ESSAYS ON DECISION PROBLEMS UNDER UNCERTAINTY

by

Mahsa Mardikoraem

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Management Science

at
The University of Wisconsin-Milwaukee

August 2021

ABSTRACT

ESSAYS ON DECISION PROBLEMS UNDER UNCERTAINTY

by

Mahsa Mardikoraem

The University of Wisconsin-Milwaukee, 2021
Under the Supervision of Professor Ehsan S. Soofi

One of the fundamental questions in many operations and decision problems is how to incorporate available information into the decision-making process in uncertain environments. In two essays, we develop tools for operations decision problems. In the first essay, we consider the case when only partial information (mode and a few percentiles) about the probability distribution of the variable of interest is available. In the second essay, we consider the case when a sample observation from the variable of interest and its predictors, along with prior distributions about the parameters of the probability model, are available.

Essay 1. Maximum entropy distributions with mode and quantiles for applications to operations problem

We use maximum entropy (ME) modeling to derive probability distributions for management science applications when partial information consists of quantiles and mode. In Bajgiran et al. (2021), we developed ME modeling when partial information includes quantiles and moments. Previous research has shown that respondents are able to assess the mode and quantiles more accurately than the mean and variance. The ME methodology, given mode and other partial information, was developed in the 1990s. This essay provides formulas for implementing this methodology, which depends on the relationships between the mode and each pair of given consecutive quantiles.

This essay presents applications that include two management science problems. The first application is a model for the stochastic project planning using Project

Evaluation and Review Technique (PERT). Recently, there was an attempt to use ME distribution for PERT with mode information. We present a more extensive information-theoretic modeling for PERT. We find distributions of the duration of the paths in a network of multiple tasks by convolutions of the ME models of the duration of activities at each path. Given the distributions of the paths, the minimal and maximal distributions of project completion time are derived. We also develop a new probabilistic approach for computing the distribution of the expected completion time by incorporating Dirichlet prior distribution. We report applications of our methodologies using real-world data for two pre-construction projects based on information provided by a national construction firm.

The second application develops ME models for demand distribution in the Newsvendor (NV) problem. In NV, a profit maximizing solution is a quantile of the demand distribution. When the demand distribution is unknown, some rules (e.g. minimax regret) are used to derive optimal order quantity based on partial information. Then some arbitrary probability models are chosen for the demand distribution where its profit maximizing quantile can be different from the proposed optimal order quantity. This essay illustrates the perfectly robust ME models for the demand distribution given by the NV minimax regret rule when partial information includes the mode with and without the median.

Essay 2. Bayesian prescriptive framework for complementary products: tariff strategies and channels inefficiencies

This essay considers the problem faced by an omnichannel retailer of complementary products that encountered on average a 35 percent increase in tariffs and attempted to pass a portion of tariff cost to customers through price and freight charge. The omnichannel retailer offers various integrated channels and touchpoints, including media, purchase, and delivery channels. We develop a new Bayesian predictive and prescriptive model for omnichannel retailers to study revenue management strategies of complementary products for the tariff when demand is not observable.

We define retailer sales inefficiency as the deviation of demand (maximum potential sales) from observed sales. Inefficiency can be due to various reasons, including shortage of products and salespersons' performance; occurrences of these cases are not recorded. We develop a new Stochastic Frontier Model (SFM) to model sales in the presence of inefficiency. The basic SFM in the econometrics literature assumes that firms are inefficient in meeting the frontier. We combine two existing extensions of the basic SFM to model sales of complementary products by a system of zero-inflated SFMs. We use copula to model the dependency of models for the different products. The use of a zero-inflated distribution for the inefficiency relaxes the restrictive assumption of the basic SFM to allow for the possibility of an efficient firm.

We implement the model using data provided by a retailer that operates nationwide. In our empirical model, the demands for each complementary product are specified as a stochastic function of price, freight charge, discounts, tariff, and media channels (advertising, catalogs, and website visits). The inefficiency is a stochastic function of delivery channels and purchase channels characteristics, including salesperson's performance and contract types. Our results provide the probability of the firm's full efficiency and probability distribution for the inefficiency of each purchase channel.

We use the Bayesian approach for the inference, which includes posterior intervals and the Bayes Factor for evidence about directional hypotheses for predictors of the demands and inefficiencies of the delivery and purchase channels. Moreover, the Bayesian approach enables deriving a predictive probability distribution of the profit. This approach provides more insightful decision-making tools for managers in complex environments. These tools are more versatile than the common approach of making decisions based on estimates of the expected profit. We show applications of our model to study various scenarios of passing a portion of tariff cost to customers. The results enable examine stochastic ordering between profit distributions of various scenarios to select the best scenario.

To

Women

,

LGBTQ community

and

many others who have to fight more . . .

TABLE OF CONTENTS

Abstract **ii**

List of figures **x**

List of tables **xii**

Acknowledgements **xiii**

1 Maximum entropy distributions with mode and quantiles for applications to operations problem **1**

1.1 Introduction . . . 1

1.1.1 Motivation and objective . . . 2

1.1.2 Information optimal models . . . 5

1.2 Optimal information models with a given mode . . . 9

1.2.1 Mode with moments . . . 9

1.2.2 Mode with quantiles . . . 15

1.3 Application to Newsvendor problem . . . 21

1.3.1 Robust ME model for minimax regret rule with mode . . . 24

1.3.2 Robust ME model for minimax regret rule with mode and median 27

1.4 Application to PERT . . . 29

1.4.1 ME approach to PERT . . . 34

1.4.2 Maximum, minimum, and expected distributions of paths . . . 37

1.4.3 ME models for construction projects . . . 40

1.4.3.1 Completion time of the construction project . . . 41

1.4.3.2 Completion time of the pre-construction project . . . 42

1.5 Concluding remarks . . . 44

2 Bayesian prescriptive framework for complementary products: tariff strategies and channels inefficiencies **48**

2.1 Introduction . . . 48

2.1.1 Literature review - Operations . . . 53

2.1.2 Literature review - Methodology . . . 56

2.2 Original SFM and its existing extensions . . . 59

2.2.1 Variants of original SFM . . . 60

2.2.2 Zero-inefficiency SFM . . . 61

2.2.3 Multi-output SFM . . . 62

2.3 New SFM and prescriptive analysis 64
2.3.1 Multi-product, zero-inefficiency SFM 65
2.3.1.1 Modeling omnichannel retailing 65
2.3.1.2 MZI-SFM specification 67
2.3.1.3 Bayesian inference 69
2.3.2 Predictive distribution of profit 73
2.3.3 Notations and variables of MZI-SFM 74
2.4 Empirical results 74
2.4.1 Empirical results 77
2.4.1.1 Model parameters 77
2.4.1.2 Comparison of purchase channels 81
2.4.1.3 Comparison of delivery channels 84
2.4.2 Revenue management strategies 85
2.5 Concluding remarks and future research 89
2.5.1 Summary of findings and implications 91
2.5.2 Future research 95

3 Appendix 111
3.1 CDF of compound deviation of the performance from the frontier (Essay 2) 111
3.2 Variables descriptions (Essay 2) 112

4 Curriculum Vitae 113

LIST OF FIGURES

1.1 ME density and CDF with mode and finite range. 13

1.2 The MDI PDFs and CDFs with the goal PDF g_1 given mode ($m = 0.5$). 13

1.3 The MDI PDFs and CDFs with the goal PDF g_2 given mode ($m = 0.5$). 14

1.4 ME density and CDF with a quantile ($q = 0.3$), finite range, without smoothening function (blue), and with smoothening function g_2 with $\sigma = 0.2$ (orange). 15

1.5 The ME PDF and CDF with given mode ($m = 0.4$) and median ($M = 0.5$). 18

1.6 The MDI PDFs and CDFs with the goal PDF g_1 and given mode ($m = 0.4$) and median ($M = 0.5$). 19

1.7 The MDI PDFs and CDFs with the goal PDF g_2 and given mode ($m = 0.4$) and median ($M = 0.5$). 19

1.8 The ME PDF and its CDF with quantiles ($q_1 = 0.33, \alpha_1 = 0.2$, median $M = 0.67$), and mode ($m = .5$). 20

1.9 The MDI PDFs and CDFs with the goal PDF g_1 and quantiles ($q_1 = 0.33, \alpha_1 = 0.2$, median $M = 0.67$), and mode ($m = .5$). 21

1.10 Robust PDFs and CDFs of demand with the goal PDF g_1, given mode $m = 1/3$ and minimax regret order quantity $q^* = 0.22$ (left), 0.33 (middle), 0.43 (right) for $\beta = 0.8, 0.6, 0.4$, respectively. 26

1.11 The MDI PDF and CDF of demand with the goal PDF g_1 given mode ($m = 0.5$), median ($M = 0.6$) and minimax regret order quantity $q^* = 0.47$ for $\beta = 0.7$. 29

1.12 Project flow. 30

1.13 Distributions of completion times for the paths a-d (left), b-e (middle), and a-c-e (right) using ME distributions for activities (blue) and beta distributions for activities (orange). 36

1.14 Survival functions of completion time of paths using ME distributions for activities (left) and PERT distributions using beta (right). 36
1.15 Survival functions of minimal, expected under Dirichlet prior, and maximal for completion time of paths. 40
1.16 Construction project completion time PDF and survival function using ME approach and g_1 smoothing function (blue), and beta distribution (orange) 42
1.17 Pre-construction completion time PDFs and survival functions using ME and beta distributions. 44

2.1 Development of the Multi-output-Zero-Inefficiency SFM and Bayesian prescriptive analysis. 65
2.2 Survival functions of the inefficiency for the orders received for product A (left) and product B (right) through purchase channels. 84
2.3 Survival functions of the posterior distributions of the probability of full efficiency for delivery of products A (left) and B (right) by three distribution channels. 85
2.4 Predictive distributions of profit for passing tariff cost to customers through adjusting the price and/or freight charge according to five scenarios. . . . 87
2.5 Predictive distributions of profit for passing and not passing tariff costs to customers through increasing price and/or discount for either one or both products according to five scenarios. 89

LIST OF TABLES

1.1 Comparison between the profit maximizing quantile q of the ME distribution given range $[0, 300]$ and mode $m = 100$ and the minimax regret order quantity q^*. 25

1.2 Comparison between the profit maximizing quantile q of the ME distribution given range, mode ($m = 150$) and median ($M = 180$ and 90), and the minimax regret quantity q^*. 29

1.3 Activities duration data. 34

1.4 Percentiles of the minimal, maximal, and expected distributions. 39

1.5 Construction project data. 40

1.6 Construction project completion time percentiles. 42

1.7 Pre-construction project's activity duration data (months). 43

1.8 Completion time of pre-construction project's percentiles. 44

2.1 Jeffreys' scale for the grade of evidence (Kass & Raftery, 1995). 72

2.2 Description and notations for the variables of the profit function. 74

2.3 Description and notations for the variables of MZI-SFM. 75

2.4 Posterior statistics for parameters of frontier. 79

2.5 Posterior statistics for inefficiency variables. 80

2.6 Posterior statistics for the inefficiency purchase channels. 83

2.7 Summary statistics for predictive distributions of profit for passing and not passing tariff costs to customers through increasing price or freight for either one or both products according to five scenarios. 87

2.8 Summary statistics for predictive distributions of profit for passing and not passing tariff costs to customers through increasing price and/or discount for either one or both products according to five scenarios. 89

ACKNOWLEDGEMENTS

Throughout the writing of this dissertation I have received a great deal of support and assistance.

I am thankful to my dissertation advisor, Ehsan S.Soofi, for his infinite guidance and support throughout the course of my Ph.D. studies. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. He helped me to be a better researcher and a better person.

I am also very grateful to Layth C.Alwan for guiding me through my professional development. I would also like to thank my other committee members, Amit Bhatnagar, Kaan Kuzu, and Refik Soyer, for their support, for taking the time to read and evaluate my work, and for their insightful comments.

Most importantly, none of this would have been possible without the unconditional love and support of my family in all aspects of my life. I would like to thank my mother Farideh, my father Gholam, my sisters Mehrsa and Melina, for being the source of encouragement, patience, and strength throughout my research studies. Finally, I could not have completed this dissertation without the support of my partner, Mohsen.

Chapter 1

Maximum entropy distributions with mode and quantiles for applications to operations problem

Keywords: Mode information; Quantiles; Maximum Entropy; PERT; Newsvendor; Bayesian priors

1.1 Introduction

Assigning probability distribution in the absence of information in favor of particular outcomes dates back to Bernoulli and Laplace. The principle of insufficient reason, also known as the principle of indifference, stipulates assigning equal probabilities unless there is a reason to assign unequal probabilities. Jaynes (1957) extended this to the maximum entropy (ME) principle, which stipulates assigning probabilities by minimally adjusting the uniform distribution in light of partial information about the variable of interest. This essay follows the ME principle to develop probability models for management science problems when partial information is given in terms of mode and quantile of the probability distribution.

1.1.1 Motivation and objective

In various operations decision problems, partial information about a probability distribution is assumed in the forms of range (minimum, maximum), moments (mean, variance), quantiles, and mode. A few examples are as follows. The Newsvendor (NV) problem concerns with methods for the derivation of an optimal order quantity. The NV profit function is asymmetrically linear and the optimal solution is a quantile of the demand distribution. Scarf's rule is based on the mean and variance of the demand distribution (Scarf, 1958). Solutions to the minimum regret are derived based on various combinations of the range, mean, median, mode, and variance of demand distribution (Perakis & Roels, 2008). The Program Evaluation and Review Technique (PERT) for stochastic project planning problems is based on the specification of the optimistic (minimum), pessimistic (maximum), and most likely value (mode) of the project completion time (Malcolm et al., 1959). Partial information in various forms is solicited from experts for constructing Bayesian prior distributions (Winkler, 1968; Sun & Berger, 1998; Van Dorp & Mazzuchi, 2021). Survey instruments solicit forecast probabilities for a set of given intervals of a random variable, in addition to a question that asks the respondent to give a point or an interval estimate for the forecast. Central banks and other institutions use the results of these surveys for producing economic outlook reports for decision makers. Scholars use these results to study uncertainty and disagreement of economic forecasters.

Partial information does not provide a unique probability distribution for decision and inference. Consequently, researchers use parametric probability distributions which satisfy the partial information. The choice is made arbitrarily, instead of systematically according to a principle. According to the ME principle, parametric distributions fitted to partial information can induce extraneous information into the results or exclude existing information (Bajgiran et al., 2021). However, ME distributions are "maximally noncommittal with regard to missing information" (Jaynes, 1957).

The objective of this essay is to use the ME principle for developing probability models based on partial information for a few management science problems. More specifically, the focus of this essay is on problems where the partial information includes the mode and quantile. Researchers over several years have conducted experiments on the ability of individuals to estimate the mean, median, and mode of a distribution (Garthwaite et al., 2005). In these experiments, subjects were asked to estimate measures of the central tendency of the distribution based on a sample of numbers. When the sample was drawn from an approximately symmetric distribution, subjects were able to estimate these measures with a high degree of accuracy. When the sample was drawn from a highly skewed distribution, "assessments of the median and mode were again reasonably accurate, but assessments of the mean were biased towards the median" (Garthwaite et al., 2005).

Survey instruments that ask respondents to assign probabilities to a set of pre-assigned intervals provide quantile information about the respondents' probability distributions of the underlying stochastic object of interest. This method is called fixed interval solicitation. Solicitation of subjective probabilities is also done by the variable interval method, where the subject assesses points corresponding to given probabilities (Garthwaite et al., 2005). An effective variable method is called bisection, where the subject is asked to assess the median, the medians of the lower and upper halves (first and third quartiles), and so on.

Engelberg et al. (2009) present a comprehensive study of the Federal Reserve of Philadelphia's Survey of Professional Forecasters (SPF) that examines the consistency of respondents' probability and point forecasts. The SPF began in 1968, but these authors used data from Quarter 1 of 1992 to Quarter 4 of 2004, in which the respondents are asked to assign probabilities to eight intervals of equal lengths of 1 percent and two open-end intervals for the inflation rate and gross domestic product (GDP) growth of the current year and the following year. The authors assumed "that the mode is contained in the interval with the greatest probability mass" (modal interval of the histogram) and developed bounds for the mean and median. They report that 2,930 of 3,173 respondents' point forecasts fell within the modal intervals of the respondents' histograms. They showed that for all four quarters aggregated over the years, the percentages of respondents whose point forecasts of the inflation rate were within the modal intervals was higher than for the mean and median falling within their bounds. For the GDP growth, the percentages of respondents whose point forecasts of the inflation rate were within the modal intervals were higher than those for the mean and median falling within their bounds for the first three quarters.

Soofi et al. (2009) analyzed the data collected via a survey of business executives after the 9/11 terrorist attack. A question asked respondents to assign probabilities to three categories of the GDP growth in third quarter of 2002 (a year after the attack) and a different question asked to indicate the GDP growth on a more refined scale. These authors observed that 87 of the 93 respondents' forecast categories fell into the modal intervals of their distributions. This and the Engelberg et al.'s (2009) study confirm that respondents more often think of mode for the center of a distribution.

1.1.2 Information optimal models

The information theory provides methodologies for deriving unique distributions that are solely consistent with given partial information. The literature in this area has mainly considered developing distributions based on various types of moments (Jaynes, 1957; 1968). A few papers have used mode (Brockett et al., 1995; Mauris, 2010; Hernández-Bastida & Fernández-Sánchez, 2019). Quantiles also have been used in a few papers (Hernández-Bastida & Fernández-Sánchez, 2019; Bajgiran et al., 2021). The use of both quantiles and mode has been also limited (Brockett et al., 1995; Brockett et al., 1997).

Two information measures are used for deriving models. The Shannon entropy of a quantity X with continuous distribution F and a probability density function (PDF) f is defined by

$$H(X) = H(f) = -\int_{\mathcal{R}} f(x) \log f(x) dx, \tag{1.1}$$

where $\mathcal{R}$ is the support of f, provided that the integral is finite. The Kullback-Leibler information divergence between f and g is defined by

$$K(f:g) = \int_{\mathcal{R}} f(x) \log \frac{f(x)}{g(x)} dx \geq 0, \tag{1.2}$$

provided that f is absolutely continuous with respect to g ($f(x) = 0$ whenever $g(x) = 0$); the inequality becomes equality if and only if $f(x) = g(x)$ almost everywhere. Unlike the entropy, $K(f:g)$ is invariant under one-to-one transformations of X for the continuous case. This measure has been used in operations and decision problems (see, for example, Alwan et al., 1998; Saghafian & Tomlin, 2016; Asadi et al., 2018). The Minimum Discrimination Information (MDI) principle minimizes $K(f:g)$ where f is in a class of distributions Ω and the reference PDF $g \notin \Omega$. When g is uniform (proper PDF or improper), the MDI coincides with the ME principle (Jaynes, 1957). For the continuous case, the ME is defined as minimization of $-K(f:g)$, where g is interpreted as the "invariance measure" function (Jaynes, 1968). In the game theory problem of selecting a distribution where the utility is defined in terms of the score function, the MDI and ME models are minimax decisions (Smith, 1974; Grünwald & Dawid, 2004).

In Bajgiran et al. (2021) we considered information models in the following classes of distributions:

$$\Omega_\theta = \{f_\theta : E_f[T_j(X)] = \theta_j,\ j = 1, \cdots, J\}\,, \tag{1.3}$$

$$\Omega_\alpha = \{f_\alpha : E_f[\mathbf{1}(X \leq q_k)] = \alpha_k, k = 1, \ldots, K\}\,, \tag{1.4}$$

$$\Omega_{\theta,\alpha} = \{f_{\theta,\alpha} : E_f[T_j(X)] = \theta_j,\ E_f[\mathbf{1}(X \leq q_k)] = \alpha_k\}\,, \tag{1.5}$$

where f is a PDF on a continuous support $\mathcal{R}$, $\theta_j,\ j = 1, \cdots, J$ are given moments, $\mathbf{1}(E)$ is the indicator function of the event E, and $q_k = F^{-1}(\alpha_k), k = 1, \ldots, K$ are given quantiles.

Optimal information models in (1.3) are studied extensively in the literature used in various problems in many fields. The ME model in Ω_θ, if exists, is unique and has PDF in the following form:

$$f^*_\theta(x) = C_\lambda \exp\left\{-\sum_{j=1}^{J} \lambda_j T_j(x)\right\}, \quad x \in \mathcal{R}, \tag{1.6}$$

where $\lambda_1, \ldots, \lambda_J$ are Lagrange multipliers and

$$C_\lambda = C(\lambda_1, \ldots, \lambda_J) = \left[\int_{\mathcal{R}} \exp\left\{-\sum_{j=1}^{J} \lambda_j T_j(x)\right\} dx\right]^{-1} < \infty, \tag{1.7}$$

is the normalizing factor for f^*_θ, and $\theta_j = -\partial \log C_\lambda / \partial \lambda_j$, see, for example, Soofi et al. (1995). The existence of the ME model is determined by the finiteness condition (1.7). Many families of probability distributions are ME models with various types of moment information. Some of the known examples include the uniform distribution

$[A, B]$ with the range information $E_f[\mathbf{1}(A \leq X \leq B)] = 1$, beta distribution [0,1] with two geometric means: $T_1(x) = \log x, T_2(x) = \log(1-x)$; on the non-negative range, the exponential distribution with $T_1(x) = x$, the gamma distribution with $T_1(x) = x$ and $T_2(x) = \log x$, and the log-normal distribution with $T_1(x) = \log x, T_2(x) = (\log x)^2$; and on the unrestricted range, the ME model with $T_j(x) = x^j, j = 1, \ldots, J$ when J is an odd number does not exist. The ME model with $J = 2$ is a normal distribution. The MDI model in Ω_θ relative to a PDF $g(x)$, if exists, is unique and has PDF in the following form:

$$f_\theta^*(x) = C_\eta g(x) \exp\left\{\sum_{j=1}^{J} \eta_j T_j(x)\right\}, \quad x \in \mathcal{R}, \tag{1.8}$$

where $\eta_1, \ldots, \eta_J$ are Lagrange multipliers and C_η is the normalizing factor. When g is uniform, (1.8) gives (1.6), hence the ME model is closest to the uniform distribution that satisfies the partial information in (1.3).

The ME model in Ω_α only exists on finite support and has PDF in the following form:

$$f_\alpha^*(x) = C_\lambda \exp\left\{-\sum_{j=1}^{K} \lambda_j \mathbf{1}(x < q_j)\right\}, \; A < x < B, \tag{1.9}$$

where $A < q_1 < \cdots < q_K < B$. This is a piecewise uniform PDF (a density histogram with unequal bins); (Bajgiran et al., 2021). The ME distributions in $\Omega_{\theta,\alpha}$ has the following form:

$$f_{\theta,\alpha}^*(x) = C_{\zeta,\lambda} \exp\left\{-\sum_{j=1}^{J} \lambda_j T_j(x) - \sum_{k=1}^{K} \zeta_k \mathbf{1}(x < q_k)\right\}, \; -\infty < x < \infty,$$

where $C_{\zeta,\lambda}$ is determined by the Lagrange multipliers $(\lambda_1, \ldots, \lambda_J, \zeta_1, \ldots, \zeta_K)$ for the moment and quantiles constraints in $\Omega_{\theta,\alpha}$. Optimal information models in (1.4) and (1.5) are studied and applied to developing Bayesian priors, forecast distributions, and the NV problem by Bajgiran al. (2021).

This essay presents the ME and MDI models with a given mode in the sets of partial information defined in (1.3)-(1.5) and illustrate their applications in the NV problem and PERT.

1.2 Optimal information models with a given mode

This section gives an overview of the methodology from Brockett, Charnes, & Paick (1995), hereafter written as BCP, with some clarifications and extensions.

1.2.1 Mode with moments

BCP derived the unimodal ME PDF using a well-known characterization of the unimodal density due to Khintchine (1938). According to Khintchine's theorem, a "zero" unimodal random variable X can be characterized as $X = UY$, where U and X are independent and U is uniformly distributed over $[0, 1]$; see Jones (2002) for other constructive illustrations of this Khintchine's theorem.

Kemperman (1971) showed that moments of Y can be computed from moments of X as follows. For any function $h(x)$, there exist a function $h^*(y)$ such that

$$\int h(x) f_X(x) dx = \int h^*(y) f_Y(y) dy, \tag{1.10}$$

where

$$h^*(y) = E_U \left[h(UY) | Y = y \right] = \frac{1}{y} \int_0^y h(t) dt. \tag{1.11}$$

If X is unimodal with mode m, then $X - m$ is zero unimodal. Application of (1.10) to a moment of $X - m$ gives

$$E_Y[T_j^*(Y)] = E_X[T_j(X)], \tag{1.12}$$

where $T_j^*(Y)$ is the transformed moment for the auxiliary variable Y, given by (1.11) as

$$T_j^*(y) = \frac{1}{y} \int_0^y T_j(t+m) dt. \tag{1.13}$$

For example, BCP presented the moment case $T_j(x) = x^j$ which transforms as follows:

$$T_j(y) = \sum_{k=0}^{j} \binom{j+1}{k} \frac{x^k m^{j-k}}{j+1}.$$

Consider the following class of distributions:

$$\Omega_\theta^* = \left\{ f_Y : E_Y[T_j^*(Y)] = \theta_j,\ j = 1, \cdots, J \right\}.$$

By (1.6), the PDF of the ME model in Ω^*_θ is

$$f^*_Y(y) = C_\lambda \exp\left\{-\sum_{j=1}^{J}\lambda_j T^*_j(y)\right\}. \tag{1.14}$$

The existence of the ME PDF is determined according to (1.7) with $T^*_j(y)$.

The relationship in (1.12) gives the following subclass of probability distributions:

$$\Omega_{\theta,m} \subset \Omega_\theta = \left\{f_X : E_X[T_j(X)] = \theta_j,\ \arg\max f_X(x) = m, j = 1, \cdots, J\right\}, \tag{1.15}$$

where Ω_θ is defined by the moments given in (1.3) without the additional mode constraint.

BCP found the ME model $f^*_X \in \Omega_{\theta,m}$ by viewing (1.13) as a one-to-one transformation between X and the auxiliary variable Y, and the inverse transformation f^*_Y. This approach gives

$$f^*_X(x) = \begin{cases} \int_{-\infty}^{x-m} C_\lambda \exp\left(-\sum_{j=0}^{J}\lambda_j T^*_j(y)\right)\frac{dy}{|y|}, & x \le m, \\ \int_{x-m}^{\infty} C_\lambda \exp\left(-\sum_{j=0}^{J}\lambda_j T^*_j(y)\right)\frac{dy}{y}, & x \ge m. \end{cases} \tag{1.16}$$

The ME PDFs f^*_Y and f^*_X are not smooth at the mode. BCP derived smoothened versions of these models through computing the MDI model with a reference a PDF g in

(1.8). In BCP, g is called the "goal density" and the following PDF's are used:

$$g_0(y) = 1 \tag{1.17}$$

$$g_1(y) = \begin{cases} \exp\left[-(|y|-\delta)^2 - \left(\frac{1}{|y|} - \frac{1}{\delta}\right)^2\right] & \text{for } |y| \leq \delta \\ 1 & \text{for } |y| \geq \delta, \end{cases} \tag{1.18}$$

$$g_2(y) = \frac{y^2}{\sqrt{2\pi}\sigma^3} \exp\left(\frac{-y^2}{2\sigma^2}\right), \quad \sigma > 0. \tag{1.19}$$

On an unbounded support, g_0 and g_1 are not proper PDFs. The uniform goal PDF leaves f_Y^* and f_X^* unchanged. The goal PDF g_1 deviates from the uniform PDF locally in a neighborhood $|y| \leq \delta$ and smoothly approaches zero as $|y| \to 0$. The MDI PDF for X with reference g_1 becomes smooth in a neighborhood of the mode. The goal PDF g_2 seeks the MDI PDF of X in $\Omega_{\theta,m}$ to be close to the PDF of the normal distribution. This goal PDF is the PDF of the random variable $Y = \sqrt{V}$, where distribution of V is Chi-Square with three degrees of freedom.

The following example illustrates the ME PDF with a given mode and its smoothened versions.

Example 1.1 (Unimodality, mode, and finite range)**.** Consider the case when partial information only includes given mode. Then (1.16) gives

$$f_X^*(x) = \begin{cases} \int_{-\infty}^{x-m} C_\lambda \frac{dy}{|y|}, & x \leq m, \\ \int_{x-m}^{\infty} C_\lambda \frac{dy}{y}, & x \geq m. \end{cases}$$

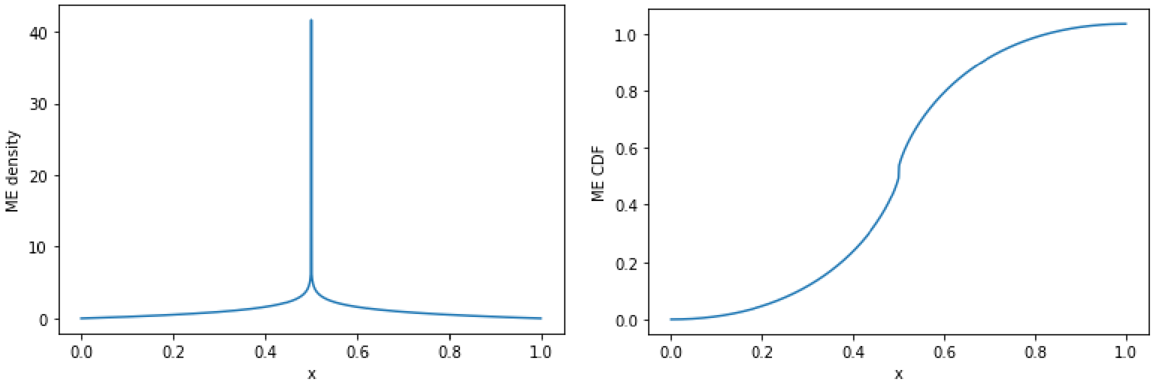

FIGURE 1.1: ME density and CDF with mode and finite range.

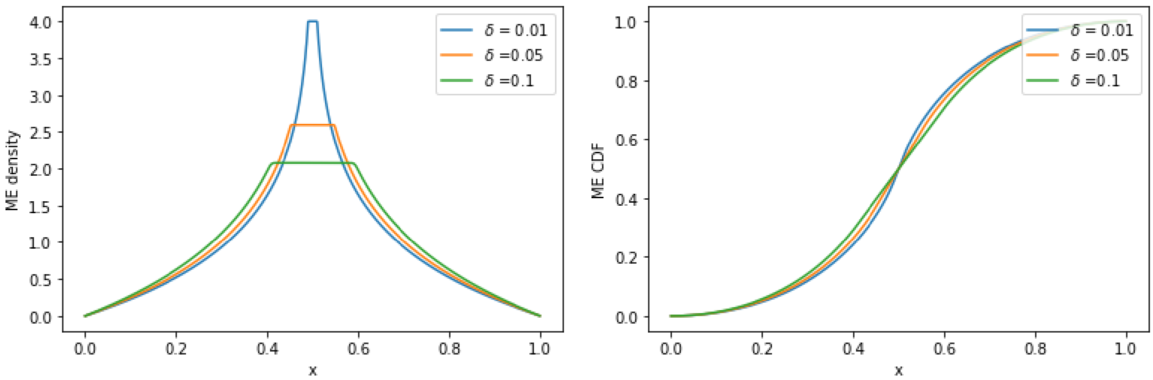

FIGURE 1.2: The MDI PDFs and CDFs with the goal PDF g_1 given mode ($m = 0.5$).

This is not a proper PDF on an unbounded support. Suppose that the range of X is the unit interval $[0, 1]$. Then f_Y^* is uniform and the transformed PDF is

$$f_X^*(x) = \begin{cases} \log \frac{m}{m-x} & \text{if } 0 \le x < m, \\ \log \frac{1-m}{x-m} & \text{if } m < x \le 1. \end{cases} \tag{1.20}$$

This PDF is not defined at the mode. The left panel of Figure 1.1 presents f_X^*. The infinite discontinuity at the mode is apparent. However, cumulative distribution function (CDF), F_X^*, is continuous as shown in the right panel of Figure 1.1. This model is used by Hernández-Bastida & Fernández-Sánchez (2019) for PERT application. Figure

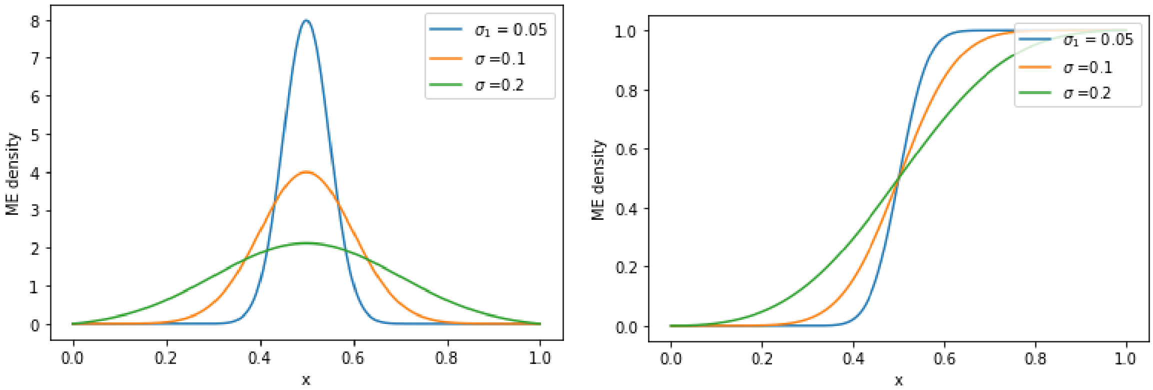

FIGURE 1.3: The MDI PDFs and CDFs with the goal PDF g_2 given mode ($m = 0.5$).

1.2 presents the smoothened version of (1.20) with goal PDF g_1 and its CDF. As seen in the left panel of Figure 1.2, the smoothing parameter δ controls the sharpness of the PDF near the mode. As δ decreases the PDF grows sharper around its mode. Figure 1.3 presents the smoothened version of (1.20) with goal PDF g_2 and its CDF. In the absence of any constraint, g_2 gives symmetric truncated normal distributions for the smoothened PDF. In this case, as the smoothing parameter σ decreases the PDF becomes sharper around mode.

In Bajgiran et al. (2021) we showed that the ME PDF with quantile constraints is discontinuous at the quantiles, while its CDF being continuous. However, the extension to the MDI approach with goal PDFs g_1 and g_2 is effective for smoothening the ME PDF with quantile constraints. The following example illustrates the ME distribution with a single quantile constraint and its smoothened version with the goal PDF g_2.

Example 1.2 (A quantile)**.** Consider Ω_α defined in (1.4) with a single quantile, $q = 0.3$. The ME PDF (1.9) gives a two-piece uniform distribution with the change point at $x = 0.3$. The left panel of Figure 1.4 shows this ME PDF in blue its smoothened version

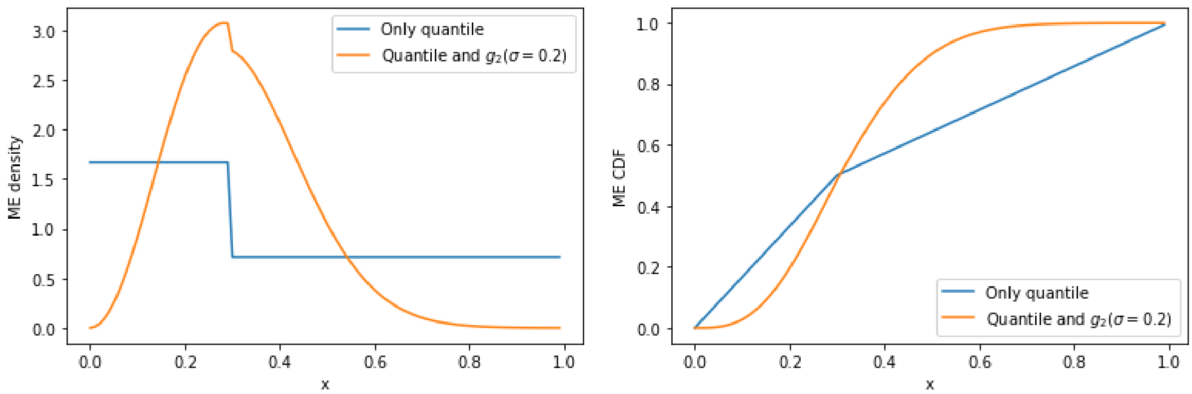

FIGURE 1.4: ME density and CDF with a quantile ($q = 0.3$), finite range, without smoothening function (blue), and with smoothening function g_2 with $\sigma = 0.2$ (orange).

with the goal PDF g_2 in orange. The smoothened PDF is also piecewise normal (and asymmetric), however with a smaller jump at the $x = .3$ and high concentration around the quantile. Corresponding CDF is shown in the right panel of Figure 1.4, which is continuous.

1.2.2 Mode with quantiles

BCP presented plots of ME PDFs with given mode, median, and interval information, and displayed transformation (1.13) for the following simple example:

$$P(X \geq q) = E_{X_{\theta,m}}[\mathbf{1}(X \geq q)] = E_{Y_\theta}[T^*_\beta(Y)] = \beta.$$

The transformed moment for this case is

$$T^*_\beta(y) = \frac{1}{y}\int_0^y \delta(t \geq q-m)\, dt = \begin{cases} 0 & y \leq q-m, \\ 1 - \frac{q-m}{y} & y \geq q-m. \end{cases} \tag{1.21}$$

This holds when $q - m \geq 0$ and suggests that transformations of quantiles depend on the relationship between the mode and the quantiles. We address the problem of including quantile constraints at the general level of relationships between the mode and a set of quantiles.

Consider the following class of probability distributions on a finite support:

$$\Omega_{m,\alpha} = \{f_{m,\alpha} : E_{m,\alpha}[\mathbf{1}(X \leq q_i)] = \alpha_i,\ \arg\max\ f_{m,\alpha}(x) = m,\ i = 1, \ldots, n\}, \tag{1.22}$$

where $A < q_1 < \cdots < q_n < B$; the finite support is required for the existence of the ME model in this class, and without loss of generality we use the unit interval.

Application of (1.13) gives the following transformed moments for three cases based on the relationships between the mode and the pairs of consecutive quantiles in (1.22):

1. For $m \leq q_i$,

$$T_i^*(y) = \frac{1}{y}\int_0^y \delta(q_i - m \leq t \leq q_{i+1} - m)\ dt = \begin{cases} 0 & y \leq q_i - m \\ 1 - \frac{q_i - m}{y} & q_i - m \leq y \leq q_{i+1} - m \\ \frac{q_{i+1} - q_i}{y} & y \geq q_{i+1} - m. \end{cases} \tag{1.23}$$

2. For $q_i \leq m \leq q_{i+1}$,

$$T_i^*(y) = \frac{1}{y}\int_0^y \delta(q_i - m \leq t \leq q_{i+1} - m)\ dt = \begin{cases} \frac{|q_i - m|}{y} & y \leq q_i - m, \\ 1 & q_i - m \leq y \leq q_{i+1} - m, \\ \frac{q_{i+1} - m}{y} & y \geq q_{i+1} - m. \end{cases} \tag{1.24}$$

3. For $q_{i+1} \leq m$,

$$T_i^*(y) = \frac{1}{y}\int_0^y \delta(q_i - m \leq t \leq q_{i+1} - m)\ dt = \begin{cases} \frac{q_{i+1} - q_i}{y} & y \leq q_i - m, \\ \left|\frac{q_{i+1} - m}{y} - 1\right| & q_i - m \leq y \leq q_{i+1} - m, \\ 0 & y \geq q_{i+1} - m. \end{cases} \tag{1.25}$$

These results are applicable to the problem with a single quantile and the mode as a special case. For example, (1.21) falls in the first case (1.23).

As noted earlier, researchers have found that individuals can assess median and mode more accurately than the mean. The following example gives the ME PDF and its

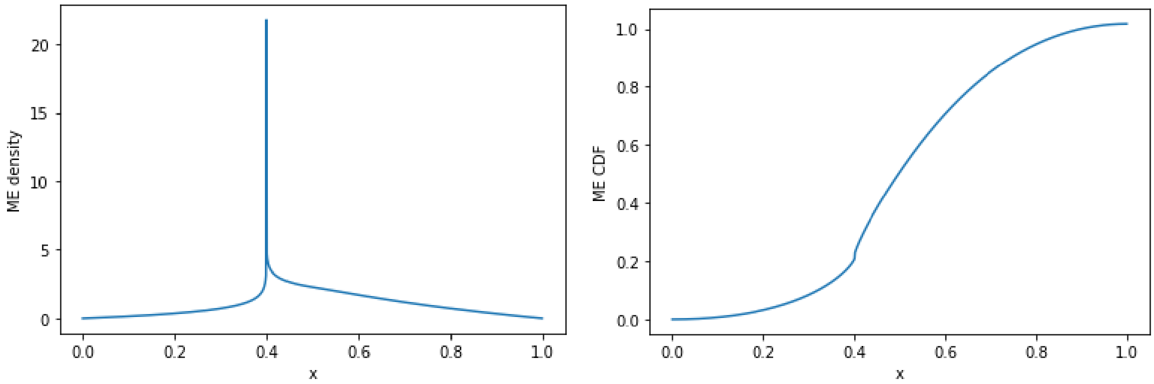

FIGURE 1.5: The ME PDF and CDF with given mode ($m = 0.4$) and median ($M = 0.5$).

smoothened versions for the important application where partial information consists of the mode and the median.

Example 1.3 (Mode and median)**.** Suppose that the partial information includes the mode m and the median M such that $m < M$. The transformed moment (1.21) is given by the first case (1.23) and ME PDF in $\Omega_{m,\alpha}$ is

$$f_X^*(x) = \begin{cases} \int_0^{x-m} C_\lambda \exp\left(-\lambda(1 - \frac{M-m}{y})\mathbf{1}(y \geq M-m)\right) \frac{dy}{|y|}x, & x \leq m, \\ \int_{x-m}^1 C_\lambda \exp\left(-\lambda(1 - \frac{M-m}{y})\mathbf{1}(y \geq M-m)\right) \frac{dy}{y}x, & x \geq m. \end{cases} \tag{1.26}$$

The ME model when $M < m$ can be obtained similarly with transformed moment (1.21) given by the first case (1.25).

Figure 1.5 shows f_X^* and its CDF when $m = 0.4$ and $M = 0.5$. This PDF has an infinite discontinuity at the mode. Figures 1.6 and 1.7 show smoothened versions of f_X^* with goal PDFs g_1 and g_2, respectively, and their CDFs. The PDF plots in these figures indicate that the smoothness issue of f_X^* is resolved to various extents.

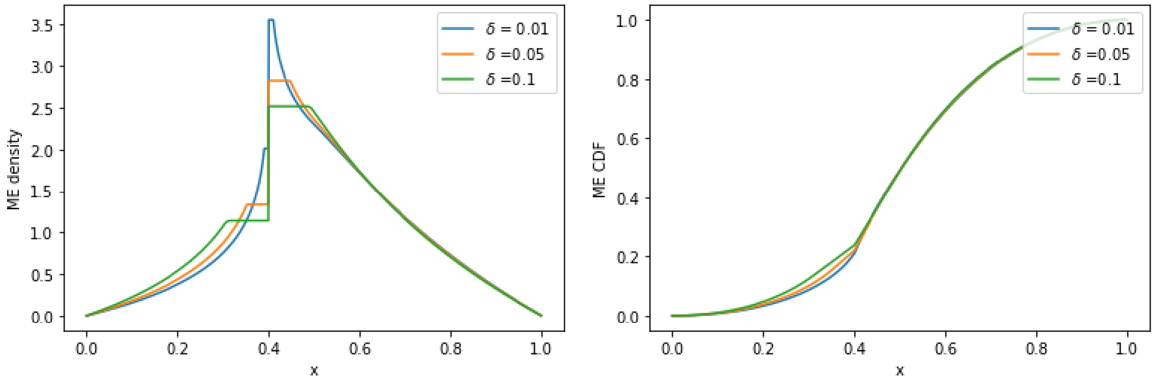

FIGURE 1.6: The MDI PDFs and CDFs with the goal PDF g_1 and given mode ($m = 0.4$) and median ($M = 0.5$).

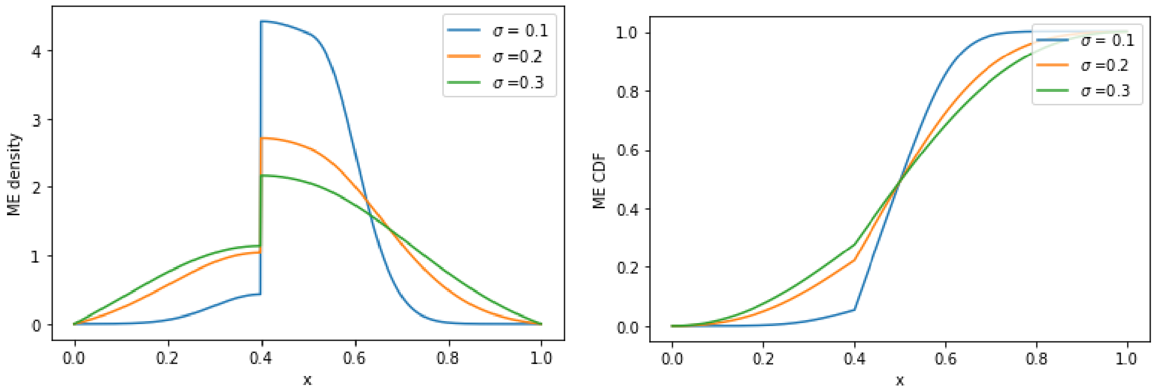

FIGURE 1.7: The MDI PDFs and CDFs with the goal PDF g_2 and given mode ($m = 0.4$) and median ($M = 0.5$).

The next example gives the ME PDF and its smoothened versions when the partial information consists of the mode and two quantiles.

Example 1.4 (Mode and two quantiles)**.** The partial information consists of the mode and two quantiles has been used to fit parametric models for variety of applications. For example, Van Dorp & Mazzuchi (2000) computed the parameters of the beta distribution based on the given mode, an upper quantile, and a lower quantile. Unlike the ME model, such a parametric model does not assign non-uniform probabilities solely based on the elicited information. Consider the case, when $q_1 \leq m \leq q_2$. The transformed moment

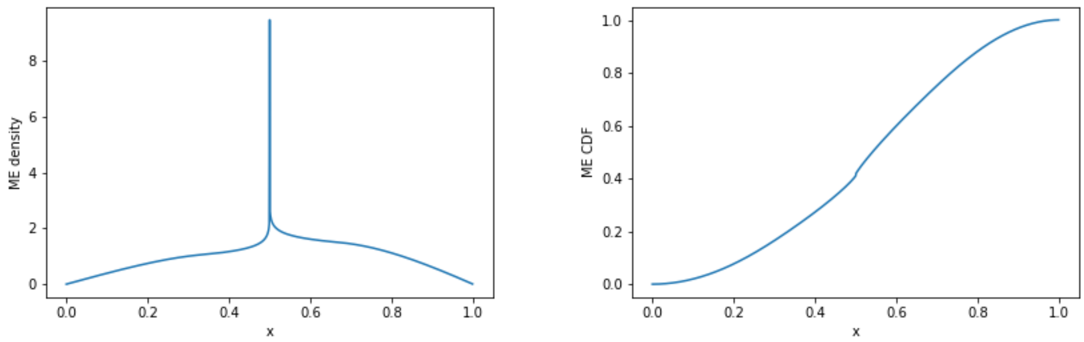

FIGURE 1.8: The ME PDF and its CDF with quantiles ($q_1 = 0.33, \alpha_1 = 0.2$, median $M = 0.67$), and mode ($m = .5$).

(1.21) is given by the first case (1.24) and ME PDF in $\Omega_{m,\alpha}$ is

$$f_X^*(x) = \begin{cases} \int_0^{x-m} C_\lambda \exp\Big(-\lambda_1(1-\frac{q_2-m}{y})\mathbf{1}(y \geq q_2 - m) \\ \qquad -\lambda_2[\frac{|q_1-m|}{y}\ \mathbf{1}(y \leq q_1 - m) + \mathbf{1}(y \geq q_1 - m)]\Big)\frac{dy}{|y|}, \quad x \leq m, \\ \int_{x-m}^{1} C_\lambda \exp\Big(-\lambda(1-\frac{M-m}{y})\mathbf{1}(y \geq M - m)\Big)\frac{dy}{y}, \qquad x \geq m. \end{cases} \tag{1.27}$$

Figure 1.8 shows f_X^* and its CDF for $q_1 = 0.33, \alpha_1 = 0.2$, the median $M = 0.67$, and $m = 0.5$. The infinite discontinuity at the mode is apparent. Figures 1.9 shows smoothened versions of f_X^* with the goal PDF g_1, and their CDFs. The PDF plots indicate that the smoothness issue of f_X^* is resolved to various extents.

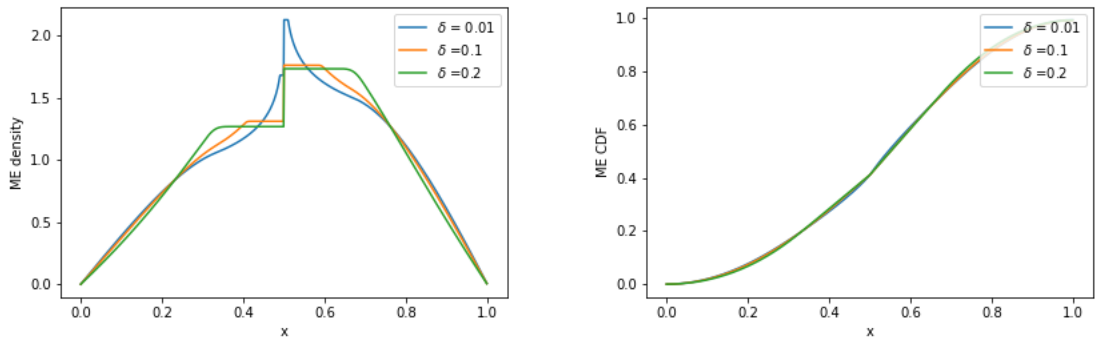

FIGURE 1.9: The MDI PDFs and CDFs with the goal PDF g_1 and quantiles ($q_1 = 0.33, \alpha_1 = 0.2$, median $M = 0.67$), and mode ($m = .5$).

Remark 1.1. The ME models on the support $[0, 1]$ can be easily transformed for applications to $[A, B]$ by the linear transformation: if $x \in [0, 1]$, then

$$y = A + (B - A), \quad x \in [A, B]. \tag{1.28}$$

The ME model on $[A, B]$ is the location-scale family with PDF

$$f_Y^*(y) = \frac{1}{B - A} f_X^* \left(\frac{y - A}{B - A} \right), \quad A \leq y \leq B. \tag{1.29}$$

1.3 Application to Newsvendor problem

The NV problem is defined by finding an optimal order quantity for a product with an uncertain demand X and fixed prices. Suppose that c is the cost per unit, r is the selling price per unit, and s is the salvage value per unit. Then $c_o = c - s > 0$ and $c_u = r - c > 0$ are the overage and underage costs, respectively. The expected profit maximizing order quantity is found by minimizing the risk (expected loss) of the decision d under the following asymmetric linear function:

$$L(d, X)] = c_o(d - X)\mathbf{1}(X \leq d) + c_u(X - d)\mathbf{1}(X > d).$$

The optimal decision that minimizes $E_f[L(d, X)]$ is given by

$$q = \arg\min_d E_f[L(d, X)], \tag{1.30}$$

where q and is the quantile of the distribution of X corresponding to $\alpha = c_u/(c_o+c_u)$; see, for example, Snyder & Shen (2011, p. 78). The risk of the decision, $E_f[L(q, X)] = \theta$, is the mean absolute deviation from the profit maximizing (risk minimizing) quantile q given by (1.30). The optimal solution is traditionally represented in terms of the βth upper quantile, $\beta = 1 - \alpha$. This optimization, however, does not offer a model for the unknown demand distribution, F. Therefore, the profit maximizing q and the profit (risk) of the decision remain unknown. For calculation and implementations purposes, it is assumed that decision makers have complete knowledge of the demand distribution in terms a parametric model.

The NV literature also offers alternatives to the profit maximizing order quantity. Optimal order quantities are derived based other criteria and assuming partial information about the demand distribution (Scarf, 1958; Gallego & Moon, 1993; Perakis & Roels, 2008). These rules are functions of the β and the assumed partial information. These approaches do not provide demand distribution. Instead, well-known probability models are chosen for the demand distribution as a separate task for the purpose of comparing the proposed solutions for the order quantity with the profit maximizing quantiles of the chosen distributions. A model for the demand distribution is said to be (perfectly) robust if its profit maximizing quantile is consistent with the optimal order quantity derived based on an alternative criterion.

The maximum entropy (ME) principle has been invoked in the NV literature for justifying the uniform, exponential, normal, and truncated normal distributions as models for the demand distribution based on partial information in terms of the range, mean, and variance of the demand. Perakis & Roels (2008) observed that, for the demand over a finite range, their proposed minimax regret rule coincides with the quantile of the uniform distribution, which is "consistent with the principle of insufficient reason (or the maximum entropy)". They noted that for the case of a given mean for the demand over the nonnegative range, "the exponential is expected to be the most robust, given that it is entropy-maximizing over the class of nonnegative distributions with known mean". They further pointed out that the ME principle provides justification for the normal distribution when the partial information includes the mean and variance. Andersson et al. (2013) compared the profit maximizing quantiles of ME models for the demand with given first two moments over a finite and nonnegative ranges with optimal order quantities given by the Scarf's and minimax regret rules, using numerical examples and simulations.

Because, an ME distribution is noncommittal to information which is not used in its derivation (Jaynes, 1968), there is no basis for its profit maximizing quantile to match the optimal order derived according to other criteria. Bajgiran et al. (2021) showed that for derivation of consistent distributions for the NV demand, the profit maximizing quantile must be included in the partial information. They constructed perfectly robust ME distributions for the Scarf's rule which assumes given mean and variance and for the minimax regret rules which assume a given mean and given mean and variance. These ME models for the demand distribution are derived by including the Scarf's and minimax optimal order quantities as the αth quantiles in of the respective partial information sets. In the current research, we construct perfectly robust ME models for the minimax regret rules with partial information given in terms of the mode with/without the median.

1.3.1 Robust ME model for minimax regret rule with mode

Perakis & Roels (2008) derived the following minimax optimal order quantity assuming unimodality with a given mode m and finite range $[A, B]$:

$$q^* = \begin{cases} A + \sqrt{(m-A)(1-\beta)[B(1-\beta) - A(1+\beta) + 2\beta m]} \\ \qquad\qquad\qquad\qquad \text{if } \; m(1-2\beta(1-\beta)) \geq (1-\beta)^2 A + \beta^2 B, \\ B - \sqrt{\beta(B - m(B(2-\beta) - \beta A - 2m(1-\beta))} \\ \qquad\qquad\qquad\qquad \text{if } \; m(1-2\beta(1-\beta)) \leq (1-\beta)^2 A + \beta^2 B. \end{cases} \tag{1.31}$$

They studied robustness of beta distributions with increasing, decreasing, and unimodal PDFs on $[0, 300]$, which were scaled or shifted to have mode $m = 100$. As noted earlier in this essay, the beta distribution is characterized as the ME model subject to two geometric moments, which for the case of a finite support $[A, B]$ are $E_f[\log(Y - A)] = \theta_1$ and $E_f[\log(B - Y)] = \theta_2$. Consequently, the use of beta distribution assumes information

TABLE 1.1: Comparison between the profit maximizing quantile q of the ME distribution given range $[0, 300]$ and mode $m = 100$ and the minimax regret order quantity q^*.

	β								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Case in (1.31)	2	2	2	2	2	1	1	1	1
q^*	211.6	176.7	150.9	130.2	112.8	97.9	83.0	66.3	45.8
q	209.4	172.2	147.6	129.6	115.8	105.9	98.1	87.3	66.6
$\lvert q^* - q\rvert$	2.2	4.5	3.3	0.6	3.0	8.0	15.1	21.0	20.8

which is not included in the derivation of q^* in (1.31). Moreover, for any ME distribution noncommittal to the quantile information $E_f[\mathbf{1}(Y > q^*) = \beta$ lack of robustness is expected. The following example illustrates this for the ME models with mode and range information.

Example 1.5.

(a) The ME model consistent with the partial information assumed in the derivation of q^* in (1.31) is given by (1.20) and (1.29). This model is not robust for the minimax regret rule. Table 1.1 compares the minimax regret order quantity q^* in (1.31) with range $[0, 300]$, mode $m = 100$, with the corresponding profit maximizing quantiles q of the ME distributions for several values of β. The results indicate that the minimax regret q^* decreases faster than q such that for $\beta \leq 0.4, q < q^*$ and for $\beta > 0.4, q > q^*$. The discrepancy between minimax regret order and the profit maximizing quantile of the ME model with mode, $|q^* - q|$, is increasing in β.

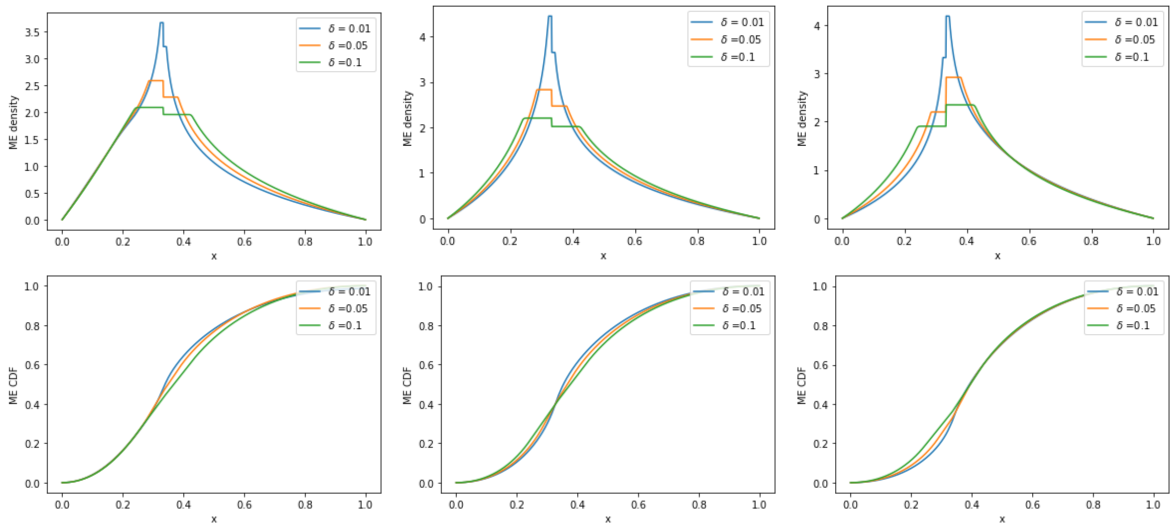

FIGURE 1.10: Robust PDFs and CDFs of demand with the goal PDF g_1, given mode $m = 1/3$ and minimax regret order quantity $q^* = 0.22$ (left), 0.33 (middle), 0.43 (right) for $\beta = 0.8, 0.6, 0.4$, respectively.

(b) For the ME model consistent with the minimax regret optimal order quantity q^*, we combine the quantile information $E_f[\mathbf{1}(X > q^*)] = \beta$ with the mode. Then use the ME model is obtained similarly to (1.26). Figure 1.10 presents plots of the smoothened versions of PDFs and CDFs of the perfectly robust ME models with $m = 1/3$ and $\beta = 0.8, 0.6, 0.4$. Equation (1.31) with $A = 0$ and $B = 1$ gives the minimax regret order quantities $q^* = 0.22, 0.33, 0.43$. For the $[0, 300]$ range, these values correspond to $m = 100$ and $q^* = 66, 98, 130$ obtained from (1.28) and the ME PDFs can be easily found by (1.29). These perfectly robust models for the minimax regret rule can be used for inference about the demand such as constructing interval forecasts.

1.3.2 Robust ME model for minimax regret rule with mode and median

Perakis & Roels (2008) derived the following minimax optimal order quantity assuming unimodality with a given mode m and median $M \geq m/2$:

$$q^* = \begin{cases} M + (1-2\beta)\sqrt{m(m-M)} & \text{if } M \leq m, 1 - \frac{m}{2M} \leq \beta \leq \frac{1}{2}, \\ 2M\sqrt{\beta(1-\beta)} & \text{if } M \leq m, \beta \geq \frac{1}{2}, \\ 2\sqrt{m(1-\beta)(2\beta m - \beta M + M - m)} & \text{if } m \leq M \leq mQ(\beta), \beta \geq \frac{1}{2}, \\ M - \sqrt{(M-m)(2\beta-1)(4\beta m - 2\beta M - 4m + 3M)} & \text{if } M \geq mQ(\beta), \beta \geq \frac{1}{2}, \\ \text{Not defined} & \text{otherwise,} \end{cases} \tag{1.32}$$

where $Q(\beta) = \frac{8\beta^2 - 12\beta + 5}{4(\beta-1)^2}$. They studied robustness of the lognormal, gamma, and negative binomial distributions for this rule. The lognormal and gamma distributions are ME models that are noncommittal to the profit maximizing quantile constraint, so the lack of robustness for the minimax regret rule is expected. The ME characterization of the negative binomial distribution is not available in the literature.

The following example illustrates ME models with mode and median with/without the quantile information in terms of q^* in (1.32).

Example 1.6.

(a) The ME model with the same partial information assumed in the derivation of q^* in (1.32) is given by (1.26) and (1.29). This model is not robust for the minimax regret rule. Table 1.2 compares the minimax regret order quantity q^* for several values of β with $m = 150$ and median as shown in the table ($M = 90, 180$) with the profit maximizing quantile q of the ME distribution given the same information obtained from (1.32) and (1.29). With $M = 90$ and $\beta < 0.5$ the conditions of case 1 of (1.32) are required which do not hold for $\beta = 0.1$. The discrepancy between minimax regret order and the profit maximizing quantile of the ME model $|q^* - q|$ decreases for $0.2 \leq \beta \leq 0.5$, then increases for $0.5 \leq \beta \leq 0.8$, and drops at $\beta = 0.9$. With $M = 180$ the conditions in case 4 of (1.32) hold. The discrepancy between minimax regret order and the profit maximizing quantile of the ME model $|q^* - q|$ increases for $0.5 \leq \beta \leq 0.8$, and drops sharply at $\beta = 0.9$.

(b) Figure 1.11 presents plots of the smoothened versions of PDF and CDF of the perfectly robust ME model (1.27) on the support $[0, 1]$ with $m = 0.5, M = 0.6$ and $\beta = 0.7$. Equation (1.32) with $A = 0$ and $B = 1$ gives the minimax regret order quantities $q^* = 0.47$, which is used for the quantile information included for calculating the robust ME model. (For the $[0, 300]$ range, the scaled values correspond to $m = 150, M = 180$ and $q^* = 141$) obtained from (1.28)). This perfectly robust model for the minimax regret rule can be used for inference about the demand such as constructing interval forecasts.

TABLE 1.2: Comparison between the profit maximizing quantile q of the ME distribution given range, mode ($m = 150$) and median ($M = 180$ and 90), and the minimax regret quantity q^*.

		β								
M		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
90	Case in (1.32)	5	1	1	1	2	2	2	2	2
	q		130.5	116.6	104.4	90.0	79.5	66.0	51.4	35.1
	q^*		134.1	119.4	104.7	90.0	88.2	82.5	72.0	54.0
	$\lvert q^* - q \rvert$		3.6	1.8	0.3	0	8.6	16.5	21.0	18.9
180	Case in (1.32)	5	5	5	5	4	4	4	4	4
	q					180	172.5	163.5	151.5	123.0
	q^*					180	157.5	144.0	131.2	118.8
	$\lvert q - q^* \rvert$					0	14.9	19.5	20.2	4.2

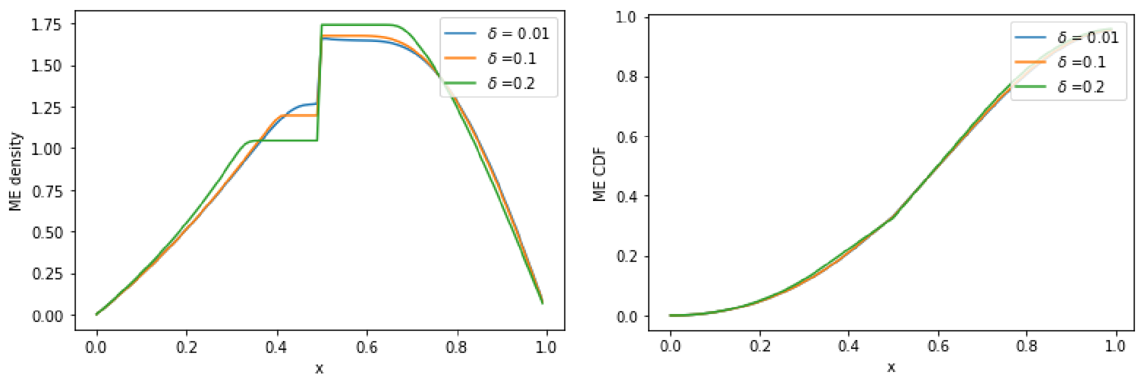

FIGURE 1.11: The MDI PDF and CDF of demand with the goal PDF g_1 given mode ($m = 0.5$), median ($M = 0.6$) and minimax regret order quantity $q^* = 0.47$ for $\beta = 0.7$.

1.4 Application to PERT

PERT was developed in the mid-1950s in the U.S. Navy. Malcolm et al. (1959) summarized the development and application of the technique used in the Navy. Figure 1.12 presents the network graphical representation (flow) of an example for a simple project

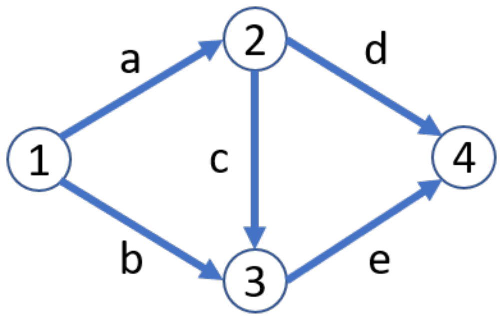

FIGURE 1.12: Project flow.

widely used in the literature since Malcolm et al. (1959). Events are depicted by circled numbers and they show a point in time that is beginning and end of a specific task or activity required for a project. The arrows show an activity necessary to achieve an event. An Activity cannot be initiated until the immediately preceding event has been accomplished. The project in Figure 1.12, includes five activities, a, b, c, d and e, and four events, 1-4. The duration of activities are assumed to be stochastic.

In the classical PERT, it is assumed that an expert provides for the optimistic completion time (A_i), pessimistic completion time (B_i), and the most likely completion time of each activity (m_i). The completion time Y_i of each activity is assumed to have an independent beta distribution with PDF

$$f_i(y) = \frac{(B_i - A_i)^{v_i+w_i-1}}{B(A_i, B_i)}(y - A_i)^{v_i-1}(B_i - y)^{w_i-1}, \quad A_i \leq y \leq B_i,\ i = 1, \ldots, n, \quad (1.33)$$

where $B(A_i, B_i) = \frac{\Gamma(A_i)\Gamma(B_i)}{\Gamma(A_i+B_i)}$ is the beta function and n is the number of arrows in the graph.

The expected completion time of each activity is given by

$$E(Y_i) = \frac{B_i v_i + A_i w_i}{v_i + w_i}.$$

The pair of shape parameters, (v_i, w_i) cannot be determined by the given mode, hence $\mu_i = E(Y_i)$ cannot be computed by the triplet of elicited information (m_i, A_i, B_i). For resolving this issue, in the classical PERT analysis it is assumed that the standard deviation of Y_i is the following simple function of (m_i, A_i, B_i) (Malcolm et al., 1959; Littlefield & Randolph, 1987):

$$\sigma_i = \frac{B_i - A_i}{6}. \tag{1.34}$$

This simplifying assumption implies that

$$\mu_i = \frac{4m_i + A_i + B_i}{6}. \tag{1.35}$$

In a project that completion of the project requires that all activities in the network to be completed, the path with maximal time of completion determines the project completion time. PERT analysis includes the expected values and variances of activities of the paths that connects the first and last nodes. Suppose that there are K such paths and let $Y_{i1}, \ldots, Y_{in_k}$ denote the times required for completion times of activities in the kth path. Then the expected completion time of the kth path is $X_k = \sum_{i=1}^{n_k} Y_{ki}$ and

$$E(X_k) = \sum_{i=1}^{n_k} E(Y_{ki}).$$

The expect completion time of the project is defined by

$$E(X_{k^*}) = \arg\max_k E(X_k). \tag{1.36}$$

In the PERT literature, the path k^* is called the *critical path*. The use of (1.34) for determining the beta parameters needed for computing (1.35) induces non-elicited information about the expected project completion time (1.36).

In the PERT literature, the variance of the kth path is also computed as

$$\mathrm{V}(X_k) = \sum_{i=1}^{n_k} \mathrm{Var}(Y_{ki}).$$

This calculation assumes that the completion times of activities are uncorrelated. Furthermore, distributions of completion of each path X_k is assumed to be normal. This assumption is based on the Central Limit Theorem which requires the number of activities in a path to be large. However, X_k can be easily simulated for any n_k.

Many authors have modified the classical PERT approach. For example Hahn (2008), Trietsch et al. (2012), and Perez et al. (2016) proposed distributions that are more flexible than the beta distribution. Van Dorp & Mazzuchi (2000) included quantile information for determining the beta parameters. These parametric distributions induce non-elicited information in the analysis. Haneveld (1986) relaxed the assumption that the distribution of duration of the activity is known and proposed a minimax approach. However, like the case of NV problem, this approach does not provide a distribution for the activity completion time.

The use of ME in PERT dates back to Kotiah & Wallace (1973). These authors used (1.34) and (1.35) for the mean and variance constraints which give the truncated normal distribution on (A_i, B_i) as the ME model. The result of this approach is contrary to their motivating statement that the ME distribution "makes use of all prior information and nothing more". Hernández-Bastida & Fernández-Sánchez (2019) attempted to use mode for deriving the ME distribution for PERT. These authors repeated the derivation of BCP for PERT, without citing the original source and did not include quantile information with the mode. Also, they did not use the smoothening method. As was illustrated in Section 1.2, the ME distribution with mode contains an infinite discontinuity at the mode, which can be corrected by the adjustment methods proposed by BCP.

This section accomplishes the following objectives.

- We compute distributions of completion times of paths based on ME distributions for durations of activities and include BCP's correction for the infinite discontinuity of the PDF around the mode.

- We compute the distributions of the minimum, maximum, and expected completion of completion times of paths. These distributions provide probabilistic extensions for the analysis of duration of paths.

- We report applications of our methodologies using real-world data for two construction projects based on information provided by a national construction firm. This application includes the median with the mode for the ME modeling of PERT.

TABLE 1.3: Activities duration data.

Activity	Minimum	Maximum	Most Likely
a	0	2	0.8
b	0	4	2.5
c	0	2	0.9
d	0	3	2
e	0	2	0.8

1.4.1 ME approach to PERT

We compare path completion times according to the notion of *stochastic order* of random variables defined as follows. Let X_1 and X_2 be two random variables with survival functions $S_k(x) = P(X_k > x)$. Then, X_1 is said to be stochastically less than X_2, denoted by $X_1 \leq_{st} X_2$, if $S_1(x) \leq S_2(x)$ for all x is the support of the distributions. In economics and decision analysis, this notion of stochastic order is called the first order stochastic dominance. The stochastic dominance provides stronger comparisons than the mean for completion times of paths. The stochastic ordering implies that $E(X_1) \leq E(X_2)$ and every αth quantile of X_1 is smaller than the αth quantile of X_2.

The following example illustrates distributions of completion times of paths computed based on smoothened ME distributions for activities.

Example 1.7. Consider the project network shown in Figure 1.12 and suppose that Table 1.3 gives the data for the completion of the activities. We use the smoothened version of the ME PDF (1.20) with the goal PDF g_1 shown in Figure 1.2 to compute the ME model for the completion time of each activity. The support of distributions in

Figure 1.2 is $[0, 1]$. The ME distribution of Y_{ki} can be obtained from the distribution of X with support on the unit interval by (1.28). Then we obtain the distribution of the completion time of each path by simulating from the ME distributions of the completion times of activities in the path.

Figure 1.13 presents plots of the completion time distributions for the three paths obtained from the ME models (blue) and beta models (orange). These PDF plots indicate that the PDFs implied by the beta distributions have smaller variances with more concentration around the mode than the PDFs implied by the ME distributions. This is expected from the ME principle which relies only on elicited information A_i, B_i, and m_i, whereas, beta distributions for the activities, in addition to the elicited information, use the non-elicited information induced via the assumed variance (1.34). The lower panels show survival plots of the distributions obtained by the two approaches. For each path, the two survival functions cross each other, which indicate the lack of stochastic order between the distributions obtained by the two methods.

Figure 1.14 presents survival plots for the three paths using the ME (left) and beta (right) approaches. The plots in the left panel indicate that the distribution obtained by the ME approach for path $a - c - e$ is stochastically dominated by the distributions for the other two paths. Consequently, the mean, median, and all other quantiles of the duration of the path $a - c - e$ are smaller than the corresponding measures for the other two paths. However, the survival functions for paths a-d and b-e cross each other near $x = 3$, below which the survival function for a-c-e dominates the survival function for b-e and for $x > 3$ the dominance is reversed. The plots in the right panel indicate that the

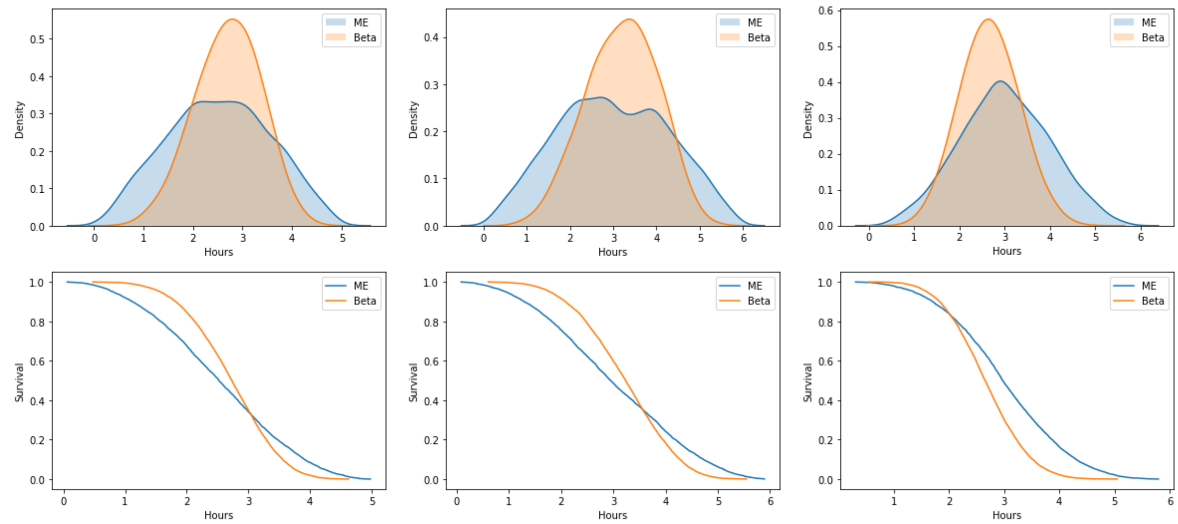

FIGURE 1.13: Distributions of completion times for the paths a-d (left), b-e (middle), and a-c-e (right) using ME distributions for activities (blue) and beta distributions for activities (orange).

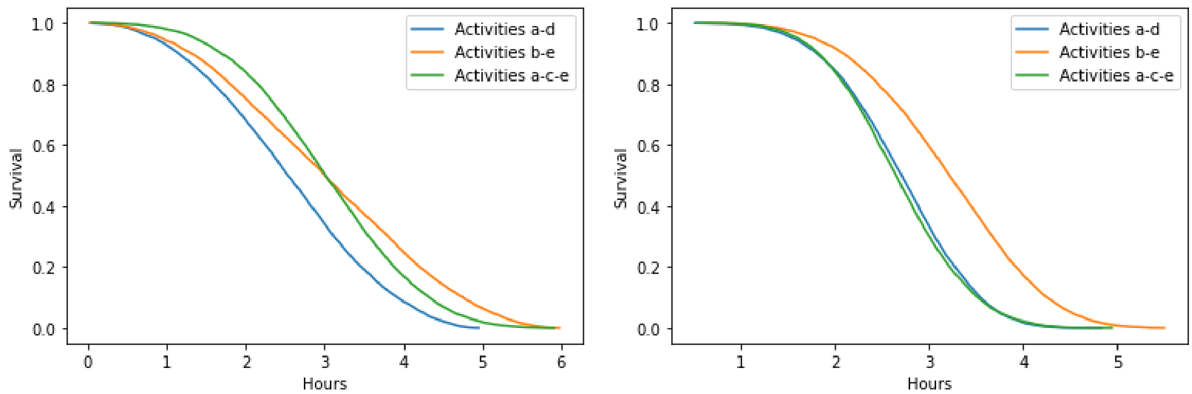

FIGURE 1.14: Survival functions of completion time of paths using ME distributions for activities (left) and PERT distributions using beta (right).

distribution obtained by the beta approach for path $b-e$ dominates the distributions for the other two paths. Consequently, the mean, median, and all other quantiles of the duration of the path $b-e$ are larger than the corresponding measures for the other two paths. The survival function of completion times of a-d and a-c-e are indistinguishable, indicating that the completion times of these paths are stochastically nearly equal.

1.4.2 Maximum, minimum, and expected distributions of paths

Consider the maximum and minimum of the set of completion times of paths:

$$Z_{\max} = \max\{X_1, \ldots, X_K\},$$
$$Z_{\min} = \min\{X_1, \ldots, X_K\}.$$

The distributions of $Z_{\max}$ and $Z_{\min}$ are determined by the joint distributions of $(X_1, \ldots, X_K)$. In the absence of information about dependence among the durations of paths, $X_1, \ldots, X_K$, the ME principle implies that they are independent. Then, the distributions of the maximal and minimal paths are given by the marginal distributions F_k of X_k as follows:

$$F_{\max}(z) = P(Z_{\max} \leq z) = P(X_1 \leq z, \ldots, X_K \leq z) = \prod_{k=1}^{K} F_k(z),$$
$$F_{\min}(z) = P(Z_{\min} \leq z) = 1 - P(Z_{\min} > z) = 1 - P(X_1 > z, \ldots, X_K > z) = 1 - \prod_{k=1}^{K}(1 - F_k(z)).$$

Random variables $Z_{\max}$ and $Z_{\min}$ provide stochastic bounds for the completion times of all paths: $Z_{\min} \leq_{st} X_k \leq_{st} Z_{\max}, k = 1, \ldots, K$. In particular, if we denote the project completion time by X^*, then

$$Z_{\min} \leq_{st} X^* \leq_{st} Z_{\max}. \tag{1.37}$$

We call $Z_{\max}$ and $Z_{\min}$ the maximal and minimal "paths". Note that these stochastic bounds may not be attainable by the completion time of any path in a project network. An important implication of (1.37) is that $E(Z_{\max}) \geq E(X_k)$ for all $k = 1, \ldots, K$, and every αth quantile of $Z_{\max}$ dominates the αth quantile of X_k for all $k = 1, \ldots, K$. Consequently, optimal decisions that minimize the risks (expected losses) according to the quadratic loss function and symmetric and asymmetric linear loss functions based on the distribution of the maximal path are higher than those based on the distributions of duration times of all other paths $X_k, k = 1, \ldots, K$. Similarly, (1.37) implies that optimal decisions that minimize the risks (expected losses) according to the quadratic loss function and symmetric and asymmetric linear loss functions based on the minimal path are smaller than those X_k for all $k = 1, \ldots, K$.

The expected distribution of the completion time of the paths is given by the following mixture of the distributions of the completion times of paths:

$$F_{\text{mix}}(x) = \sum_{k=1}^{K} \pi_k F_k(x), \quad \pi_k > 0, \sum_{k=1}^{K} \pi_k = 1,$$

where π_k is an unknown probability associated with the distribution of the completion time X_k. We use the Dirichlet prior for the unknown probabilities $(\pi_1, \ldots, \pi_K)$ with PDF

$$p(\pi_1, \ldots, \pi_K) = \frac{1}{B(\nu_1, \ldots, \nu_K)} \prod_{k=1}^{K} \pi_k^{\nu_k - 1}, \quad \pi_k > 0, \; \sum_{k=1}^{K} \pi_k = 1$$

where

$$B(\nu_1, \ldots, \nu_K) = \frac{\prod_{k=1}^{K} \Gamma(\nu_k)}{\Gamma(\sum_{k=1}^{K} \nu_i)}, \quad \nu_k > 0, k = 1 \ldots, K$$

is the multivariate beta function.

The following example illustrates the maximal, minimal,and expected distributions of completion times of paths in Figure 1.12.

Example 1.8. Figure 1.15 shows plots of the survival functions in left panel of Figure 1.14 superimposed by the survival plots of $F_{\max}, F_{\min}$, and F_{mix}. We have used $\nu_k = 1, k = 1, 2, 3$ which produces flat Dirichlet distribution $p(\pi_1, \pi_2, \pi_3) = 2$; this prior reflects lack of information that favors any of the paths. These plots illustrate that the maximal distribution stochastically dominates and the minimal distribution is stochastically dominated by the distributions of the completion times of all three paths. Consequently, the maximal and minimal distributions provide upper and lower bounds for the completion times of all three paths. The expected distribution of the completion time of the paths provides a stochastic average for the completion times of all paths.

The expected, minimal, and maximal distributions are useful for learning about the project. Table 1.4 reports selected percentiles of the maximal, minimal, and expected distributions. The median completion times of the paths is between 1.02 and 3.52 with a weighted average of 2.86. The 90 percent intervals for minimal, expected, and maximal distributions are [0.11, 2.76], [1.54, 4.16], and [1.77, 4.81].

Table 1.4: Percentiles of the minimal, maximal, and expected distributions.

Percentiles	5 %	10 %	25 %	50 %	75 %	90 %	95 %
Minimal	0.11	0.21	0.51	1.02	1.71	2.39	2.76
Expected	1.54	1.83	2.31	2.86	3.42	3.88	4.16
Maximal	1.77	2.14	2.81	3.52	4.22	4.62	4.81

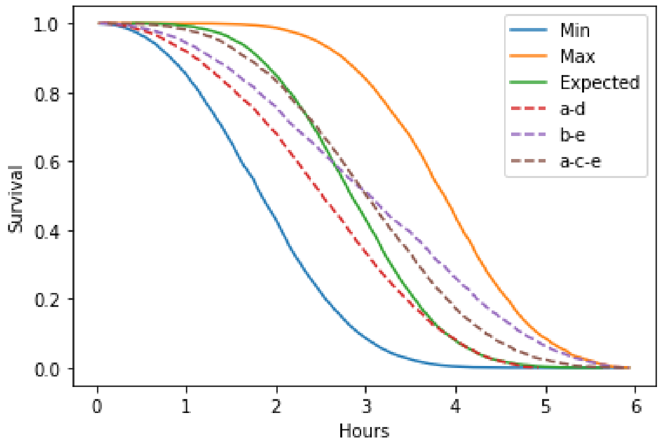

FIGURE 1.15: Survival functions of minimal, expected under Dirichlet prior, and maximal for completion time of paths.

TABLE 1.5: Construction project data.

	Duration	Common factors causing delay	Other factors causing delay
Minimum	27	No delays	
Most likely	30	Shortage of labor, Re-design	
Median	31	Shortage of labor, Re-design, Weather, Additional city required work	Shortage of material
Maximum	38	Shortage of labor, Re-design , Weather, Additional city required work, Lack of coordination between trades	Shortage of material, Pandemic, Project shutdown by community, Fire/Accidents

1.4.3 ME models for construction projects

We report applications of our methodology for two real-world projects based on information provided by a national construction firm. The company is planning to construct an apartment complex with 250 units and 500 stall garages in 2021. Information for these two projects (construction and pre-construction) is provided by a Senior Pre-construction Manager.

1.4.3.1 Completion time of the construction project

In the construction project, information on minimum, maximum, most likely, and the median of project completion time is given in month. The minimum estimation is based on the assumption of no delays. The median value is estimated by taking into account some common factors, such as shortage of labor and re-design delays, which can cause delays. The most likely value is based on the same factors and some other factors, such as bad weather and additional city required work. The maximum value is estimated by considering all the common factors and other factors, such as shortage of material, project shutdown by the community, and the pandemic. Table 1.5 presents the information about the project and factors that can cause delays.

Figure 1.16 presents PDFs and survival functions of the construction completion time obtained by the smoothened version of the ME model (1.26) with smoothening function g_1 (shown in blue) and by the basic PERT approach (shown in orange). The ME PDF is sharper around the mode. The reason for this is that the mode and median are close and beta model does not use the given median and uses other types of extraneous information. The survival plots show that the ME distribution stochastically dominates the beta distribution. Thus, the mean and every αth quantile of the ME model for the construction completion time dominate the mean and the αth quantile of the beta model. Consequently, the tradition PERT provides a more optimistic model for the construction completion time than the ME model which uses every piece of partial information available for the project. Table 1.6 compares some selected percentiles of the ME and beta distributions. We note that the median of beta distribution is smaller than the median

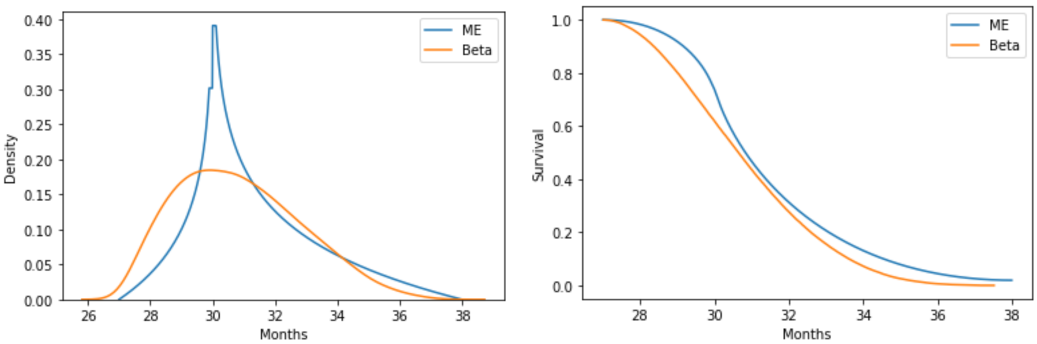

FIGURE 1.16: Construction project completion time PDF and survival function using ME approach and g_1 smoothing function (blue), and beta distribution (orange)

TABLE 1.6: Construction project completion time percentiles.

Percentiles	5 %	10 %	25 %	50 %	75 %	90 %	95 %
ME	28.79	29.36	30.11	31.00	32.86	34.72	35.83
Beta	27.97	28.37	29.29	30.63	32.23	33.62	34.41

provided by the manager. In this case, the median of beta model is close to the median given by the manager, however, such proximity is not general.

1.4.3.2 Completion time of the pre-construction project

In this problem the manager is interested to estimate the total time of three activities required before starting the construction. The activities consist of obtaining city approval for the project, completing land purchase, and obtaining construction permit. These activities must be completed sequentially, which makes this a single path problem. The manager provided the minimum, maximum, and most likely values of each activity as shown in Table 1.7.

Figure 1.17 presents PDFs and survival plots of the completion time of pre-construction project obtained by the ME approach (blue) and basic PERT approach

TABLE 1.7: Pre-construction project's activity duration data (months).

Activity	Minimum	Maximum	Most Likely
City approval	1	4	2
Land purchase	12	20	14
Permit approval	1	4	2

(orange). The ME distribution for each activity is computed using the smoothened version of the ME PDF (1.20) with the goal PDF g_1. The distribution of the pre-construction completion time is obtained via simulating from the ME distributions of the activities. The PDF plots indicate that the distribution obtained from the ME models is substantially less concentrated than the distribution obtained from the beta distributions. The concentration of the distribution obtained from the beta distributions is due to the fact that the beta model uses non-elicited information in addition to those provided by the manager. The survival plots display the lack of stochastic dominance between the distributions obtained by the two approaches.

Table 1.8 compares selected percentiles of completion time given by the two models. The model obtained by the ME approach is less concentrated and provides more prudent interval estimates for the pre-construction project completion time than the model obtained by the basic PERT approach. For example, the ME model gives the 90 percent interval as [15.75, 23.65] which is wider than interval estimate [16.49, 21.82] given by the basic PERT model.

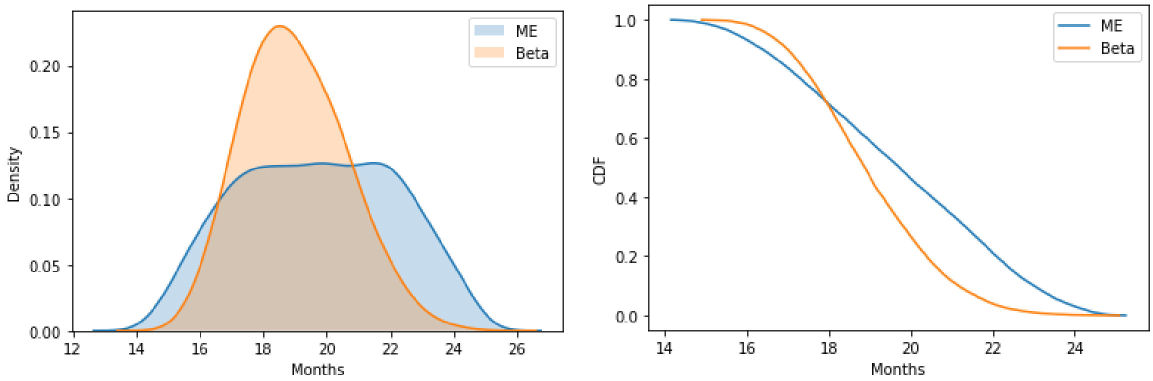

FIGURE 1.17: Pre-construction completion time PDFs and survival functions using ME and beta distributions.

TABLE 1.8: Completion time of pre-construction project's percentiles.

Percentiles	5 %	10 %	25 %	50 %	75 %	90 %	95 %
ME	15.75	16.39	17.69	19.65	21.66	22.98	23.65
Beta	16.49	16.92	17.79	18.89	20.97	21.22	21.82

1.5 Concluding remarks

The ME modeling with moments has been used in various management science problems. Recently, in Bajgiran et al. (2021), we developed ME modeling when partial information includes quantiles and moments, and we illustrated its applications in constructing Bayesian prior, measuring uncertainty and disagreement of economic forecasters, and the Newsvendor problem. However, research in methods of eliciting information about a probability distribution has shown that respondents are able to assess the mode and quantiles more accurately than the mean, and assessing the variance is a formidable task. Brockett et al. (1995) developed the ME model with mode.

This essay has continued this line of research when partial information includes the mode and quantiles for applications in management science problems. Brockett et al. (1995) posed the problem in terms of the mode and moments, and briefly pointed out the applicability of their methodology for including a quantile. Their methodology is based on transforming moments of the variable of interest to the moments of an auxiliary variable. They displayed only an example of the transformation for a single quantile and presented plots of ME PDFs with given mode, median, and interval information.

This essay posed the problem at the more general level of relationships between the mode and quantiles. For each pair of consecutive quantiles, we identified three cases and derived the transformation of the quantiles for each case. The ME PDF with given mode and quantiles is discontinuous at the mode and sharp in the neighborhood of the mode. However, the CDF is continuous. Brockett et al. (1995) proposed an information-theoretic method for smoothening the ME PDF. We presented several examples of smoothening of ME models with a given mode, with and without given quantiles. In addition, an example illustrated that the smoothening method is also effective for ME models when the partial information includes only quantiles.

Applications to two management science problems have been presented. The first application develops ME models for demand distribution in the newsvendor problem. The newsvendor's profit function is asymmetrically linear and the optimal solution is a quantile of the demand distribution. Perakis & Roels (2008) derived the minimax optimal order quantity based on various types of partial information about the demand, including cases of given mode with and without the median. They examined the robustness of some well-known parametric models by comparing the profit maximizing quantiles of the chosen distributions with the minimax optimal order quantity. Bajgiran et al. (2021) illustrated that adding the profit maximizing quantile to the set of partial moment information (mean with/without the variance) assumed for deriving the minimax optimal orders produces perfectly robust models for the demand distribution. This essay illustrated developing perfectly robust ME models for the minimax rule when partial information includes the mode with/without the median.

The second application develops ME models for completion times of activities in the stochastic project planning problem, known as PERT. This problem usually involves partial information in terms of the minimum, maximum, and mode of the completion time of each activity. Inclusion of quantiles has also been considered (Van Dorp & Mazzuchi, 2000). Hernández-Bastida & Fernández-Sánchez (2019) repeated the derivation of Brockett et al. (1995) for PERT, without citing the original source. These authors did not use the smoothening method and did not include a quantile. This essay illustrated a more extensive information-theoretic modeling for PERT, which includes smoothening of ME models with the given mode and a quantile.

This essay also has offered a new methodology for the PERT problem that includes a network of activities between the start and completion of the project. The convolution of probability distributions of completion times of all activities in a path gives distribution of the completion time of the path. Operations researchers have studied the expected completion times for the minimal or/and maximal paths in terms of the expected value. We extended this approach in terms of stochastic order between random variables and provided stochastic upper and lower bounds for the completion times of paths and the project. We also introduced the notion of expected distribution of completion time for a network of completion times in terms of Dirichlet mixture. This formulation assumes that distributions of completion times in the space of all paths are distributed according to a discrete probability distribution with unknown probabilities. Then the Dirichlet prior distribution for the unknown probabilities gives the expected distribution. Using the ME distribution of the completion times of activities and a non-informative Dirichlet prior, we computed the expected distribution. The methodology of computing the stochastic bounds and expected distribution of completion time is applicable to other distributions used in PERT. Derivations of the stochastic bounds are based on the assumption that completion times of paths are independent. Investigating this assumption and extending it to the dependence case provide challenging research problems for future.

This essay reported applications of our methodologies to new real-world data. We elicited assessments of completion times for two projects from a Senior Preconstruction Manager of a national construction firm. The assessment for one project included the minimum, maximum, mode, and median of the completion time of a construction project. For this project, we developed the smoothened ME distribution for the completion time. The assessment for the second project included the minimums, maximums, and modes for the duration of each task in the sequence of obtaining the city permit for the project, purchasing the lot, and obtaining the construction permit. For this project, we developed the smoothened ME distribution for each task and the total time until receiving the construction permit.

Chapter 2

Bayesian prescriptive framework for complementary products: tariff strategies and channels inefficiencies

Keywords: Prescriptive analysis; Bayesian Modeling; Firm's inefficiency; Omnichannel retailing; Tariff; Pricing

2.1 Introduction

During the last five decades, five of the seven United States presidents have imposed tariffs or surcharges on imports. The latest in this series is $283 billion in 2018, with rates ranging between 10 to 50 percent (Amiti et al., 2019). These tariffs caused firms to face increasing costs across their products. For protecting their profit margins, retailers that rely on imports needed to adjust their pricing strategies to pass tariff costs to customers through pricing elements. Determining the effect of the reduction of the tariff burden through pricing elements on the demand became a challenging decision issue for retailers. The prompt decision required for addressing the problem of a sudden sharp spike of the tariff cost often did not allow the usual experimentation to learn customers' responses to major pricing adjustments. Decisions become more complex when a retailer receives orders that include complementary products, where pricing adjustments for one of the products affect the other products. More complexity arises when omnichannel retailers do not record incidences of lost sales due to the inefficiency to meet the demand. This essay studies these issues through developing a new econometric model for sales which allows probable inefficiency of meeting demands for complementary products.

Omnichannel retailers offer various channels and touchpoints, including websites, online advertising, and customized catalogs. The channels in the customers' journey can be categorized into three groups: media channels where customers get exposed to the products, purchase channels where customers make a purchase, and delivery channels from which the retailer distributes the products. From a customer's perspective, it is more convenient to have access to multiple touchpoints and purchase options that provide information about products, prices, and promotions. These conveniences for customers have prompted many retailers to transform to omnichannel retailing. Successful omnichannel retailers look for ways to identify sources of inefficiencies and improve all the touch-points in the customer's journey. For omnichannel retailers, it is challenging to coordinate tariff decisions with price, discounts, and freight of complementary products while monitoring the customer purchase experience in an integrated framework.

This research is motivated by contacts with a national omnichannel retailer that faced on average a 35 percent increase in tariffs in late 2018. This retailer offers varieties of complementary products through multiple media channels, purchase channels, and delivery channels. Our retailer, like many other retailers that faced tariffs, looked for solutions to protect the profit margin. The retailer first compensated for the tariff costs by increasing the freight charge to customers, which led to push back from customers. The reduction of demand was perceived in terms of the ratio of freight charge to the total price of the order being too high. Then the pricing strategy was adjusted by keeping the freight charge as before and increasing the price of the product. This strategy also led to the reduction of demand. To mitigate the demand reduction, retailer offered more discounts.

In this essay, we develop a new model for omnichannel retailing and prescriptive analysis for tariff strategies of complementary products. Our retailer has provided data that enables us to study issues empirically in the presence of probable latent inefficiency (lost sales) due to characteristics of the purchase and delivery channels.

The lost sales historically has been viewed in terms of shortage of products in the firms. But in the new retailing trend, in which the focus is on the customer as well as the product, shortage is not the only reason anymore for lost sales. We expand this view and conceptualize the notion of lost sales in terms of inefficiency of the retailer to meet the maximum possible sales (demand) for a product. A variety of factors in the customer journey can deviate sales from maximum potential sales. Sales of different types of products can be less than the demand for a product under a pricing structure due to factors such as salesperson's performance, website design, and delivery channels (shortage which may not be recorded by the firm such as our focal retailer). Such observable factors can lead to inefficiency in terms of unobservable random lost sales.

We propose a new econometrics model where the notion of inefficiency is defined as the deviation of the firm's actual sales from its potential maximum sales, given the firm's input variables for the demand. The proposed model combines variants of existing stochastic frontier models (SFMs) in the production economics literature. The SFM originally was proposed for studying the technical inefficiency of firm's production by the deviation of the firm's output from the maximum output that should have been produced from the production inputs, called production frontier (Aigner et al., 1977; Meeusen & Van Den Broeck, 1977). Applications of the SFM later have been expanded beyond production to the more general problem of performance. The econometric specification for the original SFM, in addition to the random noise, includes a strictly positive random term for the inefficiency. This specification is not applicable to problems where some firms could probably be perfectly efficient, such as when a retailer sometimes is able to meet the demand. This essay combines two existing extensions of the original SFM to provide a new model for sales of complementary products, which allows probable perfect efficiency of the firm to meet the demand. This model enables us to study the following problems.

- Relationships between the firm's demands for complementary products with the imposed tariff and media channels (online advertising, website visits, and catalogs), given the pricing strategies in terms of the price, freight charge, and discounts while controlling for proprietary items.

- Relationships between the probabilities that the firm can meet demands for complementary products (perfectly efficient or not) and the delivery channels.
- Relationships between characteristics of purchase channels (including the channel type and salesperson's performance) and the probability distributions of inefficiencies to meet demands for complementary products.
- Develop a probability distribution for the firm's profit which allows prescriptive analysis and comparison of various pricing strategies under the tariff regime.

We utilize a Bayesian methodology which is known to be more flexible than the frequentist approach. This approach includes uncertainty about the via probability distributions for the unknown parameters (coefficients of various predictor variables, variances of the random inefficiency terms, and correlation between the random terms of SFMs for the complementary products). The data updates prior distributions to posterior distributions which provide inferences about the model parameters. We use Bayes Factor for the evidence (Kass & Raftery, 1995) that data provides about the directions of relationships between the demand for each product and its predictors, between the inefficiency and characteristics of the purchase channel, and between the delivery channel and the probabilities that the firm can meet demand. The Bayesian methodology enables prescriptive analysis in terms of the predictive probability distribution of the firm's profit. The firm's profit function from filling an order includes sales of each type of products and components of the price (price, freight charge, and discounts). Our model gives sales as random outputs whose distributions are induced by the posterior distributions of the model parameters. The uncertainty about the sales in turn induces in the profit function. The Bayesian predictive profit distributions is defined by the expected profit distribution under all outcomes of the parameters' posterior distributions. The predictive distribution is a versatile decision-making tool that enables moving beyond relying on the expected profits for prescriptive analysis. Having profit distributions, we compare predictive profit distributions for multiple strategies for passing tariff costs to customers based on quantiles of profits, such as the median, and according to stochastic order. The predictive distributions enable computing the expected profits when so desired.

2.1.1 Literature review - Operations

The management science literature includes very few research papers about tariff impacts on firms. Dong & Kouvelis (2020) discussed tariff impacts in a global supply chain network at a general level and showed the application of the responsive Newsvendor network model for the firms affected by tariff fluctuations to adjust their ordering and pricing. In practice, retailers' decisions involve more factors than only price. For example, our focal retailer in Midwest first increased the freight charge, then examined adjustments of price, and eventually offered more discounts to mitigate the demand reduction. We use real data from this retailer to study the firm's decisions to remedy tariff costs for complementary products.

Omnichannel retailing is an expanding practice. Many companies are transforming from multichannel retailing to omnichannel retailing. This transformation provides new opportunities and new operations problems. In a survey of developments in multichannel and omnichannel retailing, Verhoef et al. (2015) categorized research in multichannel and omnichannel retailing in terms of the following problems: (1) impacts of channels on the retailer's performance; (2) shopper's behavior across channels; and (3) retail mix across channels. Our work contributes to the first problem, by developing a new model for the retailer's performance which includes probable inefficiency in meeting the demand as a function of characteristics of purchase and delivery channels.

Several papers in the literature have studied roles of channels in omnichannel retailing. Gao & Su (2017a) proposed a model to study "buy online, pick up in store," and Gao & Su (2017b) studied the use of online and offline channels for information provision. Gallino et al. (2019) developed a model to design omnichannel customer offerings. We categorize channels into three groups and study relationships of their characteristics with demand and inefficiency of the retailer to meet the demand. More recently, Caro et al. (2020) suggested some future directions for retail operations. They noted that "in the omnichannel perspective, demand must be forecasted taking the new complexities that arise into consideration". Our model can be used to forecast sales and unobservable demand, where complexities are present (unobservable inefficiency to meet the demand, multiple channels, and multiple complementary products). They also suggested that demand forecasting should avoid aggregate demand or panel data to predict future aggregate sales for inventory. Our model enables estimating sales, demand, and inefficiency at a disaggregate level. Furthermore, our empirical analysis for predicting sales of complementary products is at the order-level.

Pricing papers for omnichannel retailers are limited. Harsha et al. (2019a) studied price optimization of omnichannel retailers in the presence of cross-channel interactions in supply and demand, where cross-channel fulfillment is exogenous. Harsha et al. (2019b) studied price optimization for omnichannel demand with channel substitution. There are key differences between our model and theirs. First, our model considers the demand as a function of the product and media channel characteristics. Second, our model includes inefficiency in terms of characteristics of purchase and delivery channels. Third, our model includes dependence between sales of complementary products. Fourth, our statistical inference approach provides a new revenue management framework that enables learning from profit predictive distribution. This framework gives the customary point estimation of the profit in terms of the expected profit. The expected value minimizes the risk under the squared error loss. The profit predictive distribution enables minimum risk decisions under various types of loss functions.

This essay also contributes to the literature on the coordination of complementary products. Wang (2006) studied joint pricing-production decisions of complementary products with uncertain demand. Li (2019) studied jointly conducting intertemporal price discrimination for hardware and software. Banciu and Ødegaard (2016) used copula to study the problem of pricing a bundle of products when the underlying valuations of the bundle components are dependent. We model sales of complementary products with the demand being a function of the product characteristics which includes the price and characteristics the firm's media channels.

2.1.2 Literature review - Methodology

This essay contributes to the literature of demand estimation. Many papers in the operations literature study demand estimation with censored data in various problems, including inventory management (Huh et al., 2011; Zhang et al., 2018; Shi et al., 2016), pricing strategies (Rhuggenaath et al., 2019), and joint pricing and fulfillment decisions (Chen 2017). Ferreira et al. (2016) used sales data on items that did not sell out in Ru La La to estimate lost sales of items that did sell out. Subramanian & Harsha (2017) estimated the unobserved market-shares of competitors when the unobserved lost sales are attributable to no-purchase, or a purchase from one of the competitors. Similarly, Cho et al. (2020) estimated the size of the no-purchase customer population when transaction data for a single firm's set of products is available. These studies followed censored data analysis methodologies that censored cases are known and recorded by a binary indicator variable. These methodologies do not apply to the problem when inefficiency to meet the demand is not so observable. In this essay, lost sale is defined as the difference between sales which is an observable quantity and the demand defined as the maximum possible sales for a retailer's product which not observable and cannot be recorded. Hence, the instances of lost sales cannot be recorded.

Methodologically, this essay relies on the SFM literature. The origin of SFM dates back to 1977 for the purpose of studying technical efficiency in manufacturing at the industry level (Aigner et al., 1977, and Meeusen & Van Den Broeck, 1977). The model decomposes a firm's production output as follows: production frontier (maximum potential output given the firm's production inputs), a firm-specific stochastic inefficiency term with a continuous distribution on the strictly positive support, and a random noise term. The model has been applied to various economics and management problems, such as health care utilization (Griffin & Steel, 2004), the performance of bank assets (Greene, 2005) and bank management skills (Bos et al., 2009), the performance of individual (Griffiths et al., 2014), strategic management (Lieberman & Dhawan, 2005), productivity shocks (Tsionas & Mallick, 2019), and excess inventory (Chuang et al. 2019).

The SFM literature has developed various assumptions and models for the inefficiency term. Koop et al. (1997) modeled the inefficiency term as a random effect. Osiewalski & Steel (1998) modeled the parameters of the distribution of inefficiency to be dependent on firm-specific characteristics. Amsler et al. (2017) specified inefficiency to be a function of firm-specific characteristics. We model the inefficiency term as a function of purchase and delivery channels characteristics.

In the original SFM, the strictly positive support of the inefficiency implies that all firms are always inefficient in meeting the frontier. Kumbhakar et al. (2013) extended the original SFM by introducing the zero inefficiency SFM, referred here as the ZI-SFM, where the distribution of the random inefficiency term includes a non-zero probability at zero and a continuous distribution on strictly positive support; this type of distribution is called hurdle model (Cragg, 1971). This model relaxes the restrictive assumption that all firms are inefficient and allows some firms to be perfectly efficient with a given probability. However, unlike the censored data, in this model it is not known which firms are perfectly efficient and which ones are inefficient.

Application to complementary products requires a system of simultaneous SFM that allows for dependence between the models for individual products. Several papers have proposed extensions of the original SFM for multiple products. For example, Ferreira & Steel (2007) used multivariate skewed elliptical distribution for modeling multi-output stochastic frontiers. Carta & Steel (2012) used copula for modeling the dependence between the firm's inefficiencies across the products. These authors also developed Bayesian inferences for their proposed models. We use copula for modeling the dependence between compound deviations (inefficiency plus noise) of the frontiers and the firm's performances of the individual product models. Copula models are used in the operations literature to model demand dependence by Corbett et al. (2006) to study the inventory pooling effect for non-normal dependent demand models, and by Yang et al. (2020) to study ordering decisions for multilocation newsvendor problems.

We contribute to prescriptive analysis literature. Prescriptive analysis is defined as the prescription of a decision in anticipation of the future. Bertsimas & Kallus (2019) studied the conditional-stochastic optimization problem of minimizing expected costs, given the observed data. They proposed various methods for constructing weights in the cost function, based on kernel regression, k-nearest neighbors, local linear regression, classification and regression trees, and random forests. Bertsimas & McCord (2019) extended the prescriptive learning framework to solve finite-horizon multiperiod optimization problems, including the inventory management problem. In the context of inventory management, Ferreira et al. (2018) used Thompson sampling for learning the elasticities and optimization of prices at the same time. Other research on prescriptive analysis includes Bertsimas et al. (2019), and Harsha et al. (2019c). Research on prescriptive analysis in omnichannel retailing is very limited. Our prescriptive framework is new, in that it includes firm's inefficiency and complementary products for predictions and decisions based on Bayesian predictive distributions for profit.

This essay is organized as follows. Section 2 gives the econometric specifications of the original SFM model and its existing extensions, which provide backgrounds for developing our new SFM. Section 3 presents specification of our model for the application to omnichannel retailing problems. Section 4 presents the empirical results of the proposed model using data supplied by our focal retailer.

2.2 Original SFM and its existing extensions

Classical production theory assumes that firms are fully efficient. In reality, some firms may not be efficient in utilizing their factors of production which leads to the discrepancy (inefficiency) between their observed output and the maximum possible output. The SFM methodology is developed for estimating the underlying maximum output (production frontier) while accounting for firm's random inefficiency.

2.2.1 Variants of original SFM

The original SFM model relates a performance variable y_i of the firm i to a vector of input variables $\boldsymbol{x}_i$ in the following form:

$$y_i = \mathcal{F}(\boldsymbol{x}_i'\boldsymbol{\beta}) - u_i + \nu_i, \quad i = 1, \dots, n, \tag{2.1}$$

where $\boldsymbol{\beta}$ is a vector of coefficients for the inputs, $\mathcal{F}(\boldsymbol{x}_i'\boldsymbol{\beta})$ represents the maximum performance that can be achieved by the given inputs, called the frontier, u_i is a stochastic component that represents the firm's performance inefficiency, and ν_i is the random noise. All firms are assumed to be inefficient, so the distribution of u_i is assumed to have a continuous probability density function (PDF) f_u on the strictly positive support. The noise terms are assumed to be independent and identically distributed (IID) and $f_\nu = N(0, \sigma_\nu^2)$. The distribution of inefficiency is assumed to be half-normal on the positive support, $f_u = N_+(0, \sigma_u^2)$. Similar to the random components models, it is assumed that the inefficiency and noise are independent. The PDF of the compound deviation of the firm's performance from the frontier, $\epsilon_i = \nu_i - u_i$, is

$$f_\epsilon(\epsilon) = \frac{2}{\sigma_\epsilon}\phi\left(\frac{\epsilon}{\sigma_\epsilon}\right)\Phi\left(\frac{-\lambda\epsilon}{\sigma_\epsilon}\right), \tag{2.2}$$

where ϕ is the PDF and Φ is the cumulative distribution function (CDF) of the standard normal distribution, $\lambda = \sigma_u/\sigma_\nu$, and $\sigma_\epsilon = \sqrt{\sigma_\nu^2 + \sigma_u^2}$; see, for example, Greene (2005).

Some variants of the original SFM include firm specific characteristics in the model. Most of empirical research in the SFM literature utilizes panel data. The following panel data SFM model was introduced by Greene (2005):

$$y_{it} = \mathcal{F}(\mu + \boldsymbol{x}'_{it}\boldsymbol{\beta}, \boldsymbol{w}'_i\boldsymbol{\zeta}) - u_{it} + \nu_{it}, \quad i = 1, \ldots, n, \ t = 1, \ldots, T, \tag{2.3}$$

where $\boldsymbol{x}_{it}$ is the input vector (such as prices), $\boldsymbol{w}_i$ is a vector of firm specific characteristics that represent heterogeneity unrelated to the inefficiency.

The SFM literature also offers various models for relating the random inefficiency term to a vector of covariates that represents environmental variables such as a firm's specific characteristics (Amsler et al., 2017; and Kutlu et al., 2019; among others). For example, Amsler et al. (2017) offered the following model for the inefficiency term:

$$u_i = u_i^0 \exp\{\boldsymbol{z}'_i\boldsymbol{\eta}\}, \tag{2.4}$$

where u_i^0 is the "basic inefficiency term" with distribution $f_{u_i^0} = N_+(0, \sigma_u^2)$, $\boldsymbol{z}$ is a vector of firm specific variables unrelated to the frontier, and $\exp\{\boldsymbol{z}'_i\boldsymbol{\eta}\}$ is called the "scaling function".

2.2.2 Zero-inefficiency SFM

The zero-inefficiency SFM introduced by Kumbhakar et al. (2013) accommodates both efficient and inefficient firms. Let $\alpha \in [0, 1]$ denote the probability that the firm is perfectly efficient. Then the PDF of the inefficiency term in (2.1) is the following mixture:

$$f_u(u) = \alpha \mathbf{1}(u = 0) + (1 - \alpha) f_c(u) \mathbf{1}(u > 0), \tag{2.5}$$

where $\mathbf{1}(A)$ is the indicator function of set A and f_c is a continuous PDF on strictly positive support. Specification of f_c completes the specification of the extension of (2.1) to accommodate perfectly efficient firms. With the normal distribution for f_ν and half-normal distribution for f_c, the PDF of the compound deviation of the firm's output from the frontier is the following mixture,

$$f_\epsilon(\epsilon) = \frac{\alpha}{\sigma_\nu} \phi\left(\frac{\epsilon}{\sigma_\nu}\right) + \frac{2(1-\alpha)}{\sigma_\epsilon} \phi\left(\frac{\epsilon}{\sigma_\epsilon}\right) \Phi\left(\frac{-\lambda\epsilon}{\sigma_\epsilon}\right). \tag{2.6}$$

A natural extension of zero-inefficiency SFM is the specification of the probability of full efficiency α as function of a set of covariates. Kumbhakar et al. (2013) suggested the Probit or logit models:

$$\alpha = \Phi(\boldsymbol{v}'\boldsymbol{\delta}), \quad \text{or} \quad \alpha = \frac{\exp(\boldsymbol{v}'\boldsymbol{\delta})}{1 + \exp(\boldsymbol{v}'\boldsymbol{\delta})}, \tag{2.7}$$

where $\boldsymbol{v}$ is a vector of covariates and $\boldsymbol{\delta}$ is a vector of coefficients.

2.2.3 Multi-output SFM

Modeling multi-output SFM requires a system of simultaneous SFMs that allows for dependence between the SFMs for individual outputs. Carta & Steel (2012) used a copula to model dependence between the firm's inefficiencies across the outputs. Consider $Y_1, \ldots, Y_K$ to be K random outputs with marginal CDFs $F_1, \ldots, F_K$. The copula is a multivariate function, $C : [0,1]^K \to [0,1]$ that transforms the marginal CDFs to a joint CDF for $(Y_1, \ldots, Y_K)$ as follows:

$$F(y_1, \ldots, y_K) = C(F_1(y_1), \ldots, F_K(y_K)).$$

The function C, called copula function, is a multivariate distribution function with uniform marginals on $[0,1]$. The joint PDF of $(Y_1, \ldots, Y_K)$ is given by the marginal PDFs f_i and the nth-variate copula PDF as follows:

$$f(y_1, \ldots, y_K) = c(F_1(y_1), ..., F_K(y_K))\ f_1(y_1)\ ...\ f_K(y_K).$$

Numerous examples of copula are available and are used in many fields for various application; see https://en.wikipedia.org/wiki/Copula and references therein. Any multivariate continuous CDF can be used as a copula. For example, the Gaussian copula is on $(u_1, \ldots, u_K) \in [0,1]^K$

$$C_{\text{Gauss}}(\boldsymbol{u}) = \Phi\Big(\Phi_1^{-1}(u_1), \ldots, \Phi_K^{-1}(u_K)\Big), \quad \boldsymbol{u} = (u_1, \ldots, u_K) \in [0,1]^K, \tag{2.8}$$

where Φ is the multivariate normal CDF with correlation Γ and Φ_k is the univariate standard normal CDF. This copula with $u_k = F_k(y_k)$ is used to model dependence by a given correlation structure among $(Y_1, \ldots, Y_K)$. The Gaussian copula function does not have simple analytical formula. However, approximations are available for integrating the Gaussian copula PDF.

$$c_{\text{Gauss}}(\boldsymbol{u}|\Gamma) = |\Gamma|^{-1/2} \exp\left\{\frac{1}{2}\Big(\Phi_1^{-1}(u_1), \ldots, \Phi_K^{-1}(u_K)\Big)'(\Gamma^{-1} - I_K)\Big(\Phi_1^{-1}(u_1), \ldots, \Phi_K^{-1}(u_K)\Big)\right\}, \tag{2.9}$$

where I_K is the K-dimensional identity matrix. Even when the copula function and its PDF have simple analytical formulas, computational challenge arises when a marginal distribution F_k does not have simple analytical formula. For example, the analytical expressions for CDFs corresponding to the PDFs of compound error (2.2) and (2.6) are not available. However, various approximations and simulation methods have been developed for CDFs of (2.2) and (2.6) which can be used for implementing copulas for SFMs; see, Amsler et al. (2019) and Tsay et al. (2013).

2.3 New SFM and prescriptive analysis

This section introduces a new SFM model that integrates the variants (2.3) and (2.4) and the two extensions of the original SFM presented in the preceding section into a single model, called Multi-product-Zero-inefficiency SFM (MZI-SFM). Figure 2.1 shows the visualization of the exiting SFM models that are integrated to develop our proposed SFM for performance of omnichannel retailers that can meet the demands for complementary products in some, but not all, occasions. The Bayesian predictive distribution of profit for prescriptive analysis will be obtained from Bayesian estimation of the MZI-SFM.

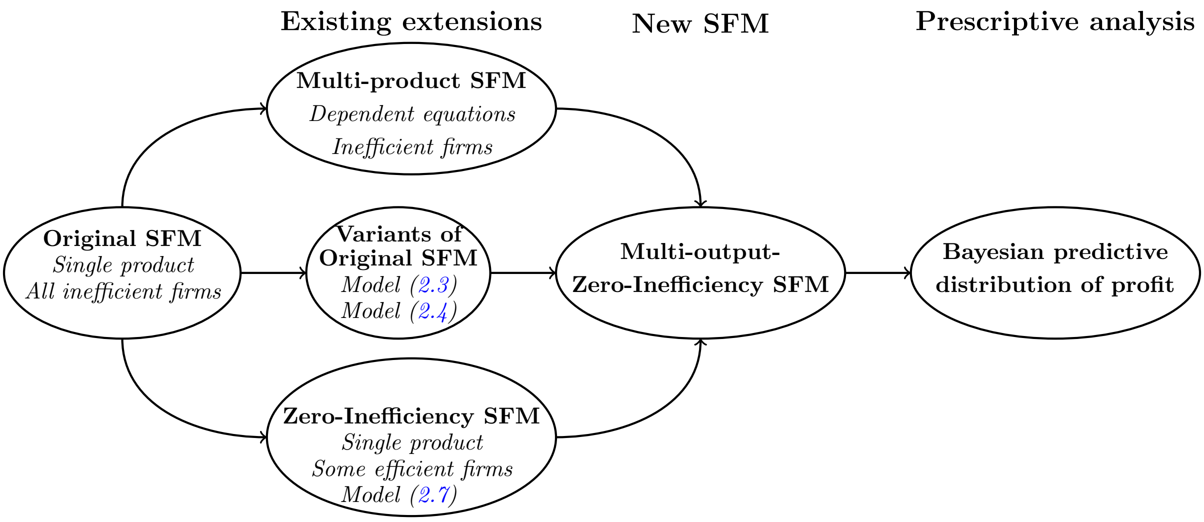

Figure 2.1: Development of the Multi-output-Zero-Inefficiency SFM and Bayesian prescriptive analysis.

2.3.1 Multi-product, zero-inefficiency SFM

2.3.1.1 Modeling omnichannel retailing

In the omnichannel retailing, costumers are exposed to advertising messages from online media and traditional media, including email and catalog. Danaher et al. (2020) measures advertising effectiveness for retailing multiple brands in multimedia and multichannel environment. The media channels have relationships with demand. Heterogeneity of the media channels' characteristics is different than product characteristics. This suggests that media channels' characteristics should be included in the frontier (demand) along with the pricing components. We adapt Greene's (2005) variant of original SFM, (2.3) to accommodate heterogeneity of the media channels' characteristics.

A retailer is efficient if the firm's purchase channels always can generate the maximum possible sales and its delivery channels always can fulfill purchase orders. However, a retailer may not be able to do so because of inefficiencies of the purchase channels and/or distribution channels due to shortage or other factors. As such inefficiency to meet the demand is not observable to be recorded. For example, our focal firm indicated that there could be rare cases of shortage, which are not recorded. Let κ denote the latent indicator of instances that a retailer does not meet demand:

$$\kappa = \begin{cases} 1 & \text{if demand} \leq \text{sales} \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

We use the zero-inefficiency SFM (2.5) with $\alpha = P(\kappa = 0)$ for the probability of the case when the firm is efficient (the firm's purchase channels generate the maximum possible sales and its delivery channels fulfill purchase orders). This gives $P(\kappa = 1) = 1 - \alpha$ for accommodating instances that the firm's purchase channels do not generate the maximum possible sales. Recall that if κ were observable, (2.10) would have simply served as a censoring indicator with α being the proportion of non-censored observations. We adapt Amsler et al.'s (2017) variant of original SFM, (2.4) in terms of purchase channel characteristics and Kumbhakar et al.'s (2013) suggested logit model (2.7) in terms of delivery channel characteristics.

For purchasing complementary products, customers' decisions about one product are influenced by the other product. Examples include hardware and software, desk and chair, and tennis racket and tennis balls. A few scenarios to demonstrate the importance of taking into account the dependency between SFM for each product are as follows:

- Promotion on product A can trigger the demands for both complementary products A and B.
- Firms can benefit from coordinating the two price instruments for A and B, because the demands for the complementary products are closely linked.
- Product A has more varieties than product B, which can effect the salesperson's performance on the purchase inefficiency of this product. Products A and B are purchased together, which makes the salesperson's performance on the purchase inefficiency of complementary products dependent.
- Shortage in product A can generate lost sales for both products A and B.

2.3.1.2 MZI-SFM specification

Without loss of generality, we present econometric specifications of our MZI-SFM for the performance of a potentially efficient retailer with two complementary products. The performance is defined by the sales of product type $k, k = 1, 2$ generated from order $i, i = 1, \ldots, n$ and will be denoted by s_{ik}.

$$y_{ik} = \gamma_k + \boldsymbol{x}'_{ik}\boldsymbol{\beta}_k + \boldsymbol{w}'_{ik}\boldsymbol{\zeta}_k - u_{ik} + \nu_{ik}, \quad \nu_{ik} \sim N(0, \sigma^2_{\nu_k}), \tag{2.11}$$

$$u_{ik} = u^o_k \exp\left(\boldsymbol{z}'_{ik}\boldsymbol{\eta}_k\right), \tag{2.12}$$

$$\alpha_{ik} = \frac{\exp(\boldsymbol{v}'_{ik}\boldsymbol{\delta}_k)}{1 + \exp(\boldsymbol{v}'_{ik}\boldsymbol{\delta}_k)}, \tag{2.13}$$

where $y_{ik} = \log s_{ik}$ and $u_k^o \sim N_+(0, \sigma^2_{u_k^0})$. The logarithmic transformation comes from the original SFM for transforming the Cobb-Douglas production frontier to additive function of factors of production inputs.

In (2.11), the frontier $\gamma_k + \boldsymbol{x}'_{ik}\boldsymbol{\beta}_k + \boldsymbol{w}'_{ik}\boldsymbol{\zeta}_k$ is for the maximum potential sales of product k that could have been generated for purchase order i. This model for the frontier is adapted from (2.3) to include two types of variables represented as $\boldsymbol{x}_{ik}$ for product characteristics and $\boldsymbol{w}_{ik}$ for media exposure variables. The first component, $\boldsymbol{x}'_{ik}$, includes products' characteristics such as price of product k, discount of product k, freight charge of product k, and tariff indicator of product k. The second component, $\boldsymbol{w}'_{ik}$, includes media channels' characteristics such as advertising platforms, catalogs, and website exposure variables, which are common for both products and change periodically.

The inefficiency term u_{ik} in (2.11) captures lack of meeting the maximum potential sales of product k that could have been generated for purchase order i. The model for the inefficiency term (2.12) is adapted from (2.4) for the inefficiency variables, $\boldsymbol{z}_{ik}$, which includes variables related to purchase channels, among others.

Distribution of inefficiency is the following adaptation of (2.5):

$$f_u(u_{ik}) = \alpha_{ik}\mathbf{1}(u_{ik} = 0) + (1 - \alpha_{ik})[1 - \mathbf{1}(u_{ik} = 0)]f_c(u_{ik}),$$
$$f_c(u_{ik}) = N_+\left(0, \exp(2\boldsymbol{z}'_{ik}\boldsymbol{\eta}_k)\sigma^2_{u_k^o}\right),$$

where the half-normal distribution parameter is implied by (2.12). The logit model (2.13) for α_{ik} is from (2.7) for the full efficiency variables, $\boldsymbol{v}_{ik}$, which represents delivery channels.

Our model for the dependence between the SFMs for complementary products are as follows. Stochastic terms ν_{ik} and u_{ik} are independent in the equation for kth product, $k = 1, 2$, and are correlated across the equations for the two products. This implies that the compound deviations from the frontiers, $\epsilon_{ik} = \nu_{ik} - u_{ik}, k = 1, 2$, are correlated across the equations for the two products. We use the Gaussian copula (2.8) with $K = 2$ for modeling dependence of the compound errors in the equations for the two products. Although the PDFs $f_{\epsilon_{ik}}, k = 1, 2$ are not continuous, the CDFs $F_{\epsilon_{ik}}$ are continuous and imply continuous marginal CDFs for $Y_{ik}, k = 1, 2$. This allows unique copula transformation $t_{ik} = F_k(y_{ik})$: $t_{ik} \sim U[0, 1]$.

We denote the set of parameters of the MZI-SFM by $\boldsymbol{\theta}_k = \{\gamma_k, \boldsymbol{\beta}_k, \boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\delta}_k, \sigma^2_{\nu_k}, \sigma^2_{u^o_k}\}$, and the correlation between ϵ_{i1} and ϵ_{i2} by ρ_{12}. Conditional on $\boldsymbol{\theta}_k$ and ρ_{12}, the Gaussian copula PDF (2.9) gives the joint PDF of (Y_1, Y_2) as:

$$f(y_1, y_2|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}) = c_{\text{Gauss}}(F_1(y_1), F_2(y_2)|\rho_{12}) \times \prod_{k=1}^{2} f_k(y_k|\boldsymbol{\theta}_k), \tag{2.14}$$

where $f_k(y_k|\boldsymbol{\theta}_k) = f_{\epsilon_k}(y_k - \gamma_k - \boldsymbol{x}_k\boldsymbol{\beta}_k - \boldsymbol{w}_i\zeta_k), k = 1, 2$, given by (2.6). We compute this joint PDF using the approximation formulas for the Gaussian copula and for F_1 and F_2.

2.3.1.3 Bayesian inference

Observations $(y_{i1}, y_{i2}), i = 1, \ldots, n$ from (2.14), provide the likelihood function as

$$L(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}) = \prod_{i=1}^{n} f(y_{i1}, y_{i2}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}).$$

Given priors for the vector of marginal parameters $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ and the correlation ρ_{12}, the Bayesian machinery produces posterior and predictive distributions for various inferences.

We use the following simple prior distributions for the parameters of MZI-SFM:

$$
\begin{aligned}
&(\gamma_k, \boldsymbol{\beta}_k, \boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\delta}_k) \sim N(\mathbf{0}, \sigma_k^2 I_d), \quad k = 1, 2,\\
&\sigma_{\nu_k}^{-2}, \sigma_{u_k^o}^{-2} \sim \text{Gamma}(a, b), \quad k = 1, 2,\\
&\rho_{12} \sim U(-1, 1),
\end{aligned}
$$

where all parameters are independent and the dimension d is determined by the number of components of $(\gamma_k, \boldsymbol{\beta}_k, \boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\delta}_k)$. The precision parameters $\sigma_{\nu_k}^{-2}$ and $\sigma_{u_k^o}^{-2}$ have gamma distributions with parameters a and b. The uniform prior for the correlation parameter ρ_{12} is non-informative. We assign large values for σ^2 and small values for a and b so that the normal and gamma prior distributions are nearly flat, reflecting "little information" about the parameters.

The data $\boldsymbol{D}$ (set of observations) updates the priors via the Bayes rule to posterior distribution $\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}|\boldsymbol{D})$. The effects of priors on posterior distributions vanish when $\boldsymbol{D}$ includes a large number of observations. The posterior distribution is obtained via approximation by Markov Chain Monte Carlo (MCMC) simulations. We use the Metropolis-Hastings (M-H) algorithm which does not require derivation of conditional distributions of the parameters (Chib & Greenberg, 1995). The M-H algorithm can be implemented in R and python, among other statistical packages. We have written python codes for M-H algorithm implementation of our MZI-SFM.

The posterior distribution $\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}|\boldsymbol{D})$ is approximated from the MCMC draws $\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m, m = 1, \ldots, M$. It has been shown in the literature that the M-H algorithm converges so that with a large number of draws M, the approximation is accurate (Chib & Greenberg, 1995); diagnostics for convergence of the MCMC are available.

The posterior distributions of the MZI-SFM parameters provide tools for inferences. The Bayes Factor (BF) quantifies the evidence that data provides in favor of a hypothesis. Suppose that H_1 is a hypothesis about a parameter that we want to test against H_2. Let $P(H_\ell), \ell = 1, 2$ be the prior probability of H_ℓ. The data $\boldsymbol{D}$ updates these probabilities to the posterior probabilities $P(H_\ell|\mathbf{D}), \ell = 1, 2$, along with updating the priors for the parameters to the posterior distributions. The BF in favor of H_1 is given by the ratio of the posterior to the prior odds:

$$BF_{12} = \frac{P(H_1|\boldsymbol{D})/P(H_2|\boldsymbol{D})}{P(H_1)/P(H_2)}.$$

For example, consider testing $H_1 : \beta_{1k} > 0$ against $H_1 : \beta_{1k} > 0$ under the prior $\boldsymbol{\theta}_k \sim N(\mathbf{0}, \sigma_k^2 I)$. Then $P(H_1) = P(H_2) = 0.5$ and the BF is the posterior odds given by

$$BF_{12} = \frac{P(\beta_{1k} > 0|\boldsymbol{D})}{P(\beta_{1k} < 0|\boldsymbol{D})}.$$

This index can easily be computed using the MCMC draws from the posterior distribution of each parameter.

The interpretation of BF for statistical significance is according to the Jeffreys' scale for the grade of evidence (Kass & Raftery, 1995) as shown in Table 2.1.

TABLE 2.1: Jeffreys' scale for the grade of evidence (Kass & Raftery, 1995).

Evidence	BF range	Asterisk
Bare mention	$1 < \text{BF} < 3.2$	
Substantial	$3.2 \leq \text{BF} < 10$	*
Strong	$10 \leq \text{BF} < 100$	**
Decisive	$\text{BF} \geq 100$	***

The posterior predictive distribution of new outcomes of (Y_1, Y_2) is by

$$f((y_1, y_2)|\boldsymbol{D}) = \int f((y_1, y_2)|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12})\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}|\boldsymbol{D})\, d\boldsymbol{\theta}_1\, d\boldsymbol{\theta}_2\, d\rho_{12}, \tag{2.15}$$

where $\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}|\boldsymbol{D})$ is the joint posterior distribution of all MZI-SFM. Computation algorithm for the predictive distribution is as follows. Each MCMC draw $\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m$ from the posterior distribution $\pi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \rho_{12}|\boldsymbol{D})$ provides the distribution conditional distribution $f((y_1, y_2)|\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m)$. Then $f((y_1, y_2)|\boldsymbol{D})$ is approximated by the average of the conditional distributions over all T draws.

The predictive distribution of sales (S_1, S_2) are obtained from (2.15) by the transformations $S_k = \exp(Y_k), k = 1, 2$. This predictive distribution $g((s_1, s_2)|\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m)$ is approximated by the average of $g((s_1, s_2)|\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m)$ over all MCMC draws $m = 1, \ldots, M$.

2.3.2 Predictive distribution of profit

The profit function per purchase order, Π, is function of the sales of the two complementary product and the variables defined in Table 2.2. The set of price components for each product, denoted as $\Lambda_k = \{PR_k, FR_k, QD_k, CD_k\}$, includes the price of the product PR_k at the time of order, the freight charge for the product FR_k, the quantity discount for the product QD_k, and the contract discount for the product CD_k. The costs of the order, denoted $\Psi_k = \{C_k, FC_k, TC_k\}$, includes the unit cost C_k of the product, the tariff cost of the product TC_k, and the freight cost FC_k of the product.

The profit function per purchase order, Π, is the following function of the sales of the two complementary products:

$$\Pi(s_1, s_2|\Lambda_1, \Lambda_2, \Psi_1, \Psi_2) = \sum_{k=1}^{2}\{FR_k - FC_k + S_k[(PR_k - QD_k - CD_k) - (C_k + TC_k)]\}. \tag{2.16}$$

where (S_1, S_2). The uncertainty about the sales of products in the future, (S_1, S_2), induces uncertainty about the profit of a future order. The predictive distribution of profit is obtained from the predictive distribution of sales (S_1, S_2) by the transformations $(S_1, S_2) \to \Pi$ according to (2.16). This predictive distribution is approximated by the average of the conditional distributions of the profit $\Pi(s_1, s_2|\Lambda_1, \Lambda_2, \Lambda_2, \Psi_1, \Psi_2, \boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m)$ over all MCMC draws $(\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \rho_{12}^m),\ t = 1, \ldots, T$.

Table 2.2: Description and notations for the variables of the profit function.

Type	Variable	Notation
Price components	Price	PR
Λ	Freight charge	FR
	Quantity discount	QD
	Contract discount	CD
Costs	Unit cost	C
	Freight cost	FC
	Tariff	TC

2.3.3 Notations and variables of MZI-SFM

2.4 Empirical results

A national retailer provided us order-level sales data from January 2018 to January 2020, which includes tariffs and several other variables. We use data on the variables listed in Table 2.3 to estimate our MZI-SFM for two complementary products, A and B. There are 2,463 orders both products A and B and about 48 percent of orders are impacted by tariffs. The predictor variables are four types: product characteristics, media exposure variables, inefficiency variables including salesperson efficiency in different purchase channels and contract types, and delivery channels characteristics.

Table 2.3 presents the descriptions and notations for the variables (details about the variables are given in the appendix). Predictors of the frontier (maximum potential demand) are product characteristics, media exposure variables, and seasonality. The product characteristics include tariff, price, freight charge and two types of discounts, and proprietary item. The tariff indicator takes value one if the product is impacted by

TABLE 2.3: Description and notations for the variables of MZI-SFM.

Type	Variable	Notation
Response	Log sales	y
Frontier		
Product characteristics ($\boldsymbol{x}$)	Tariff indicator	TI
	Price	PR
	Freight charge	FR
	Quantity discount	DQ
	Contract discount	DC
	Proprietary item indicator	PI
Media Exposure ($\boldsymbol{w}$)	Website visits	WC
	Catalog index	CA
	Online advertising	AD
Seasonality	August and September indicator	AS
Inefficiency variables		
Salesperson performance (z_1, z_2)	Phone salesperson	PS
	Outside salesperson	OS
Contract type (z_3, z_4)	Contract type 1	$CT1$
	Contract type 2	$CT2$
Full efficiency variables		
Delivery channels ($\boldsymbol{v}$)	Indicator variables for:	
	Distribution center 1	DC_1
	Distribution center 2	DC_2
	Distribution center 3	DC_3

a tariff, and zero otherwise. The quantity discount is offered at the time of order and the contract discount is offered based on the type of the contract that a customer has with the retailer. The proprietary item variable takes value one if the product is private-label, and zero otherwise. The seasonality variable is the indicator of the order is fulfilled during the high demand season (August and September). The retailer has three media channels for generating sales: website, catalogs, and online advertising. The media exposure variables are defined with time lags to allow time for kicking in sales (details are shown in the appendix).

Predictors of inefficiency in Table 2.3 are related to four purchase channels.

- Outside Sales: An order that is placed through a sales representative in person. This channel tends to include larger orders.
- Phone-Phone: An order that is placed solely through phone contacts.
- Web-Web: An order that is placed solely through the website without involvement of a sales representative.
- Web-Phone: An order that is resulted from the customer's visits of the firm's website and contacts with a sales representative.

Outside sales representative performance is for an order which is placed through the Outside Sales channel. Phone sales representative performance is for the salesperson of an order which is placed through Phone-Phone or Web-Phone channels. Any order that is placed through any of the above channels is under one of the two types of contracts offered by the retailer.

The predictors of the probability of the firm's full efficiency is the function of three delivery centers located across the country.

2.4.1 Empirical results

We report the means, standard deviations, and medians of the posterior distributions of all parameters. The posterior mean and median are the Bayes estimates of the coefficient under the squared error and the absolute error loss functions, respectively. We also report the BF for the evidence provided by the data in favor of the directional hypothesis that the parameter is positive when the median is positive or negative when the sign of the median is negative (in all cases the signs of the posterior mean and medians are the same; the Asterisks signify the grade of evidence according to Table 2.1).

2.4.1.1 Model parameters

Table 2.4 shows the results for the parameters of the frontier. The results for product characteristics are as follows. The Bayes estimates of the coefficients of tariff for both products are negative. The data provides a decisive evidence for the coefficient being negative for product A and strong evidence for product B. These results confirm that, given all other variables in our MZI-SFM, imposing tariffs is associated with decrease in demands for these products. The Bayes estimates of the coefficients of the price for both products are negative. The grade of evidence for the coefficient being negative is decisive for product A and strong for product B. The Bayes estimates of the coefficients of freight for both products are negative with a decisive grade of evidence. The Bayes estimates of the coefficients of both types of discounts are positive for both products, all with decisive evidence. These results for the tariff, price, freight, and discounts are in accord with the economic theory of negative relationship between quantity demanded and price. Passing the tariff and freight costs to customers is in effect increasing the purchase price and offering discounts is in effect decreasing the purchase price that the customer must pay a product. However, as will be shown in sequel, the mix of these variables in a given purchase price for the consumer can have different effects on the firm's bottom line (profit).

Table 2.4 shows that the Bayes estimates of the coefficients of the propriety item is negative for product A with a decisive grade of evidence, and positive for product B with a substantial grade of evidence. These results, in the context of all variables present in the MZI-SFM, suggest that private-label items for the two products have opposite effects on their demands.

The results for the seasonality indicator shows that the MZI-SFM captures the high demand months.

The results shown for the media exposure channels in Table 2.4 shows that the Bayes estimates of the coefficients for the all three media channel variables are positive. These results indicate that exposures for the products can be effective to generate demands for these product. However, the data provides various levels of evidence for the coefficients being positive. For product A, the evidence for the coefficient of catalog being positive is substantial. This also holds for the coefficient of the website visits, but the evidence for the coefficient of online advertising being positive is decisive. For product B,

TABLE 2.4: Posterior statistics for parameters of frontier.

	Product A			Product B		
Parameter	Mean (SD)	Median	BF	Mean (SD)	Median	BF
Intercept	0.79(0.204)	0.79		0.62(0.157)	0.62	
Product characteristics						
Tariff	−0.18(0.043)	−0.18	> 100***	−0.05(0.035)	−0.05	17.18**
Price	−0.38(0.121)	−0.38	> 100***	−0.18(0.126)	−0.18	10.77**
Freight	−2.30(0.277)	−2.30	> 100***	−1.71(0.229)	−1.71	> 100***
Quantity discount	18.80(0.640)	18.80	> 100***	15.73(0.611)	15.74	> 100***
Contract discount	2.74(0.241)	2.73	> 100***	1.89(0.167)	1.89	> 100***
Proprietary item	−0.09(0.046)	−0.10	> 100***	0.04(0.039)	0.04	4.00*
August and September	0.07(0.043)	0.07	> 100***	0.03(0.034)	0.03	3.88*
Media channel variables						
Catalog	0.07(0.062)	0.07	5.90*	0.02(0.048)	0.02	2.28
Website visits	0.09(0.072)	0.09	7.00*	0.17(0.063)	0.17	> 100***
Online advertising	0.91(0.249)	0.92	> 100***	0.04(0.191)	0.04	1.25
Overall media efforts	0.74(0.074)	0.72	> 100***	0.06(0.053)	0.04	> 100***

the evidence for the coefficient of website visits being positive is decisive, but for the other media channels is not more than a bare mention. These findings provide insights for the managers to make strategic decisions for improving the effectiveness of media channels to generate demand for these products. The table also includes summary statistics for the posterior distribution of the three media channels together for the two products computed using $\boldsymbol{w}'_{ik}\boldsymbol{\eta}_k$. The Bayes estimates for both products are positive with a decisive grade of evidence. These results seemed to be derived from online advertising for product A

Table 2.5: Posterior statistics for inefficiency variables.

	Product A			Product B		
Parameter	Mean (SD)	Median	BF	Mean (SD)	Median	BF
Inefficiency variables						
Intercept	−1.15(0.476)	−1.13		−1.37(0.497)	−1.33	
Salesperson performance						
Outside sales	−0.99(0.871)	−0.99	5.66*	−2.05(0.809)	−2.07	> 100***
Phone sales	−0.85(0.634)	−0.83	6.69*	−0.01(0.996)	−0.01	1.00
Contract type						
Contract type 1	−0.66(0.824)	−0.65	0.72	−0.33(0.884)	−0.30	1.76
Contract type 2	−0.34(0.642)	−0.26	2.66	0.10(0.528)	0.16	1.32
Perfect efficiency variables						
Distribution center 1	0.18(0.743)	−0.65	1.33	−0.51(0.799)	−0.30	1.86
Distribution center 2	−1.47(0.692)	−0.26	> 100***	−1.41(0.744)	0.16	> 100***
Distribution center 3	0.77(0.764)	−0.65	3.55*	−0.33(0.889)	−0.30	1.25
Correlation, variation						
$\sigma_{u_k^o}$	0.97(0.461)	0.86		0.96(0.467)	0.84	
σ_{ν_k}	0.83(0.015)	0.83		0.65(0.012)	0.65	
ρ_{12}	mean(SD): 0.28(0.019), Median: 0.28 , $BF > 100$					

and website visits for product B. Nonetheless, the positive association between the firm's overall media efforts and the demand suggested by our model should be good news for the management.

Table 2.5 gives the posterior results for other parameters of the MZI-SFM. The

Bayes estimates of coefficients of all inefficiency variables are negative with one exception, the contract type 2 for product B. The grades of evidence, however, are different. The grades of evidence for the coefficients of the outside salesperson performance being negative are substantial for product A and decisive for product B. These results of our MZI-SFM suggest that the outside sales channel and inefficiency are negatively associated. The grade of evidence for the coefficient of the phone salesperson performance being negative is substantial for product A, but for product B the grade of evidence for the negativeness of the corresponding coefficient just border to the bare mention level. These results suggest that the salesperson's performance in Web-Phone & Phone-Phone channels and inefficiency are negatively associated, data does not provide any evidence in favor of the negative association for the product B. This suggests that the company could benefit from providing training workshops for the phone sales representatives. The data does not provide evidence for the negative coefficients of contracts and inefficiency.

Table 2.5 also shows the summary statistics for the posterior distribution of the standard deviations of the base inefficiency and noise, and the correlation between the compound deviations of the sales from the demand for these two products. The Bayes estimate of the correlation coefficient is positive which confirms that products A and B are complementary.

2.4.1.2 Comparison of purchase channels

We compare the purchase channels in terms of the stochastic order between random variables defined in the first Essay as follows. Let X_1 and X_2 be two random variables with survival functions $S_k(x) = P(X_k > x), k = 1, 2$. Then, X_1 (or its distribution) is said to be stochastically less than X_2 (or its distribution) if $S_1(x) \leq S_2(x)$ for all x is the support of the distributions. In economics and decision analysis, this notion of stochastic order is called the first order stochastic dominance. The stochastic dominance provides stronger comparisons than the mean for the purchase channels. The stochastic ordering implies that $E(X_1) \leq E(X_2)$ and every αth quantile of X_1 is smaller than the αth quantile of X_2.

Figure 2.2 compares plots of the survival functions of the inefficiency for the orders received through each of the four purchase channels. The following results for product A are evident.

- The inefficiency of the Web-Web channel stochastically dominates the inefficiencies of the other channels. This result suggests that for this product the involvement of a salesperson can be more efficient than relying entirely on online shopping.
- The inefficiency of the Web-Phone channel is dominated by the inefficiencies of the other channels. This result suggests that for this product combinations of the online and a salesperson's involvement can be more efficient than the other three options.
- The inefficiencies of the Phone-Phone and Outside purchase channels lack stochastic dominance.

Figure 2.2 gives the following results for product B.

- The inefficiency of the Outside-sales is dominated by the inefficiencies of each of the other three channels, indicating that this channel is the most efficient channel for

Table 2.6: Posterior statistics for the inefficiency purchase channels.

	Number	Product A			Product B		
Channel	of orders	Mean	2.5^{th}	97.5^{th}	Mean (SD)	2.5^{th}	97.5^{th}
Outside-Sales	811	0.25 (0.130)	0.06	0.56	0.18 (0.099)	0.04	0.40
Phone-Phone	779	0.24 (0.123)	0.07	0.54	0.22 (0.118)	0.06	0.49
Web-Web	492	0.27 (0.136)	0.08	0.59	0.22 (0.114)	0.06	0.48
Web-Phone	381	0.23 (0.125)	0.05	0.53	0.23 (0.164)	0.05	0.59

product B. This can be due to the fact that product B includes more expensive items and customers prefer to purchase expensive items via person-to-person contacts.

- The inefficiencies of the other three purchase channels lack stochastic dominance.

The stochastic order results displayed in Figure 2.2 imply that the expected value and every αth quantile are ordered accordingly. Table 2.6 illustrates this property for the means and two quantiles (2.5th and 97.5th percentiles) of the inefficiency distributions for the purchase channels. For product A, the mean and quantiles for the Web-Web channel are larger than the other channels and the mean and quantiles for the Web-Phone channel are smaller than the other three channels. For product B, the mean and quantiles for the outside channel are smaller than the other channels. Such consistent relationships do not hold for the inefficiency distributions that lack stochastic order. For example, the mean and 97.5th percentile for the Outside sales channel are larger than those for the Phone-Phone channel while the reverse holds for the 2.5th percentile. Table 2.6 also shows the standard deviations of the inefficiency distributions, which are unrelated to the stochastic order.

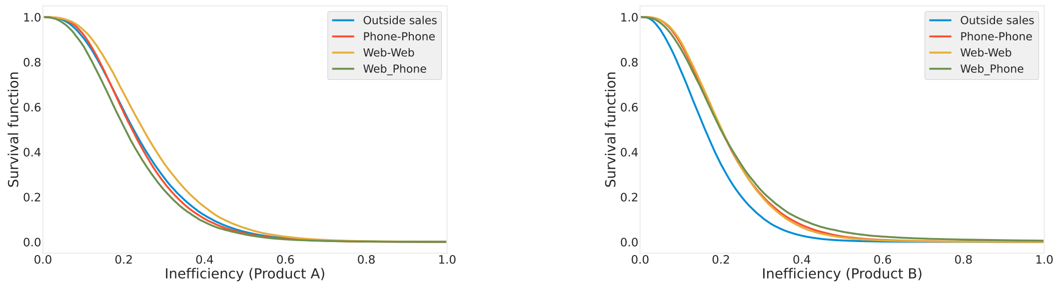

FIGURE 2.2: Survival functions of the inefficiency for the orders received for product A (left) and product B (right) through purchase channels.

2.4.1.3 Comparison of delivery channels

In our MZI-SFM, the probability of full efficiency is determined by the delivery channels. We also compare the posterior distributions of this probability in terms of the stochastic order.

Figure 2.3 compares plots of the survival functions of the posterior distributions of the probability of full efficiency for the orders delivered through each of the three distribution centers. Stochastic dominance for probabilities of full efficiency is evident for both products. The distribution for Center 3 dominates the distributions for the other two channels and the distribution for Center 2 is dominated by the distributions for the other two channels. That is, Center 3 is the most probable delivery channel and Center 2 is the least probable delivery channel to be fully efficient.

Center 2 is the main delivery channel, which is close to the firm's headquarter and fulfills more orders than the other centers, as shown in Table 2.5. This table also gives the means, standard deviations, and two quantiles (2.5th and 97.5th percentiles) of

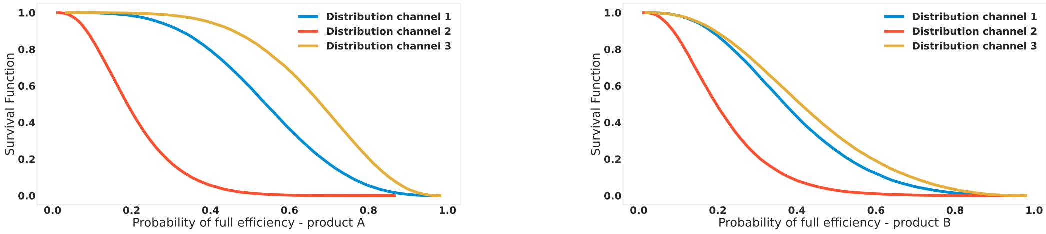

FIGURE 2.3: Survival functions of the posterior distributions of the probability of full efficiency for delivery of products A (left) and B (right) by three distribution channels.

the posterior distributions of the probability of full efficiency for the delivery channels. On average, an order from Center 1 is fully efficient with a probability of 0.54 for product A and 0.38 for product B, while the corresponding probabilities for Center 2 are 0.38 for product A and 0.22 for product B, and for Center 3 are 0.66 for product A and 0.43 for product B. Note that for product A, the means and standard deviations are ordered reversely, but for product B these two measures order similarly.

2.4.2 Revenue management strategies

Our focal retailer indicated an interest in studying profit functions under some of the scenarios for combinations of charges to customers (elements of Λ in the profit function). The retailer provided us several scenarios. We illustrate applications of our model for revenue management using examples of the scenarios. Our comparison is based on the stochastic order of the predictive distributions of the profit under the given scenarios.

1. The retailer provided scenarios to pass a portion of the tariff costs to customers through increasing price and/or freight charge for either one or both products. We

compare predictive distributions of profit for the following scenarios to pass \$150 of tariff cost to customers.

(a) Current scenario: The retailer's most recent values of Λ.

(b) Scenario 1: Increases prices of both products by \$75.

(c) Scenario 2: Increases the freight charges of both products by \$75.

(d) Scenario 3: Increase the price of product B by \$150. Recall that product B is the more expensive one.

(e) Scenario 4: Increase the freight charge of product B by \$150.

Figure 2.4 shows the plots of the survival functions of the predictive distributions of profit under the above five scenarios. The profit under Scenario 1 stochastically dominates the profit for all other scenarios, followed by the stochastic dominance of the profit under Scenario 3 over the profit under the other three scenarios. The profit under the current scenario, Scenario 2, and Scenario 4 lack stochastic dominance. These results suggest that passing the tariff costs to customers by increasing prices of two products by a smaller amount is more beneficial for the retailer rather than increasing either the price of each product or the freight charge by a larger amount. Table 2.7 gives the means, standard deviations, and several percentiles of the predictive distributions of the profit under these scenarios in dollar.

2. The retailer provided scenarios to change the mix of the current prices and discounts and for passing a portion of the tariff costs to customers through increasing price

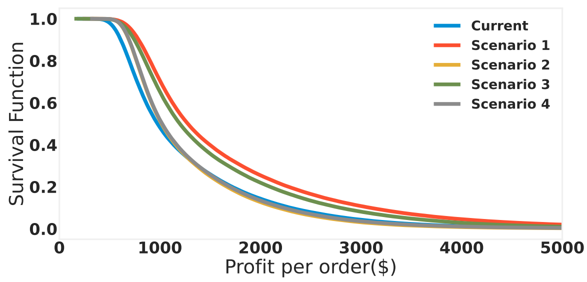

FIGURE 2.4: Predictive distributions of profit for passing tariff cost to customers through adjusting the price and/or freight charge according to five scenarios.

TABLE 2.7: Summary statistics for predictive distributions of profit for passing and not passing tariff costs to customers through increasing price or freight for either one or both products according to five scenarios.

	Mean	SD	2.5^{th}	25^{th}	50^{th}	75^{th}	97.5^{th}
Current scenario	1265.84	857.52	511.38	733.62	976.51	1531.47	3428.68
Scenario 1	1683.76	1242.14	656.03	953.01	1277.73	2021.21	4715.23
Scenario 2	1267.97	751.67	603.28	801.58	1016.21	1499.16	3156.24
Scenario 3	1539.85	1011.07	621.61	899.82	1198.18	1864.90	4117.77
Scenario 4	1292.00	827.80	605.51	804.69	1022.85	1522.80	3289.27

and/or discount for either one or both products. We compare predictive distributions of profit for the following scenarios.

(a) Current scenario: The retailer's most recent values of Λ.

(b) Scenario 1: Increase the price and contract discount for product B by \$50.

(c) Scenario 2: Increase the price and contract for product B by \$100 and \$50, respectively.

(d) Scenario 3: Increase prices of both products by \$50 and discount for product A by \$100.

(e) Scenario 4: Increase the price of product A by \$50, price of product B by \$100 and discount for product B by \$50.

Scenarios 1 and 3 change the mix of the current prices and discounts without passing additional cost to customers. Scenarios 2 and 4 pass \$50 costs of tariff to customers.

Figure 2.5 shows the plots of the survival functions of the predictive distributions of profit under the above five scenarios. The following patterns are emerging.

- The profit under Scenario 4 stochastically dominates the profit for all other scenarios, followed by the stochastic dominance of the profit under Scenario 2 over the profit under the three scenarios which do not pass additional tariff costs to customers. These results indicate that passing tariff costs by increasing the price of more expensive product B can be more beneficial for the retailer than increasing the price of this product twice as much a discount for it.

- Among the three scenarios which do not pass additional tariff costs to customers, the profit under Scenario 1 stochastically dominates the profit for all other scenarios. This result suggests that increasing the price and contract discount for product B in equal amounts can be beneficial for the retailer.

- The profit under the current scenario and Scenario 3 lack stochastic dominance.

Table 2.8 gives the means, standard deviations, and several percentiles of the predictive distributions of the profit under these scenarios in dollar.

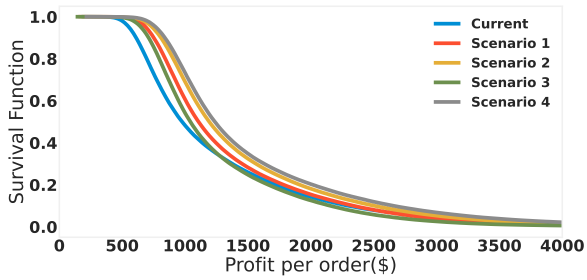

FIGURE 2.5: Predictive distributions of profit for passing and not passing tariff costs to customers through increasing price and/or discount for either one or both products according to five scenarios.

TABLE 2.8: Summary statistics for predictive distributions of profit for passing and not passing tariff costs to customers through increasing price and/or discount for either one or both products according to five scenarios.

	Mean	SD	2.5^{th}	25^{th}	50^{th}	75^{th}	97.5^{th}
Current scenario	1265.84	857.52	511.38	3428.68	733.62	976.51	1531.47
Scenario 1	1374.37	970.85	639.13	3333.19	882.35	1107.03	1598.21
Scenario 2	1474.04	1013.62	686.72	3581.13	949.34	1192.41	1719.39
Scenario 3	1273.63	800.83	602.34	3012.27	831.66	1041.91	1493.22
Scenario 4	1553.66	1201.08	714.01	3868.15	985.34	1236.89	1798.73

2.5 Concluding remarks and future research

Forecasting sales is vital for developing revenue management strategies. This essay has developed an econometric model for forecasting sales for omnichannel retailers that face the burden of tariff costs on complementary products. The tariff has been a continuing serious problem for retailers in the U.S. During the last five decades, five of the seven United States presidents have imposed tariffs or surcharges on imports. The latest in the series was $283 billion in 2018. This study was motivated by the experience of a national retailer in the Midwest that encountered on average a 35 percent increase in tariffs in 2018 and made various attempts to protect its profit margin.

The operations literature usually assumes that firms are efficient to meet the maximum potential demands for their products and sales can be less than the demand due only to shortage of inventory. Econometric literature offers models for problems where full efficiency assumption may or may not hold. We have used this approach to develop an econometric model for sales of complementary products of retailers that operate multi-channels for the exposure, multi-channels for purchase, and multi-channels for delivery of the products. Our focal retailer provided us data enabling us to examine revenue management strategies that pass tariff costs to customers.

The class of econometric models that do away from the assumption of full efficiency of firms is called the Stochastic Frontier Model (SFM). Variants of SFM decompose a performance variable into three parts: the frontier defined by the maximum potential output based on given inputs, a random term with a distribution on the strictly positive or nonnegative that captures the inefficiency to meet the frontier, and a random noise term. We contribute to the SFM literature by developing a new SFM, called Multi-product-Zero-Inefficiency SFM (MZI-SFM), which adapts and combines a few exiting variants for sales of complementary products of an omnichannel retailer. Recently, Caro et al. (2020) have suggested that demand forecasting should avoid aggregate demand or panel data to predict future aggregate sales. We have used order-level data to forecast sales at the disaggregated level for each product while allowing for correlations between SFMs for the individual products.

The inefficiency term of our MZI-SFM has generalized the usual definition of lost sales to a random deviation of sales (observed output) from the maximum potential demand that could have been realized given the product's characteristics (price, freight charge, tariff cost, among others) and the exposure of the product through the firm's media channels. The inefficiency occurs for various reasons, such as shortage, salesperson's performance, and the delivery channel's performance. The zero-inefficiency feature of our MZI-SFM allows for the possibility that the firm can be fully efficient in fulfilling some orders. Correlations between SFMs for the individual products are modeled by a copula.

We used the Bayesian approach for our empirical analysis, which includes uncertainty about the model parameters. In this approach, entire posterior distributions of the model parameters are available for estimation and testing. For each parameter, we reported the posterior mean and median, which are Bayes estimates of the parameter as the optimal decisions under the squared error and absolute error loss functions. We also reported the Bayes Factor as the evidence for the directional hypothesis that the parameter being positive or negative according to the sign of the median.

2.5.1 Summary of findings and implications

The results for all coefficients of variables in the frontier are in accord with economic theory and marketing literature. The data provides decisive evidence that the maximum potential demand that could have been materialized is negatively associated with every variable that represents a cost to the customer. These findings provide insights for the managers to make pricing decisions. The evidence for the positive association between the maximum potential demand and the firm's overall media efforts is decisive, which should be good news for the management of our focal firm. The results for each of the two complementary products that we considered, suggest that the result for the overall media efforts seemed to be derived by online advertising for one product and by website visits for the other more expensive product. These findings provide useful information for the management in making strategic decisions about the firm's media efforts in terms of the media channels.

The probability distribution of the random inefficiency term in our MZI-SFM includes for purchase channel variable (salesperson's performance), contract type of the purchase order, and delivery channel that fulfills the order. A salesperson is involved in orders generated in person called "Outside sales" or in orders generated via "Phone sales" which can be solely phone contacts (called "Phone-Phone" channel), or a combination of phone and online contacts called (called "Web-Phone" channel). The Bayes estimates of the coefficients for both types sales for both products are negative, which reflects higher salesperson's performance scores are associated with lower inefficiency in meeting the maximum potential demand that could have been generated for orders. However, the grades of evidence for such negative association vary for the type of the sales (Outside sales, Phone sales) and the type of the product. For the case of the more expensive product, the data provides decisive evidence in favor of the negative association between the outside salesperson's performance and inefficiency. Our model relates the probability of full efficiency in meeting the maximum potential demand to the delivery channels that fulfill the order. Among the three delivery channels of the firm, the date provides decisive evidence for the negative association between the distribution center located close to the firm headquarter and the probability of full efficiency.

The posterior distributions of our MZI-SFM's parameters allowed computing the posterior distributions of inefficiency for each of the firm's purchase channels (Outside sales, Phone-Phone (phone contacts only), Web-Phone combination, and Web-Web (online only)) and the posterior distribution of the probability of full efficiency for each of the firm's three delivery channels. Having these distributions on hand, we compared channels based on the notion of stochastic order between random variables. The stochastic order comparison is stronger than the comparison of the expected value (which is widely used in the operations literature), in that the stochastic ordering implies the same order for the expected values and for every corresponding quantiles of the random variables. Our findings suggest that for the less expensive of two products under study, the involvement of a salesperson can be more efficient than relying entirely on online shopping. Furthermore, for this product combinations of the online and a salesperson's involvement can be more efficient than the other three options. For the case of more expensive products, customers prefer the Outside-sales channel which relies solely on person-to-person contacts. The stochastic order comparisons of delivery channels revealed that the distribution center close to the firm's headquarter is the least probable to be fully efficient.

The Bayesian framework provides predictive distribution for sales of each product, given by the sales forecast model which is the average over of all sales forecast models whose parameters are the random outcomes of the posterior distributions of the

MZI-SFM parameters. The firm's profit is a function of the random outcomes of random sales outputs for the products. In this framework, the profit is endowed with a probability distribution induced by the predictive distributions of sales of the products. We computed and compared predictive distributions of profit for several scenarios provided by our focal firm for passing tariff costs to customers. Our stochastic order comparisons of profits of the scenarios that pass a portion of tariff costs through five combinations of price and freight charges suggest that passing the tariff costs to customers by increasing prices of two products by a smaller amount is more beneficial for the retailer rather than increasing either the price of each product or the freight charge by a larger amount. The stochastic comparisons of profits of five scenarios that pass a portion of the tariff costs to customers through increasing price and/or discount for either one or both products. The results suggest that passing tariff costs by increasing the price of more expensive product B can be more beneficial for the retailer than increasing the price of this product twice as much as a discount for it. Among the three scenarios which do not pass additional tariff costs to customers, the profit under Scenario 1 stochastically dominates the profit for all other scenarios. Our results also suggest that, among three scenarios with a different mix of price and discount increases without passing additional tariff costs to customers, increasing the price and contract discount for the more expensive product in equal amounts can be beneficial for the retailer.

Overall, this essay opened new directions of modeling and empirical analysis in the management science literature. The MZI-SFM can be applied to various applications in management science problems where firm's performance is based on multiple outputs and the firm may or may not be fully efficient. Many management science problems include a profit or a loss function, which is a function of the firm's performance. The Bayesian approach used in this essay provides decision-making tools in terms of the distribution of the profit or loss function rather than relying on a single summary measure, such as the expected value.

2.5.2 Future research

This essay has not addressed a few related problems which provide interesting and in part challenging topics for future research. Alternative specifications of the MZI-SFM in terms of inclusion of other variables are presently under consideration. The inclusion of variables that differentiate characteristics of online exposures of products, solely online purchase channel, and combinations of online and salesperson purchase channel will enrich the current model. Inclusion of sales of each complementary product in the SFM for other products will also enrich the current model.

An approach for modeling inefficiency that has been studied extensively in management science is Data Envelopment Analysis (DEA). Comparison of our MZI-SFM and DEA for sales of multi-product and subsequent study of revenue management strategies remains for the future.

We studied sales, maximum potential demand, inefficiency, and profit for each order. We barely touched the time series aspects of the data generating process by including a seasonal indicator and lags of the media channel variables, mainly because of a limited data span. Extending our model to a dynamic MZI-SFM would be an interesting topic for future studies. A dynamic MZI-SFM includes correlations between the compound errors within the SFMs for individual products in addition to the correlations between the SFMs for the complementary products. Implementation of a dynamic SFM requires data over a long time span. Such a model will allow studying weekly or monthly profit distributions. Researchers mainly have focused on estimating aggregated sales. However, a disaggregated model can easily provide aggregation sales and profits.

Developing an algorithm that finds the optimal values of inputs (price, discount, freight charge, and tariff costs) simultaneously is a challenging topic for future research. Reaching an analytical solution would be extremely difficult, if not impossible. However, solving this problem computationally would be a valuable contribution.

For the Bayesian analysis, we used a simple set of prior distributions. The MZI-SFM with other existing parametric and the Dirichlet process prior will be considered in future research. Developing computation codes for implementing other priors will allow studying the robustness of our results for the choice of priors.

Bibliography

[1] Dennis Aigner, CA Knox Lovell, and Peter Schmidt. Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, 6(1):21–37, 1977.

[2] Arzum Akkas, Vishal Gaur, and David Simchi-Levi. Drivers of product expiration in consumer packaged goods retailing. *Management Science*, 65(5):2179–2195, 2019.

[3] Layth C Alwan, Nader Ebrahimi, and Ehsan S Soofi. Information theoretic framework for process control. *European Journal of Operational Research*, 111(3):526–542, 1998.

[4] Mary Amiti, Stephen J Redding, and David E Weinstein. The impact of the 2018 tariffs on prices and welfare. *Journal of Economic Perspectives*, 33(4):187–210, 2019.

[5] Christine Amsler, Artem Prokhorov, and Peter Schmidt. Endogenous environmental variables in stochastic frontier models. *Journal of Econometrics*, 199(2):131–140, 2017.

[6] Christine Amsler, Peter Schmidt, and Wen-Jen Tsay. Evaluating the cdf of the distribution of the stochastic frontier composed error. *Journal of Productivity Analysis*, 52(1-3):29–35, 2019.

[7] Majid Asadi, Nader Ebrahimi, Ehsan S Soofi, and Somayeh Zarezadeh. New maximum entropy methods for modeling lifetime distributions. *Naval Research Logistics (NRL)*, 61(6):427–434, 2014.

[8] Amirsaman H Bajgiran, Mahsa Mardikoraem, and Ehsan S Soofi. Maximum entropy distributions with quantile information. *European journal of operational research*, 290(1):196–209, 2021.

[9] Mihai Banciu and Fredrik Ødegaard. Optimal product bundling with dependent valuations: The price of independence. *European Journal of Operational Research*, 255(2):481–495, 2016.

[10] Neshat Beheshti, Jeffrey S Racine, and Ehsan S Soofi. Information measures of kernel estimation. *Econometric Reviews*, 38(1):47–68, 2019.

[11] Dimitris Bertsimas, Jack Dunn, and Nishanth Mundru. Optimal prescriptive trees. *INFORMS Journal on Optimization*, 1(2):164–183, 2019.

[12] Dimitris Bertsimas and Nathan Kallus. From predictive to prescriptive analytics. *Management Science*, 66(3):1025–1044, 2020.

[13] Dimitris Bertsimas and Christopher McCord. From predictions to prescriptions in multistage optimization problems. *arXiv preprint arXiv:1904.11637*, 2019.

[14] John R Birge and Marilyn J Maddox. Bounds on expected project tardiness. *Operations Research*, 43(5):838–850, 1995.

[15] Jaap WB Bos, Michael Koetter, James W Kolari, and Clemens JM Kool. Effects of heterogeneity on bank efficiency scores. *European Journal of Operational Research*, 195(1):251–261, 2009.

[16] George EP Box and George C Tiao. *Bayesian inference in statistical analysis*, volume 40. John Wiley & Sons, 2011.

[17] Patrick L Brockett, Linda L Golden, and Kwang H Paick. Constructing a unimodal bayesian prior distribution from incompletely assessed information. In *Applying Maximum Entropy to Econometric Problems*. Emerald Group Publishing Limited, 1997.

[18] Felipe Caro and Jérémie Gallien. Clearance pricing optimization for a fast-fashion retailer. *Operations Research*, 60(6):1404–1422, 2012.

[19] Felipe Caro, A Gürhan Kök, and Victor Martínez-de Albéniz. The future of retail operations. *Manufacturing & Service Operations Management*, 22(1):47–58, 2020.

[20] Cinzia Carota, Giovanni Parmigiani, and Nicholas G Polson. Diagnostic measures for model criticism. *Journal of the American Statistical Association*, 91(434):753–762, 1996.

[21] Alessandro Carta and Mark FJ Steel. Modelling multi-output stochastic frontiers using copulas. *Computational Statistics & Data Analysis*, 56(11):3757–3773, 2012.

[22] Boxiao Chen, Xiuli Chao, and Cong Shi. Nonparametric learning algorithms for joint pricing and inventory control with lost-sales and censored demand. *Available at SSRN 2836057*, 2017.

[23] Qi Chen, Stefanus Jasin, and Izak Duenyas. Real-time dynamic pricing with minimal and flexible price adjustment. *Management Science*, 62(8):2437–2455, 2016.

[24] Stanley F Chen and Ronald Rosenfeld. A survey of smoothing techniques for me models. *IEEE transactions on Speech and Audio Processing*, 8(1):37–50, 2000.

[25] Siddhartha Chib and Edward Greenberg. Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335, 1995.

[26] Sanghoon Cho, Mark Ferguson, Pelin Pekgun, and Jongho Im. Robust demand estimation of choice models for personalized offers using sales transaction data. *Available at SSRN 3598259*, 2020.

[27] Howard Hao-Chun Chuang, Rogelio Oliva, and Gregory R Heim. Examining the link between retailer inventory leanness and operational efficiency: Moderating roles of firm size and demand uncertainty. *Production and Operations Management*, 28(9):2338–2364, 2019.

[28] Charles J Corbett and Kumar Rajaram. A generalization of the inventory pooling effect to nonnormal dependent demand. *Manufacturing & Service Operations Management*, 8(4):351–358, 2006.

[29] John G Cragg. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, pages 829–844, 1971.

[30] Peter J Danaher, Tracey S Danaher, Michael Stanley Smith, and Ruben Loaiza-Maya. Advertising effectiveness for multiple retailer-brands in a multimedia and multichannel environment. *Journal of Marketing Research*, 57(3):445–467, 2020.

[31] Lingxiu Dong and Panos Kouvelis. Impact of tariffs on global supply chain network configuration: models, predictions, and future research. *Manufacturing & Service Operations Management*, 22(1):25–35, 2020.

[32] Nader Ebrahimi, Ehsan S Soofi, and Refik Soyer. Multivariate maximum entropy identification, transformation, and dependence. *Journal of Multivariate Analysis*, 99(6):1217–1231, 2008.

[33] Joseph Engelberg, Charles F Manski, and Jared Williams. Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business & Economic Statistics*, 27(1):30–41, 2009.

[34] José TAS Ferreira and Mark FJ Steel. Model comparison of coordinate-free multivariate skewed distributions with an application to stochastic frontiers. *Journal of Econometrics*, 137(2):641–673, 2007.

[35] Kris Johnson Ferreira, Bin Hong Alex Lee, and David Simchi-Levi. Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management*, 18(1):69–88, 2016.

[36] Kris Johnson Ferreira, David Simchi-Levi, and He Wang. Online network revenue management using thompson sampling. *Operations research*, 66(6):1586–1602, 2018.

[37] Adam J Fleischhacker and Pak-Wing Fok. An entropy-based methodology for valuation of demand uncertainty reduction. *Decision Sciences*, 46(6):1165–1198, 2015.

[38] Guillermo Gallego and Ilkyeong Moon. The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society*, 44(8):825–834, 1993.

[39] Fei Gao and Xuanming Su. Omnichannel retail operations with buy-online-and-pick-up-in-store. *Management Science*, 63(8):2478–2492, 2017.

[40] Fei Gao and Xuanming Su. Online and offline information for omnichannel retailing. *Manufacturing & Service Operations Management*, 19(1):84–98, 2017.

[41] Paul H Garthwaite, Joseph B Kadane, and Anthony O'Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):680–701, 2005.

[42] Maria Goranova, Rahi Abouk, Paul C Nystrom, and Ehsan S Soofi. Corporate governance antecedents to shareholder activism: A zero-inflated process. *Strategic Management Journal*, 38(2):415–435, 2017.

[43] Willam Greene. Fixed and random effects in stochastic frontier models. *Journal of productivity analysis*, 23(1):7–32, 2005.

[44] William Greene. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of econometrics*, 126(2):269–303, 2005.

[45] Jim E Griffin and Mark FJ Steel. Semiparametric bayesian inference for stochastic frontier models. *Journal of econometrics*, 123(1):121–152, 2004.

[46] Eugene D Hahn and Jonathan P Doh. Using bayesian methods in strategy research: an extension of hansenet al. *Strategic Management Journal*, 27(8):783–798, 2006.

[47] Eugene David Hahn. Mixture densities for project management activity times: A robust approach to pert. *European Journal of operational research*, 188(2):450–459, 2008.

[48] Willem K Klein Haneveld. Robustness against dependence in pert: An application of duality and distributions with known marginals. In *Stochastic Programming 84 Part I*, pages 153–182. Springer, 1986.

[49] Pavithra Harsha, Ramesh Natarajan, and Dharmashankar Subramanian. A prescriptive machine learning framework to the price-setting newsvendor problem. 2019.

[50] Pavithra Harsha, Shivaram Subramanian, and Markus Ettl. A practical price optimization approach for omnichannel retailing. *Informs Journal on Optimization*, 1(3):241–264, 2019.

[51] Pavithra Harsha, Shivaram Subramanian, and Joline Uichanco. Dynamic pricing of omnichannel inventories: honorable mention—2017 m&som practice-based research competition. *Manufacturing & Service Operations Management*, 21(1):47–65, 2019.

[52] Agustín Hernández-Bastida and M Pilar Fernández-Sánchez. How adding new information modifies the estimation of the mean and the variance in pert: a maximum entropy distribution approach. *Annals of Operations Research*, 274(1):291–308, 2019.

[53] Sabiölla Hosseini, Marieluise Merz, Maximilian Röglinger, and Annette Wenninger. Mindfully going omni-channel: An economic decision model for evaluating omni-channel strategies. *Decision Support Systems*, 109:74–88, 2018.

[54] Woonghee Tim Huh, Retsef Levi, Paat Rusmevichientong, and James B Orlin. Adaptive data-driven inventory control with censored demand based on kaplan-meier estimator. *Operations Research*, 59(4):929–941, 2011.

[55] Gerrit K Janssens, Lotte Verdonck, and Katrien Ramaekers. Service-oriented performance of inventory models with partial information on unimodal demand lead-time distributions. *RAIRO-Operations Research*, 55:S1207–S1228, 2021.

[56] Edwin T Jaynes. Information theory and statistical mechanics. *Physical review*, 106(4):620, 1957.

[57] Edwin T Jaynes. Prior probabilities. *IEEE Transactions on systems science and cybernetics*, 4(3):227–241, 1968.

[58] Harry Joe. Generating random correlation matrices based on partial correlations. *Journal of Multivariate Analysis*, 97(10):2177–2189, 2006.

[59] MC Jones. On khintchine's theorem and its place in random variate generation. *The American Statistician*, 56(4):304–307, 2002.

[60] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the american statistical association*, 90(430):773–795, 1995.

[61] JHB Kemperman. Moment problems with convexity conditions i. In *Optimizing methods in statistics*, pages 115–178. Elsevier, 1971.

[62] A Ya Khintchine. On unimodal distributions. *Izvestiya Nauchno-Issledovatel'skogo Instituta Matematiki i Mekhaniki*, 2(2):1–7, 1938.

[63] Gary Koop, Jacek Osiewalski, and Mark FJ Steel. Bayesian efficiency analysis through individual effects: Hospital cost frontiers. *Journal of econometrics*, 76(1-2):77–105, 1997.

[64] TCT Kotiah and Norval D Wallace. Another look at the pert assumptions. *Management Science*, 20(1):44–49, 1973.

[65] Subal C Kumbhakar, Christopher F Parmeter, and Efthymios G Tsionas. A zero inefficiency stochastic frontier model. *Journal of Econometrics*, 172(1):66–76, 2013.

[66] Levent Kutlu, Kien C Tran, and Mike G Tsionas. A time-varying true individual effects model with endogenous regressors. *Journal of econometrics*, 211(2):539–559, 2019.

[67] Hui Li. Intertemporal price discrimination with complementary products: E-books and e-readers. *Management Science*, 65(6):2665–2694, 2019.

[68] Marvin B Lieberman and Rajeev Dhawan. Assessing the resource base of japanese and us auto producers: A stochastic frontier production function approach. *Management science*, 51(7):1060–1075, 2005.

[69] TK Littlefield Jr and PH Randolph. Reply—an answer to sasieni's question on pert times. *Management Science*, 33(10):1357–1359, 1987.

[70] Gábor Lugosi, Mihalis G Markakis, and Gergely Neu. On the hardness of inventory management with censored demand data. *arXiv preprint arXiv:1710.05739*, 2017.

[71] Dominique Makowski, Mattan S Ben-Shachar, SH Chen, and Daniel Lüdecke. Indices of effect existence and significance in the bayesian framework. *Frontiers in psychology*, 10:2767, 2019.

[72] Dominique Makowski, Mattan S Ben-Shachar, and Daniel Lüdecke. bayestestr: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, 4(40):1541, 2019.

[73] Donald G Malcolm, John H Roseboom, Charles E Clark, and Willard Fazar. Application of a technique for research and development program evaluation. *Operations research*, 7(5):646–669, 1959.

[74] Gilles Mauris. Uncertainty interval expression of measurement: possibility maximum specificity versus probability maximum entropy principles. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 386–395. Springer, 2010.

[75] Wim Meeusen and Julien van Den Broeck. Efficiency estimation from cobb-douglas production functions with composed error. *International economic review*, pages 435–444, 1977.

[76] Jacek Osiewalski and Mark FJ Steel. Numerical tools for the bayesian analysis of stochastic frontier models. *Journal of Productivity Analysis*, 10(1):103–117, 1998.

[77] Georgia Perakis and Guillaume Roels. Regret in the newsvendor model with partial information. *Operations research*, 56(1):188–203, 2008.

[78] José García Pérez, María del Mar López Martín, Catalina García García, and Miguel Ángel Sánchez Granero. Project management under uncertainty beyond beta: The generalized bicubic distribution. *Operations Research Perspectives*, 3:67–76, 2016.

[79] Jason Rhuggenaath, Paulo Roberto de Oliveira da Costa, Alp Akcay, Yingqian Zhang, and Uzay Kaymak. A heuristic policy for dynamic pricing and demand learning with limited price changes and censored demand. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 3693–3698. IEEE, 2019.

[80] Soroush Saghafian and Brian Tomlin. The newsvendor under demand ambiguity: Combining data with moment and tail information. *Operations Research*, 64(1):167–185, 2016.

[81] H Scarf. A min max solution of an inventory problem, stud ies in the mathematical theory of inventory and production (chap ter 12), 1958.

[82] Cong Shi, Weidong Chen, and Izak Duenyas. Nonparametric data-driven algorithms for multiproduct inventory systems with censored demand. *Operations Research*, 64(2):362–370, 2016.

[83] Michael Stanley Smith. Bayesian approaches to copula modelling. *arXiv preprint arXiv:1112.4204*, 2011.

[84] Lawrence V Snyder and Zuo-Jun Max Shen. *Fundamentals of supply chain theory.* Wiley Online Library, 2011.

[85] Ehsan S Soofi, Nader Ebrahimi, and Mohamed Habibullah. Information distinguishability with application to analysis of failure data. *Journal of the American Statistical Association*, 90(430):657–668, 1995.

[86] Ehsan S Soofi, Paul C Nystrom, and Masoud Yasai-Ardekani. Executives' perceived environmental uncertainty shortly after 9/11. *Computational statistics & data analysis*, 53(9):3502–3515, 2009.

[87] Shivaram Subramanian and Pavithra Harsha. Demand modeling in the presence of unobserved lost sales. *Submitted, IBM Research*, 2017.

[88] D Sun and James O Berger. Reference priors with partial information. *Biometrika*, 85(1):55–71, 1998.

[89] Dan Trietsch, Lilit Mazmanyan, Lilit Gevorgyan, and Kenneth R Baker. Modeling activity times by the parkinson distribution with a lognormal core: Theory and validation. *European Journal of Operational Research*, 216(2):386–396, 2012.

[90] Wen-Jen Tsay, Cliff J Huang, Tsu-Tan Fu, and I-Lin Ho. A simple closed-form approximation for the cumulative distribution function of the composite error of stochastic frontier models. *Journal of Productivity Analysis*, 39(3):259–269, 2013.

[91] Mike G Tsionas and Sushanta K Mallick. A bayesian semiparametric approach to stochastic frontiers and productivity. *European Journal of Operational Research*, 274(1):391–402, 2019.

[92] J René van Dorp and Thomas A Mazzuchi. Three-point lifetime distribution elicitation for maintenance optimization in a bayesian context. *Expert Judgement in Risk and Decision Analysis*, pages 147–177, 2021.

[93] RenÉ J Van Dorp and Thomas A Mazzuchi. Solving for the parameters of a beta a distribution under two quantile constraints. *Journal of Statistical Computation and Simulation*, 67(2):189–201, 2000.

[94] Peter C Verhoef, Pallassana K Kannan, and J Jeffrey Inman. From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. *Journal of retailing*, 91(2):174–181, 2015.

[95] Eric-Jan Wagenmakers, Tom Lodewyckx, Himanshu Kuriyal, and Raoul Grasman. Bayesian hypothesis testing for psychologists: A tutorial on the savage–dickey method. *Cognitive psychology*, 60(3):158–189, 2010.

[96] Yunzeng Wang. Joint pricing-production decisions in supply chains of complementary products with uncertain demand. *Operations Research*, 54(6):1110–1127, 2006.

[97] Robert L Winkler. The consensus of subjective probability distributions. *Management Science*, 15(2):B–61, 1968.

[98] Chaolin Yang, Zhenyu Hu, and Sean X Zhou. Multilocation newsvendor problem: Centralization and inventory pooling. *Management Science*, 2020.

[99] Dennis J Zhang, Hengchen Dai, Lingxiu Dong, Fangfang Qi, Nannan Zhang, Xiaofei Liu, Zhongyi Liu, and Jiang Yang. The long-term and spillover effects of price promotions on retailing platforms: Evidence from a large randomized experiment on alibaba. *Management Science*, 66(6):2589–2609, 2020.

[100] Huanan Zhang, Xiuli Chao, and Cong Shi. Perishable inventory systems: Convexity results for base-stock policies and learning algorithms under censored demand. *Operations Research*, 66(5):1276–1286, 2018.

Chapter 3

Appendix

3.1 CDF of compound deviation of the performance from the frontier (Essay 2)

The CDF of compound deviation of the firm's performance from the frontier (ϵ in 2.2) does not have analytical expression. There are different approaches in the literature that provide approximations and simulation methods (Tsay et al., 2013, and Amsler et al., 2019). Exact result is available only where $\sigma_\nu = \sigma_u$. For other cases, we use closed form approximation introduced by Tsay et al. (2013). Suppose CDF function for ϵ is as follows:

$$F_\epsilon(Q) = \int_{-\infty}^{Q} f(\epsilon)d\epsilon$$

Tsay et al. (2013) show for each value of Q, $F(Q)$ can be written as $\frac{2}{\sigma}I(Q)$, where $I(Q)$'s approximation is as following:

$$I(Q) = \quad exp(\frac{a^2c_1^2}{4b^2 - 4a^2c_2})\frac{1}{4\sqrt{b^2 - a^2c_2}}[1 - erf(\frac{-ac_1 + \sqrt{2}Q(b^2 - a^2c_2)sign(Q)}{2\sqrt{b^2 - a^2c_2}}) + \frac{erf(\frac{bQ}{\sqrt{2}})}{2b}\frac{1 + sign(Q)}{2},$$

where $c_1 = 1.09500814703333$ and $c_2 = 0.75651138383854$.

3.2 Variables descriptions (Essay 2)

Var	Description
s	Number of items sold in an order
PR	Average $(\frac{\text{Price of each item at each order - Min price of the item}}{\text{Min price of the item}})$ for items of each type at each order
FR	$(\frac{\text{Freight sell}}{\text{Extended sell price}})$ of each order
DQ	Average $(\frac{\text{Unit quantity break price of each item - Unit sell price of the item}}{\text{Unit quantity break price of the item}})$ for items of each type at each order
DC	Average $(\frac{\text{Unit advertised price of each item - Unit quantity break price of the item}}{\text{Unit advertised price of the item}})$ for items of each type at each order
TI	1 if the product sold faced tariff, 0 otherwise
PI	1 if private brand, 0 otherwise
AS	1 for Aug-Sep, 0 otherwise
WC	$\frac{\text{Average daily clicks in the last 3 weeks of order - Min average of clicks over 3 weeks}}{\text{Min average of clicks over 3 weeks}}$
CA	$\frac{\text{Total catalogs sent in the previous 3 weeks - Min catalogs sent in the 3 weeks}}{\text{Max catalogs sent in the 3 weeks}}$
AD	Advertising dollar spent during the previous month of order (in \$M)
PS	$\frac{\text{Total sales a phone representative made during 2018-2019}}{\text{Total sales made in the phone-related channels}} \times 100$
OS	$\frac{\text{Total sales an outside sales representative made during 2018-2019}}{\text{Total sales made in the outside sales channel}} \times 100$
CT_1	1 if order has contract type 1, otherwise 0
CT_2	1 if order has contract type 2, otherwise 0
DC_j	$DC_j = 1$ if order delivered in channel j, 0 otherwise

Curriculum Vitae

EDUCATION

University of Wisconsin-Milwaukee
Sheldon B. Lubar School of Business
Ph.D., Management Science (August 2021)
Major : Operations and Supply Chain Management
Minor : Business Statistics

Thesis : Essays on decision problems under uncertainty

- Maximum entropy distributions with mode and quantiles for applications to operations problem
- Bayesian prescriptive framework for complementary products: tariff strategies and channels inefficiencies

College of Engineering and Applied Sciences
M.Sc., Industrial Engineering, 2016

Thesis: Manufacturing Site Selection in the Global Context using Machine learning techniques

Sabanci University, Istanbul, Turkey
Graduate work, 2014

Sharif University of Technology, Tehran, Iran
B.Sc., Industrial Engineering, 2013

RESEARCH INTERESTS

Applications: Retail operations, efficiency analysis, digital Supply Chain, Computational Advertising, sustainability

Methodology: Prescriptive analysis, Bayesian econometrics, data-driven optimization, information theory, machine learning

PUBLICATIONS

Published paper:

Maximum entropy distributions with quantile information (with A.H.Bajgiran and E.S.Soofi), *European Journal of Operational Research*, 2020

Working paper:

Bayesian prescriptive framework for complementary products: tariff strategies and channels inefficiencies (with E.S.Soofi, target journal: *Manufacturing & Service Operations Management*)

Maximum entropy distributions with mode and quantiles (with E.S.Soofi, target journal: *European Journal of Operational Research*)

Manuscript under preparation:

Recent advances in Computational Advertising (with E.Fokoue, Rochester Institute of Technology and E.Nwankwo, Duke University)

Conferences, Workshops

Bayesian prescriptive framework for complementary products pricing with channels inefficiencies, 2020 INFORMS Annual Conference, (*Session: Bayesian data science and analytics*)

Pricing strategies for tariff on complementary products: A Bayesian prescriptive model, 2020 DSI Annual Conference (*Session: Business Analytics*)

Bayesian predictive demand model with promotion, tariff, and freight costs, INFORMS Annual Conference National Harbor, MD, November 2020

Bayesian seemingly unrelated mean and quantile stochastic frontier model for demand, Joint Statistical Meetings (American Statistical Association), August, Philadelphia, PA, August 2020

Impacts of an IoT application on supply chain collaboration: A Bayesian prescriptive framework, MSOM Annual conference, June 2020 [Abstract accepted but conference cancelled due to Covid19]

Data-Driven approach to the price-setting Newsvendor, Statistical and Applied Mathematical Sciences Institute, program on Games, Decisions, Risk and Reliability, Lehigh, NC, August 2019

Sixth Symposium on Games and Decisions in Reliability and Risk, George Washington University, Washington, DC, May 2019

Bayesian quantile regression model for the price-setting newsvendor problem, invited session at Production and Operations Management Society (POMS) Annual Conference, Washington D.C., May 2019

Maximum entropy demand models with newsvendor information, presented (by Prof. Soofi) at the invited session "Statistics in Operations" of International Symposium on Business and Industrial Statistics (ISBIS Meeting), Greece, July 2018

Newsvendor Demand Models with Partial Information, Lubar School of Business Research Seminar Series, UWM, October 2017

New maximum entropy newsvendor models, invited session, INFORMS Annual Conference, Houston, TX, October 2017

Maximum entropy newsvendor models, American Statistical Association 34th Quality and Productivity Research Conference QPRC, Storrs, CT, June 2017

Other research activities

Statistical and Applied Mathematical Sciences Institute (SAMSI)

Program on Games, Decisions, Risk and Reliability
Research group: Computational advertising (under coordination of Prof. David Banks, Duke University), August 2019 - May 2020

HONORS AND AWARDS

Sheldon B. Lubar Doctoral Scholarship (5000$), UWM, 2018, 2019 and 2020

Travel award, poster presentation at the Statistical and Applied Mathematical Sciences Institute, program on Games, Decisions, Risk and Reliability, Lehigh, NC, 2019

Travel award, Sixth Symposium on Games and Decisions in Reliability and Risk, George Washington University, Washington, DC, 2019

NSF/QPRC award for presenting in American Statistical Association-Quality and productivity Conference, Storrs, CT, 2017

Graduate chancellor award, UWM, 2015

Full scholarship for one-year graduate study, Sabanci University, 2013

Ranked among the **top 0.1%** from 400,000 applicants in the nation-wide entrance exam, admission to Sharif University of Technology, 2009

TEACHING EXPERIENCE

Instructor

Lubar School of Business, UWM

Supply Chain Analytics, 1 section, Online, Fall 2021
Supply Chain Analytics, 1 section, Online, Spring 2021
Supply Chain Analytics, 1 section, Online, Fall 2020
Supply Chain Analytics, 2 sections, Hybrid, Spring 2020
Supply Chain Analytics, 2 sections, Fall 2019

Teaching Assistant

Lubar School of Business, UWM

Introduction to Management Statistics, Fall 2018 - Spring 2019
Supply Chain Management, , Fall 2017 - Spring 2018

College of Engineering and Applied Sciences, UWM

Introduction to Engineering, Fall 2014 - Spring 2016

Sabanci University

Introduction to Optimization, Fall 2013 - Spring 2014

INDUSTRY EXPERIENCE

Lubar School of Business Industry Connect program, April - June 2020

Developed and presented Bayesian predictive algorithms for 3 years data
Developed prescriptive analysis for price, discounts, promotion and freight charge of multiple products

Allen Edmonds, WI, Sep 2015 - Jan 2016

Developed forecasting models for 60+ stores nationwide given sales data over the past 10 years
Implemented inventory optimization for stores and warehouses
Analyzed strategic decisions on transition to omni-channel retailing

Professional Service Journal Reviewer: Journal of Applied Stochastic Models in Business and Industry

Selected Graduate Courses

Empirical Methods & Data Analytics:

Doctoral seminar decision sciences (Bayesian Econometrics, two courses)
Multivariate Techniques in Management Research
Pattern Recognition-Statistical-Neural and Fuzzy Approaches (Computer science department)
Statistical Analysis (linear models)
Advanced Econometrics Methods (Time series)
Machine Learning (Sabanci University)

Operations and Supply Chain Management:

Probability Models for Operations Decisions
Analytical Modeling in Operations Management
Applied Game Theory in Supply Chain Management
Optimizing Supply Chains for Environmental Sustainability
Inventory Management under Uncertain Settings
Supply Chain Logistics Management

Operations Research & Mathematics:

Applied Stochastic Processes (Mathematics department)
Linear Programming (Sabanci University)
Optimization (Sabanci University)
Advanced Operations Research Models

Computer skills Python (Pandas, Scikit-learn, PyMC, SciPy), R, SQL, Win-Bugs, SAS, C++, JMP, WEKA

References

Ehsan S. Soofi (Advisor)
Distinguished Professor of Management Science and Statistics
Sheldon B. Lubar School of Business
University of Wisconsin-Milwaukee
Email: esoofi@uwm.edu, Phone: (414) 202-6666

Layth C. Alwan
Professor of Supply Chain, Operations Management & Business Statistics
Sheldon B. Lubar School of Business
University of Wisconsin-Milwaukee
Email: alwan@uwm.edu; Phone: (414) 229-6253

David Banks
Professor of the Practice of Statistical Science
Director of Statistical and Applied Mathematical Sciences Institute
Duke University
Email: dlbanks@duke.edu; Phone: (919) 684-3743

Refik Soyer
Professor of Decision Sciences & Statistics
The George Washington University
Email: soyer@gwu.edu; Phone: (202) 994-6445