airbus, *FAST 45,* 2009.

[52] Bureau of Labor Statistics, "Consumer Price Index: All urban consumers," [Online]. Available: http://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt. [Accessed 2 June 2013].

[53] Bureau of Labor Statistics, "Labor force statistics from the current population survey," [Online]. Available: http://data.bls.gov/pdq/SurveyOutputServlet. [Accessed 2 June 2013].

[54] Institute for Supply Management, "ISM Manufacturer Report on Business PMI History," [Online]. Available: http://www.ism.ws/ISMReport/content.cfm?ItemNumber=10752. [Accessed 2 June 2013].

[55] University of Michigan, "The Index of Consumer Sentiment," [Online]. Available: http://www.sca.isr.umich.edu/data-archive/mine.php. [Accessed 2 June 2013].

[56] Federal Reserve Bank of St. Louis, "University of Michigan Inflation Expectation," [Online]. Available: http://research.stlouisfed.org/fred2/series/MICH. [Accessed 2 June 2013].

[57] Bureau of Economic Analysts, "National Income and Product Accounts Tables," [Online]. Available: http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=1&isuri=1. [Accessed 2 June 2013].

[58] T. Matsumura *et al*., "Modeling the Contribution of Accident Investigations to Airplane Safety," in *10th AIAA Aviation Technology Integration and Operations Conference*, Fort Worth, TX, 2010.

[59] M. Heron, "Deaths: Leading causes for 2010," *National Vital Statistcs Reports: National Center for Health Statistics,* vol. 62, no. 6, 2010.

[60] Research and Innovative Technology Administration (RITA), U.S. Department of Transportation, "National Transportation Statistics," [Online]. Available: http://www.bts.gov/publications/national_transportation_statistics/. [Accessed 6 August 2013].

[61] M. Mohaghegh, "Ecolution of structures design philosohpy and criteria," *Journal of Aircraft,* vol. 42, no. 1, pp. 814-831, 2005.

[62] J. Stoop and S. Dekker, "Are safety investigations pro-active?," *Safety Science,* vol. 50, no. 1, pp. 1422-1430, 2012.

[63] S. Borenstein and M. Zimmerman, "Matket Incentives for Safe Commercial Airline Operation," *American Economic Review,* vol. 78, no. 1, pp. 913-935, 1988.

[64] A. Chalk, "Market Forces and Commercial Aircraft Safety," *Journal of Industrial Economics,* vol. 36, no. 1, pp. 61-81, 1987.

[65] G. Kaplanski and H. Levy, "Sentiment and stock prices: The case of aviation disasters," *Journal of Financial Economics,* vol. 95, no. 1, pp. 174-201, 2010.

[66] N. Rose, "Profitability and Product Quality - Economic-Determinants of Airline Safety Performance," *Journal of Political Economy,* vol. 98, no. 1, pp. 944-964, 1990.

[67] E.M. Orsay *et al*., "Prospective study of the effect of safety belts on morbidity and health care costs in motor-vehicle accidents," *Journal of the American Medical Association,* vol. 260, no. 24, pp. 3598-3603, 1988.

[68] L. Robertson, "Reducing death on the road: the effects of minimum safety standards, publicized crash tests, seat belts, and alcohol," *American Journal of Public Health,* vol. 86, no. 1, pp. 31-34, 1996.

[69] A. Santos *et al*., "Summary of Travel Trends: 2009 National Household Travel Survey," U.S. Department of Transportation Federal Highway Administration, Washington, DC, 2011.

[70] Federal Aviation Administration, "General Aviation and Part 135 Activity Surveys," 2002-2009. [Online]. Available: http://www.faa.gov/data_research/aviation_data_statistics/general_aviation/. [Accessed 30 August 2013].

[71] J. Stoop, "Independent accident investigation: a modern safety tool," *Journal of Hazardous Materials,* vol. 111, no. 1, pp. 39-44, 2004.

[72] J. Stoop and J. Kahan, "Flying is the safest way to travel: How aviation was a pioneer in independent accident investigation," *European Journal of Transport Infrastructure Research,* vol. 5, no. 1, pp. 115-128, 2005.

[73] Y. Dien *et al*., "Accident investigation: From searching direct causes to finding in-depth causes - Problem of analysis or/and analyst?," *Safety Science,* vol. 50, no. 1, pp. 1398-1407, 2012.

[74] S. Roed-Larsen and J. Stoop, "Modern accident investigation - Four major challenges," *Safety Science,* vol. 50, no. 1, pp. 1392-1397, 2012.

[75] O. Ashenfelter, "Measuring the value of a statistical life: Problems and prospects," *Economic Journal,* vol. 116, no. 1, pp. C10-C23, 2006.

[76] W. Viscusi, "The value of life: Estimates with risks by occupation and industry," *Economic Inquiry,* vol. 42, no. 1, pp. 29-48, 2004.

[77] U.S. Department of Transportation, "Treatment of the Economic Value of a Statistical Life in Departmental Analyses-2011 Interim Adjustment," Memorandum to: Secretarial Officers Modal Administrators, 2011.

[78] European Transport Safety Council, "Cost effective EU transport safety measures," 2003.

[79] National Transportation Safety Board, "In-Flight Separation of Vertical Stabilizer American Airlines Flight 587 Airbus Industrie A300-605R, N14053 Belle Harbor, New York November 12, 2001," NTSB/AAR-04(04), 2004.

[80] B. Murphy *et al*., "Overview of the Structure Investigation for the American Airlines Flight 578 Investigation," in *46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Austin, TX, 2005.

[81] Federal Aviation Administration, "Airworthiness Directives; Airbus Airplanes (Final Rule), Docket No. FAA-2011-0518 (AD 2012-21-15)," 2012.

[82] National Transportation Safety Board, "Airplane Accident Report. Loss of Control and Impact with Pacific Ocean Alaska Airlines Flight 261 McDonnell Douglas MD-83, N963AS About 2.7 Miles North of Anacapa Island, California January 31, 2000," NTSB/AAR-02/01, 2002.

[83] Federal Aviation Administration, "Airworthiness Directives; Boeing Model 747 Airplanes, Docket No. FAA-2005-22624," 2005.

[84] Federal Aviation Administration, "Airworthiness Directives; Boeing Model 777 Airplanes, Docket No. FAA-2006-24270," 2006.

[85] Federal Aviation Administration, "Airworthiness Directives; McDonnell Douglas Model DC-9-10, DC-9-20, DC-9-30, DC-9-40, and DC-9-50 Airplanes; Model DC-9-81 (MD_91), DC-9-82 (MD-82), DC-9-83 (MD-83) and DC-9-87 (MD-87) Airplnaes, Docket No. FAA-2005-22254 (AD 2006-10-14)," 2006.

[86] Federal Aviation Administration, "Airworthiness Directives; Boeing Model 767 Airplanes (Final rule), Docket No. FAA-2005-22623 (AD 2008-06-06)," 2008.

[87] Federal Aviation Administration, "Airworthiness Directives; The Boeing Company Airplanes (Final rule) (Boeing-757), Docket No. FAA-2011-1093 (AD 2012-16-16)," 2011.

[88] Federal Aviation Administration, "Airworthiiness Directives; The Boeing Company Airplanes (Final rule), Docket No. FAA-2011-0518 (AD 2012-21-15)," 2012.

[89] Federal Aviation Administration, "Airworthiness Directives; The Boeing Company Airplanes (Final rule) (Boeing-737), Docket No. FAA-2008-0415 (AD 2011-27-03)," 2012.

[90] J. Claybrook, "Statement to the U.S. Senate Committee on Commerce, Science and Transportation," 12 September 2000. [Online]. Available: http://www.citizen.org/autosafety/article_redirect.cfm?ID=5414. [Accessed 5 November 2014].

[91] Bridgestone/Firestone, "Bridgestone/Firestone Announces Voluntary Recall of 3.85 million RADIAL ATX and RADIAL ATX II Tires, and 2.7 million Wilderness AT Tires [Press Release]," 2000. [Online]. Available: http://usgovinfo.about.com/blfirestone.htm. [Accessed 5 November 2014].

[92] National Highway Traffic Safety Administration, "Firestone Tire Recall: Complaints Listing," 4 October 2001. [Online]. Available: http://www.nhtsa.gov/nhtsa/announce/press/Firestone/Update.html. [Accessed 5 November 2014].

[93] C. Isidore, "Ford recalls 13 million tires," 22 May 2001. [Online]. Available: http://money.cnn.com/2001/05/22/recalls/ford/. [Accessed 5 November 2014].

[94] S. Kane *et al.*, "Update Report: Toyota Sudden Unintended Acceleration," Safety research and Strategies, Inc., [Online]. Available: http://www.safetyresearch.net/toyota-sudden-unintended-acceleration/. [Accessed 5 November 2014].

[95] U.S. Department of Transportation, "Technical Assessment of Toyota Electronic Throttle Control (ETC) Systems," 2011.

[96] Toyota, "Q3 2010 Earnings Call," 4 February 2010. [Online]. Available: http://www.123jump.com/earnings-calls/Toyota-Q3-2010-Earnings0Call-Transcript/36508/1. [Accessed 5 November 2014].

[97] T. Frank, "Safety last: Lies and coverups mask roots of small-plane carnage," 16 June 2014. [Online]. Available: http://www.usatoday.com/story/news/nation/2014/06/12/lies-coverups-mask-roots-small-aircraft-carnage-unfit-for-flight-part-1/10405323/. [Accessed 5 November 2014].

[98] B. Schoettle and M. Sivak, "A survey of public opinion about autonomous and self-driving vehicles in the U.S., the U.K., and Australia," The University of Michigan Transportation Research Institute. Report No. UMTRI-2014-21, 2014.

[99] S. Govindaraj *et al.*, "Market overreaction to product recall revisited-The case of Firestone Tires and the Ford Explorer," *Review of Quantitative Finance and Accounting,* vol. 23, no. 1, pp. 31-54, 2004.

[100] D. Raymer, Aircraft Design: A Conceptual Approach, Reston, VA: American Institute of Aeronautics and Astronautics, 2006.

[101] Airfoil Tools, "NACA 22112," [Online]. Available: http://airfoiltools.com/airfoil/details?airfoil=naca22112-jf. [Accessed 7 April 2014].

BIOGRAPHICAL SKETCH

Garrett Waycaster was born in Birmingham, AL in 1987. He graduated with a Bachelor of Science in mechanical engineering in 2009 followed by a Master of Science in mechanical engineering in 2010 from the University of Alabama. He worked in Vancouver, WA as an R&D mechanical design engineer at Hewlett-Packard as a member of the retail publishing solutions group in 2011. He enrolled as a doctoral student under Dr. Raphael T. Haftka and Dr. Nam-Ho Kim as a part of the Structural and Multidisciplinary Optimization group at University of Florida in 2012. During the summer of 2012, Garrett worked as an intern at Cessna Aircraft in Wichita, KS where he helped to develop a method for probabilistic optimization of manufacturing tolerances. In the summer of 2014, he visited the Institut Clément Ader at Université Paul Sabatier in Toulouse, France under the supervision of Dr. Christian Bes and Dr. Christian Gogu and will receive a dual PhD degree through the university. His research interests include design optimization for market systems, uncertainty quantification, and data modeling and analysis.