ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE

Overcoming Hypothetical Bias In Contingent Valuation Surveys

Yakinthi G. Pavlaki
Department of Economics

University of Crete

A thesis submitted for the degree of
Doctor of Philosophy in Economics

Rethymno, Crete, Greece

September 2022

Thesis Advisor: Margarita Genius
Associate Professor, Department of Economics, University of Crete, Greece,
genius@uoc.gr

Advisory Committee Member: Vangelis Tzouvelekas
Professor, Department of Economics, University of Crete, Greece,
v.tzouvelekas@uoc.gr

Advisory Committee Member: Georgios Tsiotas
Assistant Professor, Department of Economics, University of Crete, Greece,
tsiotas@uoc.gr

ABSTRACT

Yakinthi G. Pavlaki: OVERCOMING HYPOTHETICAL BIAS IN CONTINGENT VALUATION SURVEYS

(Under the direction of Associate Prof. Margarita Genius)

This Ph.D. thesis aims to build econometric models that can overcome hypothetical bias in Contingent Valuation surveys. Within 4 interrelated Chapters, this thesis focuses on constructing a mixture model and applying stochastic frontier analysis in order to include the existence of hypothetical bias. The main idea of the proposed model of the present thesis is that in a Contingent Valuation survey there might be two types of respondents. The first one refers to respondents that answer sincerely about their WTP and the second to respondents that overstate their WTP.

The first chapter presents a literature review regarding Contingent Valuation Method (CVM) and the problem of Hypothetical Bias. Additionally, it analyzes the theoretical framework regarding the statistical models that are constructed in Chapter 2 and 3. The second and third Chapters present the proposed model that can overcome hypothetical bias for the open-ended elicitation format and for the double-bounded dichotomous choice elicitation format. The Chapters contain the theoretical background of the corresponding format, the stochastic frontier theory, which is applied to model hypothetical bias, the mixture models theory that is used in order to allow the existence of the 2 classes and the estimation procedure, including the EM algorithm.

In order to test the validity of the proposed model, simulations took place for 1000 observations and 1000 replications for several cases. Additionally in Chapter 3, after the illustration of the simulation results an empirical application is presented with real CVM data. The fourth Chapter deals with the issue of the selection of starting values for the EM algorithm. Chapter 4 begins with the literature review about the initialization techniques and the importance of the initial values. Furthermore, application of three initialization methods were applied to the proposed model for the double-bounded format and a comparison of the initialization techniques took place in order to conclude which initialization strategy performs better.

Keywords: Environmental Econometrics, Contingent Valuation Method (CVM), Open-Ended, Double-Bounded, Hypothetical Bias, Stochastic Frontier Analysis, Composed Error, Finite Mixture Models, EM algorithm, Willingness-To-Pay, Initial Values

ΠΕΡΙΛΗΨΗ

Υακίνθη Γ. Παυλάκη: ΔΙΟΡΘΩΝΟΝΤΑΣ ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΥΠΟΘΕΤΙΚΗΣ ΜΕΡΟΛΗΨΙΑΣ ΣΕ ΕΡΕΥΝΕΣ ΠΙΘΑΝΟΛΟΓΙΚΗΣ ΑΠΟΤΙΜΗΣΗΣ (CONTINGENT VALUATION)

(Υπό την καθοδήγηση της Αναπλ. Καθηγήτριας Margarita Genius)

Η παρούσα διδακτορική διατριβή μελετά πως θα ξεπεραστεί το πρόβλημα της Υποθετικής Μεροληψίας σε έρευνες που πραγματοποιούνται με την μέθοδο της πιθανολογικής αποτίμησης ή Contingent Valuation (CV). Πιο συγκεκριμένα προτείνεται ένα εναλλακτικό μοντέλο, στοχεύοντας οι εκτιμήσεις που προκύπτουν να είναι απαλλαγμένες από το πρόβλημα της Υποθετικής Μεροληψίας, που σύμφωνα με την βιβλιογραφία, είναι βασικός παράγοντας που πλήττει την αξιοπιστία και την εγκυρότητα των αποτελεσμάτων της μεθόδου CV.

Η διδακτορική διατριβή αποτελείται από τέσσερα αλληλένδετα κεφάλαια. Στο πρώτο κεφάλαιο παρατίθεται η θεωρία που αναφέρεται στη μέθοδο CV, η βιβλιογραφική ανασκόπηση για την διάχυση της μεθόδου καθώς και οι κριτικές που έχει δεχτεί η μέθοδος συμπεριλαμβανομένου του προβλήματος της Υποθετικής Μεροληψίας. Επιπλέον η βιβλιογραφική ανασκόπηση του πρώτου κεφαλαίου αναφέρει ποικίλες έρευνες που αναζήτησαν την ύπαρξη του προβλήματος καθώς επίσης και διάφορες μεθόδους που έχουν προταθεί για να διορθωθεί η Υποθετική Μεροληψία. Επιπλέον στο τέλος του πρώτου κεφαλαίου γίνεται αναφορά του θεωρητικού υπόβαθρου του στατιστικού μοντέλου που κατασκευάστηκε στην παρούσα διατριβή.

Στο δεύτερο και στο τρίτο κεφάλαιο παρουσιάζεται η εφαρμογή του προτεινόμενου στατιστικού μοντέλου για να διορθωθεί το πρόβλημα της Υποθετικής Μεροληψίας για την Open-Ended μέθοδο εκμαίευσης δεδομένων και για την Double-Bounded Dichotomous Choice μέθοδο εκμαίευσης δεδομένων της CV μεθόδου αντίστοιχα. Πιο συγκεκριμένα, στα κεφάλαια αυτά παρουσιάζεται αρχικά η θεωρία της CV μεθόδου που αντιστοιχεί στην αντίστοιχη μέθοδο εκμαίευσης, αναλύεται λεπτομερώς η θεωρητική

βάση, η κατασκευή του στατιστικού μοντέλου, καθώς και άλλες θεωρίες που εφαρμόστηκαν κατά τον σχεδιασμό του μοντέλου, πχ. Θεωρία μικτών μοντέλων-Mixture models theory κτλ.. Στην συνέχεια των κεφαλαίων αυτών πραγματοποιήθηκαν προσομοιώσεις με σκοπό να εξετασθεί πως λειτουργεί το μοντέλο κι αν ανταποκρίνεται στον αρχικό σκοπό του. Επιπροσθέτως στο τρίτο κεφάλαιο παρουσιάζεται και μια εμπειρική εφαρμογή όπου το προτεινόμενο μοντέλο δοκιμάστηκε σε πραγματικά στοιχεία μελέτης με την εφαρμογή της μεθόδου CV.

Τέλος το τέταρτο κεφάλαιο πρόκειται για επέκταση του τρίτου κεφαλαίου καθώς ο ΕΜ αλγόριθμος, που εφαρμόζεται για την εκτίμηση του προτεινόμενου μοντέλου, αναφέρεται στην βιβλιογραφία ότι αντιμετωπίζει σοβαρή ευαισθησία στις αρχικές τιμές. Για τον λόγο αυτό, προκειμένου να προταθεί ένα ολοκληρωμένο μοντέλο για την αντιμετώπιση της Υποθετικής Μεροληψίας, έγινε μια σύγκριση τριών διαφορετικών μεθόδων προσδιορισμού αρχικών τιμών. Οι μέθοδοι που εξετάστηκαν ήταν η 1 τυχαία αρχικοποίηση, η οποία μέθοδος εφαρμόστηκε στα κεφάλαια 2 και 3, η μέθοδος των 100 τυχαίων αρχικοποιήσεων και τέλος ο αλγόριθμος k-means. Για την σύγκριση πραγματοποιήθηκαν προσομοιώσεις και εφαρμόστηκαν διάφορα κριτήρια προερχόμενα από την βιβλιογραφική ανασκόπηση.

Λέξεις Κλειδιά: Περιβαλλοντική Οικονομετρία, Πιθανολογική Αποτίμηση, Ανοιχτού-Τύπου μέθοδος εκμαίευσης, Κλειστού-Τύπου μέθοδος εκμαίευσης, Υποθετική Μεροληψία, Υπόδειγμα Στοχαστικού Συνόρου, Πεπερασμένα Μείγματα Κατανομών (Μικτά μοντέλα), Σύνθετο Σφάλμα, ΕΜ Αλγόριθμος, Προθυμία Πληρωμής, Αρχικές Τιμές

DEDICATED TO

MY PARENTS AND MY HUSBAND

AKNOWLEDGEMENTS

I would like to express my sincere gratitude to my Research Supervisor, Margarita Genius for her invaluable guidance, dedication and patience over the years. I am very grateful for the amount of things that she taught me through all these years and for the amount of time that she has taken to advise and support me. I could not have imagined having better advisor and I consider her my mentor.

I am very thankful to my other dissertation committee members: Vangelis Tzouvelekas and Georgios Tsiotas for their support and help over the years in shaping the direction of my research.

I also thank Michail Tsagris for his insightful comments regarding R language.

I acknowledge the financial support from the Special Account and Research of the University of Crete, project KA4722, during my studies.

I would like to express my gratitude to my parents Georgios Pavlakis and Georgia Pavlaki for their unlimited love and support. I thank them for always being by my side and believing in me. They taught me to have faith and never give up on my dreams.

I would also like to express my gratitude to my husband Konstantinos Boutzoukas for his unconditional love and constant encouragement the last 5 years. I thank him for being by my side and giving me strength to finish my research.

TABLE OF CONTENTS

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

1DC: Single-Bound Dichotomous Choice

2DC: Double-Bounded Dichotomous Choice

AIC: Akaike Information Criterion

ARI: Adjusted Rand Index

BIC: Bayesian Information Criterion

CV/CVM: Contingent Valuation/ Contingent Valuation Method

DC: Dichotomous Choice

EM: Expectation Maximization

HB: Hypothetical Bias

L/LL: Log-Likelihood

ML: Maximum Likelihood

MWTP: Mean Willingness To Pay

nn: No-No

ny: No-Yes

NOAA: National Oceanic and Atmospheric Administration

OE: Open-Ended

OOHB: One-and-One-Half-Bound

REBMIX: Rough Enhanced Bayes Mixture

SF: Stochastic Frontier

WTA: Willingness To Accept

WTP: Willingness To Pay

yn: Yes-No

yy: Yes-Yes

Εκτεταμένη Περίληψη

Η παρούσα διδακτορική διατριβή έχει ως αντικείμενο την επίλυση του προβλήματος Υποθετικής Μεροληψίας σε έρευνες με την μέθοδο Πιθανολογικής Αποτίμησης-Contingent Valuation (CV). Συγκεκριμένα μεταξύ τεσσάρων αλληλοεξαρτώμενων κεφαλαίων, προτείνεται ένας εναλλακτικός τρόπος μοντελοποίησης των δεδομένων προερχόμενες από CV έρευνες. Ο εναλλακτικός τρόπος μοντελοποίησης που προτείνεται αναφέρεται σε δύο εναλλακτικές μεθόδους εκμαίευσης των δεδομένων και σε διαφορετικούς τρόπους προσδιορισμού των αρχικών τιμών για την εκτίμηση του προτεινόμενου μοντέλου.

Κεφάλαιο 1: Η μέθοδος Πιθανολογικής Αποτίμησης (Contingent Valuation Method) και η Υποθετική Μεροληψία

Το πρώτο κεφάλαιο παρουσιάζει την μέθοδο Contingent Valuation και τις κριτικές που έχει δεχτεί αναφορικά με τις εκτιμήσεις της. Πιο αναλυτικά, καθώς τα περισσότερα περιβαλλοντικά αγαθά δεν έχουν κάποια τιμή στην αγορά οι οικονομολόγοι έχουν αναπτύξει μεθόδους για να τα αξιολογούν. Κάποιες μέθοδοι που χρησιμοποιούνται είναι οι αποκαλυπτόμενες προτιμήσεις όπου βασίζονται στις παρατηρήσεις της πραγματικής συμπεριφοράς και οι δεδηλωμένες προτιμήσεις που βασίζονται σε αυτά που ισχυρίζονται τα άτομα ότι θα έκαναν (Bockstael and McConnell, 2007:15). Η μέθοδος CV ανήκει στις μεθόδους δεδηλωμένων προτιμήσεων και θεωρείται ως από τα πιο ευέλικτα εργαλεία καθώς παρέχει την δυνατότητα να σχεδιαστούν ποικίλα σενάρια (Carson and Hanemann, 2005:824).

Ο ερευνητής που εφαρμόζει την μέθοδο CV χρησιμοποιεί ένα ερωτηματολόγιο που αποτελείται από τρία μέρη. Το πρώτο μέρος αποτελείται από γενικές ερωτήσεις σχετικές με το αντικείμενο της έρευνας, το δεύτερο μέρος παρουσιάζει το CV σενάριο και τις πληροφορίες αναφορικά με την υποτιθέμενη πληρωμή και τέλος το τρίτο μέρος αφορά ερωτήσεις για την συλλογή των δημογραφικών χαρακτηριστικών των ερωτώμενων (Carson and Hanemann, 2005:825).

Πιο αναλυτικά, στο δεύτερο μέρος του ερωτηματολογίου συγκεντρώνονται οι πληροφορίες για την προθυμία πληρωμής (Willingness To Pay-WTP) ή την προθυμία αποζημίωσης (Willingness To Accept-WTA) όπου υπάρχουν διαφορετικές μέθοδοι εκμαίευσης για την συγκέντρωση των απαντήσεων. Ονομαστικά κάποιες μέθοδοι εκμαίευσης είναι η Ανοιχτού Τύπου (Open-Ended), όπου η ερώτηση προθυμίας για πληρωμη/αποζημίωση είναι ανοιχτή, κι επιπλέον υπάρχουν κάποιες μέθοδοι Δημοψηφίσματος όπου δίνεται στους ερωτώμενους μια προσφορά (bid) και ο ερωτώμενος καλείται να απαντήσει με ένα Ναι ή Όχι για την προσφορά αυτή. Εάν είναι μόνο μία η ερώτηση η μέθοδος εκμαίευσης ονομάζεται Single-Bound Dichotomous-Choice, εάν στον ερωτώμενο δίνονται δύο διαδοχικές ερωτήσεις ονομάζεται Double-Bounded Dichotomous-Choice και τέλος εάν ακολουθεί και τρίτη διαδοχική ερώτηση η μέθοδος ονομάζεται Third-Bound Dichotomous-Choice (Arrow et.al., 1993:4).

Παρ’ όλη την ευρεία χρήση της μεθόδου CV, έχει δεχτεί κριτική για την αξιοπιστία των αποτελεσμάτων της. Μια από τις βασικότερες κριτικές της αναφέρεται στο πρόβλημα της Υποθετικής Μεροληψίας. Η Υποθετική Μεροληψία ορίζεται ως την διαφορά που υπάρχει ανάμεσα στο τι δηλώνουν τα άτομα ότι θα πλήρωναν και στο τι θα πλήρωναν πραγματικά (Loomis, 2014:35). Η Υποθετική Μεροληψία εμφανίζεται σε δύο μορφές, Υπερεκτίμηση ή Υποεκτίμηση του WTP ή του WTA. Για να ξεπεραστεί το πρόβλημα της Υποθετικής Μεροληψίας έχουν προταθεί ποικίλες μέθοδοι, οι οποίες χωρίζονται σε δύο βασικές κατηγορίες ανάλογα το σε ποιο σημείο της έρευνας εφαρμόζονται.

Πιο αναλυτικά, υπάρχουν μέθοδοι που εφαρμόζονται εκ των προτέρων (ex ante), δηλαδή κατά την συλλογή των δεδομένων και αποσκοπούν στην αποφυγή της Υποθετικής Μεροληψίας. Τέτοιες μέθοδοι είναι οι “cheap talk”, “solemn oath” και “scenario adjustments”(Haab et al., 2013:599), όπου επιδιώκουν μέσω ενημέρωσης για την ύπαρξη του προβλήματος ή μέσω όρκων να παροτρύνουν τους ερωτώμενους να απαντήσουν ειλικρινά. Από την άλλη υπάρχουν και μέθοδοι που εφαρμόζονται εκ των υστέρων (ex post), δηλαδή κατά την διάρκεια της εκτίμησης μέσω στατιστικών τεχνικών (Hofler and List, 2004:213).

Οι Hofler και List(2004) πρότειναν την χρήση του υποδείγματος στοχαστικού συνόρου (stochastic frontier analysis) για να συμπεριληφθεί η διαφορά ανάμεσα στις πραγματικές και στις υποθετικές προσφορές σε μια εφαρμογή που αφορούσε κάρτες του baseball. Στην συνέχεια προτάθηκε η χρήση του υποδείγματος στοχαστικού συνόρου για την μοντελοποίηση της υπερδήλωσης στην Double-Bounded μέθοδο (Chien et al.,2005). Επιπροσθέτως οι Kumbhakar, Parmeter και Tsionas (2013) πρότειναν το μοντέλο μηδενικής αναποτελεσματικότητας (zero inefficiency model) όπου σε ένα δείγμα μπορούν να υπάρχουν ταυτόχρονα, με μια πιθανότητα επιχειρήσεις που είναι αποτελεσματικές και επιχειρήσεις που είναι μη αποτελεσματικές.

Στην παρούσα διατριβή προτείνεται ένα μικτό μοντέλο υποδείγματος στοχαστικού συνόρου (mixture stochastic frontier model) ώστε να μοντελοποιηθεί η ύπαρξη Υποθετικής Μεροληψίας συνδυάζοντας τις προτάσεις των Chien, Huang και Shaw (2005) και των Kumbhakar, Parmeter και Tsionas (2013). Τέλος το μοντέλο το οποίο προτείνεται παρουσιάζεται για δύο μεθόδους εκμαίευσης, την Open-Ended και την Double-Bounded.

Κεφάλαιο 2: Εφαρμογή του υποδείγματος στοχαστικού συνόρου (stochastic frontier analysis) και μικτών (mixture) μοντέλων στην μέθοδο CV για την Ανοιχτού τύπου (Open-Ended) μέθοδο εκμαίευσης.

Στο δεύτερο κεφάλαιο της παρούσας διατριβής παρουσιάζεται το προτεινόμενο μοντέλο για την περίπτωση όπου η μέθοδος εκμαίευσης του WTP είναι η Open-Ended. Στην περίπτωση αυτή ο ερωτώμενος καλείται να απαντήσει σε μια ανοιχτή ερώτηση πόσο θα ήταν διατεθειμένος να πληρώσει για την αλλαγή που παρουσιάζει το υποθετικό σενάριο.

Η Open-Ended μέθοδος έχει κάποια πλεονεκτήματα και μειονεκτήματα. Αρχικά, ένα πλεονέκτημα είναι ότι το WTP εκμαιεύεται απευθείας (Loomis, 1990:79). Επιπλέον η μέθοδος αυτή παρέχει περισσότερες πληροφορίες αναφορικά με τις προτιμήσεις του ερωτώμενου κι επιπλέον προσαρμόζεται ευκολότερα όταν μια έρευνα πραγματοποιείται σε περισσότερα χώρες (Håkansson, 2008:186). Από την άλλη η μέθοδος παρουσιάζει και κάποια μειονεκτήματα, όπως ότι πρόκειται για μια δύσκολη διαδικασία καθώς τα άτομα δυσκολεύονται να προσδιορίσουν το ποσό που θα ήταν διατεθειμένοι να πληρώσουν με αποτέλεσμα είτε να μην απαντούν καθόλου είτε απαντούν υποτιμώντας το ποσό που θα ήταν διατεθειμένοι να πληρώσουν (Loomis, 1990:79).

Για την κατασκευή του μοντέλου χρησιμοποιήθηκε η θεωρία του υποδείγματος στοχαστικού συνόρου υπό την υπόθεση ότι η Υποθετική Μεροληψία εμφανίζεται σε μορφή Υπερεκτίμησης του WTP. Για να ενταχθεί στο απλό μοντέλο γραμμικής παλινδρόμησης η υπερεκτίμηση προστίθεται στον διαταρακτικό όρο, όπου κατανέμεται σύμφωνα με την κανονική κατανομή, ένα επιπλέον σφάλμα που ονομάζεται μονόπλευρος (one-sided) όρος σφάλματος καθώς κατανέμεται σύμφωνα με την ημι-κανονική κατανομή (Kumbhakar and Lovell, 2000:140).

Το γεγονός ότι μπορεί σε ένα δείγμα να υπάρχουν ερωτώμενοι που δεν απαντούν ειλικρινά, δεν σημαίνει ότι δεν υπάρχουν και κάποιοι που όντως δίνουν την ειλικρινή τους απάντηση. Για τον λόγο αυτό, χρησιμοποιώντας την θεωρία των μικτών μοντέλων και την ιδέα των Kumbhakar et al. (2013:67), γίνεται η υπόθεση ότι στο δείγμα υπάρχουν δύο ομάδες ταυτόχρονα, αυτοί που απαντούν ειλικρινά κι αυτοί που υπερδηλώνουν, με την πιθανότητα να ανήκει κάποιος στην ομάδα 1 να είναι η p_1 και η πιθανότητα να ανήκει κάποιος στην ομάδα 2 να είναι η $p_2 = 1 - p_1$.

Για την εκτίμηση του μικτού μοντέλου χρησιμοποιείται ο αλγόριθμος EM ο οποίος χειρίζεται το πρόβλημα μεγιστοποίησης ως πρόβλημα που λείπουν παρατηρήσεις (missing values), όπου οι παρατηρήσεις που λείπουν στην προκειμένη περίπτωση είναι οι πληροφορίες αναφορικά με το σε ποια ομάδα ανήκει ο κάθε ερωτώμενος. Ο αλγόριθμος EM αποτελείται από δύο βήματα, το βήμα E που αναφέρεται στην προσδοκία και το βήμα M που αναφέρεται στην μεγιστοποίηση (McLachlan and Peel, 2000:48).

Στην συνέχεια ακολούθησαν προσομοιώσεις 1000 επαναλήψεων και 1000 παρατηρήσεων για ένα εύρος περιπτώσεων ώστε να εξεταστεί η εγκυρότητα του μοντέλου. Επιπλέον αναφορικά με την πιθανότητα να ανήκει ο ερωτώμενος στην ομάδα 1(απαντούν ειλικρινά) και στην ομάδα 2 (υπερδηλώνουν) εξετάστηκαν δύο διαφορετικές περιπτώσεις. Στην μια περίπτωση θεωρήθηκε ότι η πιθανότητα να ανήκει κάποιος σε μία ομάδα είναι σταθερή για όλους όσους ανήκουν στην ομάδα αυτή, ενώ η δεύτερη περίπτωση που εξετάστηκε υποθέτει ότι η πιθανότητα να ανήκει κάποιος σε μια ομάδα είναι διαφορετική για κάθε ερωτώμενο καθώς η πιθανότητα εξαρτάται από μια μεταβλητή. Επιπλέον, κατά την διάρκεια της εκτίμησης δόθηκε ιδιαίτερη προσοχή στον προσδιορισμό των αρχικών τιμών και ακολουθήθηκαν κάποια βήματα για τον προσδιορισμό τους.

Τα αποτελέσματα των προσομοιώσεων και για τις δύο περιπτώσεις προσδιορισμού της πιθανότητας έδειξαν ότι το προτεινόμενο μοντέλο είναι σε θέση να διορθώσει το πρόβλημα της Υποθετικής Μεροληψίας καθώς οι εκτιμητές των παραμέτρων είναι πολύ κοντά στις αληθινές τιμές κι επιπροσθέτως η Μεροληψία είναι πολύ κοντά στο μηδέν. Επιπλέον, για λόγους σύγκρισης, τα δεδομένα εκτιμήθηκαν και με το απλό μοντέλο εκτίμησης, που δεν λαμβάνει υπόψη την ύπαρξη Υποθετικής Μεροληψίας και από την σύγκριση των δύο μοντέλων προέκυψε ότι το απλό μοντέλο δεν εκτιμά το ίδιο καλά τις παραμέτρους και η Μεροληψία που προκύπτει είναι αρκετά υψηλή.

Σε κάποιες περιπτώσεις όπου η αληθινή τιμή της παραμέτρου που αναφέρεται στον βαθμό υπερδήλωσης είχε μικρότερες τιμές, παρατηρήθηκε ότι το πρόγραμμα δυσκολεύτηκε να διακρίνει τις δύο ομάδες. Όσο μεγαλύτερος ήταν ο βαθμός υπερδήλωσης τόσο ευκολότερα το πρόγραμμα μπορούσε να διακρίνει τις δύο ομάδες. Επιπλέον στις περιπτώσεις που δυσκολευόταν το πρόγραμμα να διακρίνει τις δύο ομάδες είχε ως αποτέλεσμα να παρουσιάζονται περισσότερες επαναλήψεις με προβλήματα στο προσδιορισμό των τυπικών αποκλίσεων κάποιων παραμέτρων.

Εν συντομία, τα αποτελέσματα από τις προσομοιώσεις για το Open-Ended μοντέλο οδηγούν στο συμπέρασμα ότι το μοντέλο είναι κατάλληλο για να διορθώσει το πρόβλημα της Υποθετικής Μεροληψίας αρκεί το μοντέλο να μπορεί να διακρίνει τις δύο ομάδες. Στην συνέχεια της διατριβής παρουσιάζεται το μοντέλο που προτείνεται για την Double-Bounded μέθοδο εκμαίευσης των δεδομένων.

Κεφάλαιο 3: Εφαρμογή του υποδείγματος στοχαστικού συνόρου (stochastic frontier analysis) και μικτών (mixture) μοντέλων στην μέθοδο CV για την Double-Bounded μέθοδο εκμαίευσης.

Στο τρίτο κεφάλαιο της παρούσας διατριβής παρουσιάζεται το προτεινόμενο μοντέλο για την περίπτωση όπου η μέθοδος εκμαίευσης του WTP είναι με την μέθοδο του δημοψηφίσματος και πιο συγκεκριμένα η Double-Bounded. Στην περίπτωση αυτή ο ερωτώμενος καλείται να απαντήσει με ένα Ναι ή Όχι στην ερώτηση για το αν θα ήταν διατεθειμένος να πληρώσει για την αλλαγή που παρουσιάζει το υποθετικό σενάριο μια συγκεκριμένη προσφορά. Στην συνέχεια, εάν έχει απαντήσει με Ναι στην πρώτη προσφορά του παρουσιάζεται μια δεύτερη προσφορά, μεγαλύτερη από την πρώτη και καλείται ξανά να απαντήσει με Ναι ή Όχι. Στην περίπτωση που έχει απαντήσει Όχι στην πρώτη προσφορά τότε η δεύτερη προσφορά που δίνεται στον ερωτώμενο είναι μικρότερη της πρώτης προσφοράς και καλείται ξανά να απαντήσει με Ναι ή Όχι.

Η μέθοδος Double-Bounded προτείνεται ως η μέθοδος που θα πρέπει να χρησιμοποιούν οι ερευνητές καθώς η χρήση μεθόδου δημοψηφίσματος έχει πολλά πλεονεκτήματα (Arrow et al., 1993: 21). Αρχικά είναι πιο ρεαλιστική δεδομένου ότι η παροχή δημόσιων αγαθών είναι συνήθως με μεθόδους δημοψηφίσματος συνεπώς είναι κάτι πιο οικείο στους ερωτώμενους κι επιπροσθέτως δεν έχουν στρατηγικό λόγο να απαντήσουν μη ειλικρινά. Καθώς δεν απαιτεί πολύ σκέψη και προσπάθεια για να απαντηθούν οι ερωτήσεις συνεπώς είναι λιγότεροι εκείνοι που δεν απαντούν. Επιπλέον η μέθοδος αυτή μοιάζει με την λειτουργία της αγοράς όπου τα άτομα βλέπουν τις τιμές και αποφασίζουν αν θα αγοράσουν ή όχι. Από την άλλη, η μέθοδος αυτή παρουσιάζει και κάποια μειονεκτήματα. Ένα μειονέκτημα είναι ότι οι εκτιμήσεις για την προθυμία πληρωμής μπορεί να επηρεάζονται από υποθέσεις για την συνάρτηση χρησιμότητας ή για την κατανομή των σφαλμάτων (Loomis, 1990:79). Επιπλέον ένα βασικό μειονέκτημα της μεθόδου είναι ότι οι ερωτώμενοι επηρεάζονται από την πρώτη προσφορά με αποτέλεσμα να αποδέχονται και την δεύτερη προσφορά.

Από τις απαντήσεις που δίνει ο ερωτώμενος στις προσφορές προκύπτει ένα συμπέρασμα αναφορικά με το που ανήκει το η προθυμία πληρωμής-WTP του ερωτώμενου. Πιο αναλυτικά εάν ο ερωτώμενος έχει πει Ναι και στις δύο προσφορές τότε το WTP του ατόμου είναι μεγαλύτερο από την υψηλότερη προσφορά. Το αντίθετο ισχύει εάν το άτομο έχει πει Όχι και στα δυο προτεινόμενα ποσά καθώς συμπεραίνεται ότι το WTP του ερωτώμενου είναι μικρότερο από την χαμηλότερη προσφορά. Στην περίπτωση όπου οι απαντήσεις του ατόμου είναι Ναι στην πρώτη ερώτηση και Όχι στην δεύτερη ή Όχι στην πρώτη και Ναι στην δεύτερη, προκύπτει ένα διάστημα, ότι δηλαδή το WTP του ατόμου είναι μεταξύ των δύο προσφορών.

Για την κατασκευή του μοντέλου, όπως έχει ήδη αναφερθεί, χρησιμοποιήθηκε η θεωρία του υποδείγματος στοχαστικού συνόρου για την περίπτωση όπου η Υποθετική Μεροληψία εμφανίζεται σε μορφή Υπερεκτίμησης του WTP. Για να ενταχθεί στο μοντέλο λανθάνουσας μεταβλητής η υπερεκτίμηση ακολουθήθηκε η ίδια μεθοδολογία με αυτή του κεφαλαίου 2. Επιπλέον όπως και στο κεφάλαιο 2, μελετήθηκαν δύο διαφορετικές περιπτώσεις αναφορικά με την πιθανότητα να ανήκει κάποιος σε μια ομάδα. Στην μια περίπτωση θεωρήθηκε ότι η πιθανότητα να ανήκει κάποιος σε μία ομάδα είναι σταθερή για όλους όσους ανήκουν στην ομάδα αυτή, ενώ στη δεύτερη περίπτωση η πιθανότητα να ανήκει κάποιος σε μια ομάδα είναι διαφορετική για κάθε ερωτώμενο καθώς η πιθανότητα εξαρτάται από μια μεταβλητή.

Στην συνέχεια πραγματοποιήθηκαν προσομοιώσεις 1000 επαναλήψεων για 1000 παρατηρήσεις για την δημιουργία δεδομένων με την ύπαρξη υπερδήλωσης ώστε να εξασφαλιστεί η ύπαρξη Υποθετικής Μεροληψίας κι έπειτα ακολούθησε εκτίμηση με το προτεινόμενο μοντέλο ώστε να εξεταστεί η εγκυρότητα και η ικανότητα του μοντέλου για την διόρθωση του προβλήματος. Τα αποτελέσματα των προσομοιώσεων και για τις δύο περιπτώσεις προσδιορισμού της πιθανότητας έδειξαν ότι το προτεινόμενο μοντέλο διορθώνει σε ικανοποιητικό βαθμό το πρόβλημα της Υποθετικής Μεροληψίας. Πιο συγκεκριμένα οι εκτιμήσεις των παραμέτρων είναι πολύ κοντά στις αληθινές τιμές για όλες τις περιπτώσεις που εξετάστηκαν και η Μεροληψία που υπολογίστηκε είναι πολύ κοντά στο μηδέν. Επιπλέον καθώς τα δεδομένα εκτιμήθηκαν και με το απλό μοντέλο εκτίμησης που δεν λαμβάνει υπόψη την ύπαρξη Υποθετικής Μεροληψίας, από την σύγκριση των αποτελεσμάτων προέκυψε ότι για όλες τις περιπτώσεις εάν δεν ληφθεί υπόψη κατά την εκτίμηση η ύπαρξη της υπερδήλωσης από τους ερωτώμενους, το απλό μοντέλο επιστρέφει εκτιμήσεις που δεν είναι τόσο κοντά στις αληθινές τιμές και η Μεροληψία είναι πολύ μεγαλύτερη.

Επιπλέον στο κεφάλαιο 3 αφότου ολοκληρώθηκαν οι προσομοιώσεις, πραγματοποιήθηκε μια εμπειρική εφαρμογή του προτεινόμενου μοντέλου σε αληθινά δεδομένα από έρευνα με την μέθοδο CV. Πιο αναλυτικά, για την υλοποίηση της εμπειρικής εφαρμογής το προτεινόμενο μοντέλο τροποποιήθηκε, καθώς τα δεδομένα δεν παρουσίαζαν υπερδήλωση αλλά υποδήλωση του WTP. Συνεπώς τα δεδομένα εκτιμήθηκαν για την περίπτωση όπου η Υποθετική Μεροληψία υπάρχει υπό την μορφή υποεκτίμησης του WTP και για την περίπτωση όπου η πιθανότητα υποδήλωσης δεν είναι ίδια αλλά διαφέρει ανάμεσα στους ερωτώμενους καθώς εξαρτάται από κάποιες μεταβλητές.

Το δείγμα αναφέρεται σε 1827 παρατηρήσεις και τα δεδομένα είναι διαθέσιμα στο πακέτο “Ecdat” της γλώσσας προγραμματισμού R στο όνομα “kakadu” τα οποία δεδομένα προέρχονται από το άρθρο της Werner (1999) (Croissant and Graves, 2020:84-85). Τα αυθεντικά στοιχεία προέρχονται από ένα άρθρο των Carson, Wilks και Imber του 1994 όπου πραγματοποίησαν μια έρευνα το 1990 στην Αυστραλία με την μέθοδο CV για το εάν θα προχωρήσει η ίδρυση ορυχείων στην διατηρητέα περιοχή του Kakadu ή αν θα πρέπει να ενταχθεί κι η περιοχή αυτή στο εθνικό πάρκο του Kakadu.

Οι ανεξάρτητες μεταβλητές που χρησιμοποιήθηκαν για την ερμηνεία του WTP καθώς και οι μεταβλητές που επηρεάζουν την πιθανότητα είναι εκείνες που χρησιμοποίησε και βρήκε ότι είναι στατιστικά σημαντικές στην έρευνα της η Werner. Από την εκτίμηση του μοντέλου με το προτεινόμενο μοντέλο εντοπίστηκε ότι όντως ήταν υποεκτιμημένο το WTP. Επιπλέον εφόσον εκτιμήθηκε το απλό μοντέλο και το προτεινόμενο μικτό μοντέλο συγκρίθηκαν τα δύο μοντέλα εφαρμόζοντας τα κριτήρια BIC και AIC όπου έδειξαν ότι το προτεινόμενο μικτό μοντέλο είναι καταλληλότερο καθώς οι μεταβλητές που προστέθηκαν στο μικτό μοντέλο περιγράφουν με καλύτερο τρόπο την προθυμία πληρωμής.

Συνοψίζοντας από τα αποτελέσματα των προσομοιώσεων και της εμπειρικής εφαρμογής συμπεραίνεται ότι προτεινόμενο μοντέλο μπορεί να διορθώσει το πρόβλημα της Υποθετικής Μεροληψίας είτε είναι υπό την μορφή υπερεκτίμησης του WTP είτε υπό την μορφή υποεκτίμησης του WTP. Τέλος το τέταρτο κεφάλαιο της παρούσας διατριβής εμβαθύνει στο θέμα του προσδιορισμού των αρχικών τιμών και πιο συγκεκριμένα συγκρίνει 3 διαφορετικές τεχνικές ώστε να προσδιοριστεί ποιος τρόπος προσδιορισμού των αρχικών τιμών λειτουργεί καλύτερα για την διόρθωση του προβλήματος της Υποθετικής Μεροληψίας.

Κεφάλαιο 4: Σύγκριση διαφορετικών μεθόδων αρχικών τιμών για την Double-Bounded μέθοδο εκμαίευσης του μικτού (mixture) μοντέλου.

Στην παρούσα διατριβή το προτεινόμενο μικτό μοντέλο εκτιμάται με αλγόριθμο EM, ο οποίος αλγόριθμος παρουσιάζει κάποια μειονεκτήματα. Το πρώτο μειονέκτημα είναι ότι απαιτεί καλές αρχικές τιμές και το δεύτερο μειονέκτημα είναι ότι υπάρχει περίπτωση ο αλγόριθμος να παγιδευτεί σε τοπικά μέγιστα (Panić et al., (2020:1). Καθώς λοιπόν, ο προσδιορισμός αρχικών τιμών είναι πολύ σημαντικός για τον αλγόριθμο EM, στο κεφάλαιο αυτό εξετάζεται το θέμα αυτό λεπτομερώς και συγκεκριμένα δοκιμάζονται και συγκρίνονται διαφορετικοί τρόποι αρχικοποίησης αναφορικά με το πως χωρίζονται οι παρατηρήσεις σε δύο ομάδες. Στην βιβλιογραφία έχουν προταθεί πολλές τεχνικές ώστε να επιλέγονται οι αρχικές τιμές, πολλές εκ των οποίων βασίζονται στην ομαδοποίηση/κατηγοριοποίηση (clustering). Η κατηγοριοποίηση χρησιμοποιεί πολλές τεχνικές ώστε να χωριστούν τα δεδομένα σε υποομάδες με γνώμονα να μοιάζουν όσο το δυνατόν περισσότερο όσοι ανήκουν στην ίδια ομάδα (Mann and Kaur, 2013:43-44).

Για την παρούσα έρευνα υλοποιήθηκαν τρεις μέθοδοι αρχικοποίησης. Η πρώτη μέθοδος αναφέρεται σε τυχαία αρχικοποίηση που πραγματοποιήθηκε 1 φορά (1 random initialization), η οποία είναι η μέθοδος που εφαρμόστηκε στα κεφάλαια 2 και 3. Στην συνέχεια εφαρμόστηκε μια επέκταση της προηγούμενης μεθόδου καθώς πραγματοποιήθηκε η τυχαία αρχικοποίηση 100 φορές (100 random initializations) και στη συνέχεια επιλέχθηκαν ως αρχικές τιμές οι εκτιμήσεις όπου ανάμεσα στις 100 επαναλήψεις παρουσίασαν την μεγαλύτερη log-likelihood. Τέλος η τρίτη μέθοδος αρχικοποίησης που εφαρμόστηκε είναι ο αλγόριθμος κατηγοριοποίησης k-means.

Λίγα λόγια για την κάθε μέθοδο, η 1 τυχαία αρχικοποίηση είναι από τις πιο ευρέως χρησιμοποιούμενες μεθόδους για να προσδιοριστούν αρχικές τιμές για τον αλγόριθμο EM (Biernachi et al., 2003:566). Με την μέθοδο αυτή τα δεδομένα χωρίστηκαν σε δύο ομάδες με αναλογία 50:50 μέσω τυχαίων λήψεων που γίνονται από μια ομοιόμορφη κατανομή. Οι τυχαίες λήψεις από μία ομοιόμορφη κατανομή είναι μια γνωστή διαδικασία για τον προσδιορισμό αρχικών τιμών (Hipp and Bauer, 2006:41, Shireman et al., 2017:284).

Η τυχαία αρχικοποίηση που γίνεται περισσότερες φορές, στην περίπτωση της παρούσας έρευνας 100, είναι μια επέκταση της προηγούμενης μεθόδου που ουσιαστικά επαναλαμβάνεται η διαδικασία περισσότερες από μια φορές. Η μέθοδος αυτή είναι γνωστή ως “search/run/select” (Biernachi et al., 2003:563-66). Συνήθως για την επιλογή των αρχικών τιμών ανάμεσα στις επαναλήψεις γίνεται μέσω κριτηρίων όπως το BIC (Shireman et al., 2017:284). Η μέθοδος αυτή συνήθως συνδυάζεται με μια μέθοδο που ονομάζεται σύντομες εκτιμήσεις EM (short runs of EM) όπου δεν επιτρέπεται στον αλγόριθμο να φτάσει σε σύγκλιση αλλά ανακόπτεται η διαδικασία μεγιστοποίησης (Biernachi et al., 2003:567).

Ο αριθμός των τυχαίων αρχικοποιήσεων που προτείνεται από την βιβλιογραφία είναι γύρω στις 1000, κάτι το οποίο είναι πρακτικά αδύνατο, για τον λόγο αυτό συνηθίζεται να υλοποιούνται 100 (Shireman et al., 2017:284). Επιπλέον καθώς ανακόπτεται ο αλγόριθμος προτού φτάσει σε σύγκλιση, γίνεται λόγος στις πόσες επαναλήψεις θα ήταν ιδανικό να σταματήσει. Στην βιβλιογραφία δεν υπάρχει ένας γενικός κανόνας, συνήθως αναφέρονται σε 10 και 20 επαναλήψεις. Στην παρούσα έρευνα επιλέχθηκαν 20 επαναλήψεις όπου είναι ο αριθμός επαναλήψεων που επιλέγει αυτόματα το πρόγραμμα Stata (StataCorp, 2021:4).

Τέλος η τρίτη μέθοδος αρχικοποίησης είναι ο αλγόριθμος k-means όπου είναι ένας πολύ γνωστός αλγόριθμος για διαχωρισμό δεδομένων σε ομάδες, όπου ο αριθμός των ομάδων προσδιορίζεται από τον ερευνητή εκ των προτέρων. Η βασική ιδέα της μεθόδου αυτής είναι ότι οι διαφορές μεταξύ των στοιχείων που ανήκουν στην ίδια ομάδα έχουν ελαχιστοποιηθεί (Kassambara, 2017:36-37). Βασικό πλεονέκτημα αυτής της μεθόδου είναι ότι είναι μια απλή και γρήγορη μέθοδος όμως το βασικό της μειονέκτημα είναι ότι πρέπει να είναι εκ των προτέρων γνωστός ο αριθμός των ομάδων (Kassambara, 2017:46).

Για την σύγκριση αυτών των 3 εργαλείων κατηγοριοποίησης πραγματοποιήθηκαν προσομοιώσεις ώστε να εφαρμοστούν κάποια κριτήρια σύγκρισης. Οι προσομοιώσεις έγιναν με 500 επαναλήψεις και 1000 παρατηρήσεις, για τις ίδιες περιπτώσεις για το Double-Bounded μικτό μοντέλο που προτείνεται, για την περίπτωση όπου η πιθανότητα είναι σταθερή για όσους ανήκουν στην ίδια ομάδα.

Αρχικά μια πρώτη σύγκριση πραγματοποιήθηκε αφότου ολοκληρώθηκαν οι εκτιμήσεις για όλες τις περιπτώσεις και μεθόδους αναφορικά με τους εκτιμητές και την Μεροληψία. Στην συνέχεια εφαρμόστηκαν κάποια κριτήρια απόδοσης που προτείνονται από την βιβλιογραφία για σύγκριση μεθόδων αρχικοποίησης. Πιο συγκεκριμένα, ένα κριτήριο είναι ο χρόνος που απαιτείται για να ολοκληρώσει το πρόγραμμα τις επεξεργασίες (Meilă and Heckerman, 2001:16). Ένα δεύτερο κριτήριο είναι ο συνολικός αριθμός επαναλήψεων που απαιτούνται για να φτάσει ο αλγόριθμος σε σύγκλιση και επίσης η ικανότητα να βρει το ολικό μέγιστο, όπου για το κριτήριο αυτό πρέπει οι εκτιμητές να ικανοποιούν κάποιες συνθήκες (Karlis and Xekalaki, 2003:580-81).

Ακόμη ένα κριτήριο όπου εφαρμόστηκε είναι ο δείκτης ARI (Adjusted Rand Index) όπου μετράει πόσο καλά χωρίστηκαν τα στοιχεία από την μέθοδο κατηγοριοποίησης (Maruotti and Punzo, 2021:455-56) και επιπλέον προστέθηκαν στην σύγκριση ο αριθμός των επαναλήψεων που αφαιρέθηκαν λόγω προβλημάτων στον υπολογισμό της τυπικής απόκλισης κάποιων παραμέτρων.

Εφαρμόζοντας τα παραπάνω κριτήρια προέκυψαν κάποια ενδιαφέροντα συμπεράσματα καθώς όλες οι μέθοδοι αρχικοποίησης φέρνουν εις πέρας με επιτυχία τον αρχικό σκοπό της διόρθωσης του προβλήματος της Υποθετικής Μεροληψίας. Λαμβάνοντας όμως υπόψη τα κριτήρια που εφαρμόστηκαν, η σύγκριση έδειξε ότι η μέθοδος με τις 100 τυχαίες αρχικοποιήσεις λειτουργεί καλύτερα σε σχέση με τις άλλες δύο μεθόδους όμως ένα βασικό μειονέκτημα αυτής της μεθόδου είναι ότι απαιτεί πολύ χρόνο.

Από την παρούσα διδακτορική έρευνα, προτείνεται για την διόρθωση του προβλήματος Υποθετικής Μεροληψίας η εφαρμογή ενός μικτού (mixture) μοντέλου όπου λαμβάνεται υπόψη κατά την εκτίμηση η πιθανότητα υπερδήλωσης του WTP. Για τις αρχικές τιμές του μοντέλου, εάν δεν υπάρχει πίεση χρόνου, να εφαρμόζονται οι 100 τυχαίες αρχικοποιήσεις, ειδάλλως οι άλλες δύο μέθοδοι αρχικοποίησης προτείνονται εξίσου καθώς η διαφορά στην απόδοση των τριών μεθόδων είναι πολύ μικρή.

Chapter 1
Contingent Valuation and Hypothetical Bias

Introduction

Most environmental goods are not traded in markets and as a result economists have developed methods for their valuation. Methods such as revealed preference which is based on observations of actual behavior and stated preference which is based on hypothetical behavior (Bockstael and McConnell, 2007:15), have been applied to value environmental goods.

One of the stated preference approaches is the Contingent Valuation Method (CVM) (Carson and Hanemann, 2005:824). The CVM is considered as a flexible tool because it provides the possibility of creating several experimental scenarios (Carson and Hanemann, 2005:824).

As mentioned in Du Preez, Menzies, Sale and Hosking (2012:3), over time CVM became the most widely used method for valuing goods that are not available in a market. The CVM is a tool that is usually used to value potential effects of policy changes in the case where market-base valuation of the effect is impossible. “The results of these analyses are often intended to inform policy decisions, which are made within the context of formal policymaking institutions” (Calson et al., 2016:460).

Although CVM is a broadly accepted method, there are several problems that the researchers have to deal with and the reliability of the method has been questioned since a heated debate has been triggered through the years. As the application of CVM spread to deal with the valuation of a number of goods many problems appeared.

Namely the most popular problems are problems such as non-response bias (Berg, 2005:865), starting point bias (Boyle et al., 1985:189), information bias (Ajzen et al.,1996:44), psychological biases (Bateman et al., 1995:166), question order bias (Kartman et al., 1996:532), and hypothetical bias (Loomis, 2014:35).

In this thesis the focus is on overcoming one of the critiques namely the existence of Hypothetical Bias. “Hypothetical bias can be defined as the difference between what a person indicates they would pay in the survey or interview and what a person would actually pay” (Loomis, 2014:35).

Hypothetical bias is a major issue of the CVM and as Haab, Interis, Petrolia and Whitehead mention (2013:595) “In short, we find promise for the curious researcher that the CVM debate is not settled, important questions remain, and that a critical examination of the CVM literature will provide fertile ground for future research”.

Many techniques have been devised to overcome hypothetical bias. The approaches that have been used are separated in two main categories. The first category is ex ante and the second one is ex post approaches (Loomis, 2014:34). This thesis aims to deal with the problem of hypothetical bias by using an ex post approach. More specifically, by using stochastic frontier analysis as in the case of Hofler and List (2004) where they applied the approach for Open-Ended auctions.

Additionally in this thesis the method is going to be applied to the Double-Bounded Dichotomous Choice (DC) model while introducing also elements from latent class models theory since we will be dealing with mixture models of two classes (Normal error and composed Normal-Half-normal error).

This thesis is inspired by the work of Chien, Huang and Shaw (2005) related to the estimation of a model when yea-saying bias is present and the work of Kumbhakar, Parmeter and Tsionas (2013) that propose a model that can be applied when both efficient and inefficient firms are present in a sample with a given probability. Consequently in this thesis mixtures will be added, since we have both overstating respondents and respondents that answer sincerely, in order to combine the two basic ideas that inspired this thesis.

Furthermore, we use the results of Tsay et al. (2013) who provide a closed-form approximation for the cumulative distribution function of a composed Normal-Half-normal error. Finally, in the present thesis we use the EM algorithm estimation procedure (McLachlan and Peel, 2000).

1.1. The Contingent Valuation Method (CVM)

1.1.1. The method

Neoclassical theory was faced with the problem that most environmental goods don't have a price, since they are not generally purchased in markets like other common goods. In order to overcome this problem several valuation methods have been proposed in the literature, one such method being the CVM (Vatn, 2005:302).

The CVM is one of the most commonly used methods to value non-marketed resources, such as wildlife, recreation and environmental quality (Hanemann et.al., 1991:1255). A basic principle in a survey using CVM is that people are rational and additionally, people know their preferences for a particular good relative to other goods and they can translate their preferences into money (Green and Tunstall, 1999:242-43).

In an investigation using the CVM approach, the researcher uses a questionnaire that presents a hypothetical environmental change to respondents. In the case of an environmental improvement the respondent is asked to state the maximum amount of income he/she would be willing to pay (WTP) for the improvement or alternatively the minimum amount of income he/she would be willing to accept (WTA) to forego the improvement. The aims and the needs of each research are the determinants of the design of the questionnaire. Typically the questionnaire consists of three parts, the beginning, the middle and the end (Green and Tunstall, 1999:238).

The first part is an introduction about the general subject of the research followed by general questions. The second part concerns the CVM scenario, the purpose of the scenario, how it will be implemented, how it will be financed and afterwards the WTP or WTA questions. Finally in the third part personal questions are made so that the researcher gathers the demographic characteristics of the respondents (Carson and Hanemann, 2005:825).

More analytically, the first part, which contains the introductory questions, aims to help the respondents understand the purpose of the research. As Green and Tunstall (1999:239) mention, the language that the researchers use must be clear and furthermore the researchers must give attention to every detail in order to ensure that the respondents will understand the questions and answer correctly. Furthermore, the description of the good that is evaluated is very crucial because a wrong description may mislead the respondents' answers with a significant impact on the validity of the research. Boyle and Bergstrom (1999:193) mention that several studies have proven that more information or less information in the description of the good may have statistically significant effects in surveys using CVM.

The second part of the questionnaire, which refers to the valuation scenario, contains the questions for eliciting WTP or WTA. This part consists of two components, the first one is an introduction which informs the respondents that they can freely express their opinion about the scenario and emphasizes the importance of the respondents' participation (Green and Tunstall, 1999:245-46). Furthermore, in this part the researcher should inform the respondents that they should have in mind their disposable income. Arrow et al. (1993:9-14) mentioned that one of the problems that concerned the National Oceanic and Atmospheric Administration-NOOA Panel was that the majority of previous applications of CVM, respondents were not reminded to have in mind their budget constrain while answering, so respondents may answer without thinking carefully. The other component refers to the WTP/WTA questions. Finally, the third part of the questionnaire is related to the collection of the socioeconomic information of the respondents.

The WTP or WTA questions are the most essential part of the questionnaire. There are several elicitation methods to design in a different way the CVM questions. Mitchell and Carson (1989:98) present 9 CV elicitation methods that could be categorized in two groups, whether respondents are given a single WTP question or an iterated series of WTP questions. Namely the single question methods are, the Open-Ended (OE) or the Direct question, the Payment Card, the Sealed bid auction, the Take-it-or-leave-it offer, the spending question offer and the Interval checklist. On the other hand, the iterated series questions methods are the Bidding game, the Oral auction and the Take-it-or-leave-it offer with follow-up.

For the first CV surveys the Bidding game format was used and the respondents were called to answer with a yes or no if they are willing to pay a certain amount. If they answered yes then a bigger amount was given, alternatively a smaller if they answered no. This process continued until the respondent switched answer form yes to no (or from no to yes) (Carson and Haneman, 2005:870).

Due to concerns about the starting bid of the Bidding game method, led researchers to apply simpler approaches such as Single closed-ended questions, known as the referendum, Single-bound (1DC) or binary discrete-response format and the Double-bounded (2DC) approach where a double sampling framework, with a second binary discrete choice question depending on the answer at the first was applied (Carson and Haneman, 2005:871). Additionally, in 2002 Cooper, Hanemann and Signorello, proposed another elicitation approach called the one-and-one-half-bound (OOHB).

Besides the fact that there is a variety of elicitation methods a researcher can choose, the NOAA Panel (Arrow et al., 1993:4) suggests that the elicitation method researchers should use is a referendum method such as 1DC and 2DC. Furthermore, one other main guideline that has been given by the NOAA Panel refers to Single-bound DC and suggests that the sample size might have to be at least 1000 respondents and generally that WTP format is preferred to WTA (Bateman et al., 1995:162).

A few words for some of the elicitation formats, firstly, the Open-Ended method is "a form of an open-ended question asking what is the maximum amount they would be willing to pay for the program in question" (Arrow et.al., 1993:4). On the other hand, the Single-Bound Dichotomous Choice, the Double-Bound Dichotomous Choice and the Third-Bound Dichotomous Choice are in a form of a hypothetical referendum in which each respondent has to answer if he is willing to pay a certain amount of money with a "Yes" or "No"(Arrow et al., 1993:4).

More specifically, the Single-Bound Dichotomous Choice method includes only one question-bid in which the respondents may answer with yes or no if they agree or disagree about paying a given amount of money in order to ensure an environmental improvement. To continue with, the Double-Bounded Dichotomous Choice method includes an extra follow-up question depending on the answer given to the first question. In the case of a WTP question, if the respondent has responded to the first question with a "yes" then, for the second question, the second bid will be bigger than the first bid. On the other hand, if the answer to the first bid is "no" then the second bid will be a smaller amount (Hanemann et.al., 1991:1255-56).

In Iterative Bidding method, after the DC questions an open-ended questions follows which gives to respondents the freedom to move up or down from the given WTP starting point (Bateman et al., 1995:161-1640). Finally, in the one-and-one-half-bound (OOHB) approach, two prices are given up front to the respondents and the researcher informs them that although the exact cost of the good is unknown, it lies within the two prices. Afterwards one of the two prices is randomly selected and the respondents are asked if they are willing to pay the given price, whether a follow-up question will ensue depends on the selected initial price and the answer to the first question, since the WTP amounts must be consistent to the stated price range (Cooper et al., 2002:742).

In this thesis the main elicitation methods that will be examined are the Double-Bounded Dichotomous Choice method as well as the open-ended elicitation method.

1.1.2. The diffusion of the method

The method was proposed in 1963 by Robert Davis in his Ph.D. thesis at Harvard University. The use of the method started in the beginning of the '70s and expanded after the '80s in the US (Loomis, 1999:613), in Europe as well and from the '90s all over the world (Bonnieux and Rainelli, 1999:585-86). The initial versions of CVM proposed by Davis in 1963 and Randall et al. in 1974 focused on incentives and free-rider issues (Green et al., 1998:86).

More analytically, Davis applied an Open-Ended protocol and Randall on the other hand applied a sequential bidding protocol. Randall et al. presented a number of arguments for the use of sequential bidding instead of Open-Ended. More specifically they mentioned that the referendum task was simpler and less affected by misinterpretations. Although there were many arguments in favor of sequential bidding protocol, the Open-ended protocol was the most commonly used protocols in the early '80s (Green et al., 1998:87).

The CVM has been widely used for valuation of environmental changes in many countries in the past years. The reasons for the broad acceptance of CVM are several. Firstly, from the mid-seventies until the decade of 1990 there was an increasing use of the method which led to a substantial number of publications, for example "the article by Randall et al., in 1974 in the first issue of the Journal of Environmental Economics and Management, which introduced more differentiation of contingent valuation surveys from opinion polls by using photographs to help describe the valuation scenario" (Haab et al., 2013:594).

Second, another important reason for the acceptance of CVM was the publication of the book by Mitchell and Carson in 1989 which "first integrated economic theory, survey research methods, and social science measurement issues" (Haab et al., 2013:594). Last but not least, the Exon Valdez oil spill, which turned the awareness of numerous economists and many institutions as well towards CVM (Haab et al., 2013:594).

In 1993 the US National Oceanic and Atmospheric Administration (NOAA) created a "blue ribbon panel" that gave guidelines for researches using CVM and afterwards numerous critiques came along from researchers such as Hausman (Haab et al., 2013:594).

1.2. Hypothetical Bias

1.2.1. The beginning of the debate over CVM

CVM is a broadly accepted method, although there are several problems that the researchers have to deal with. Sugden (1999:139) mentions that even if the respondents answer the CV questions honestly, major problems can arise. For example, the design of the scenario, especially for a public good, is not that easy.

The respondents on the other hand, as Green and Tunstall (1999:209-11) mention, are also facing difficulties in their attempt to give the right answers. Namely the respondents deal with memory problems, communication problems which lead to difficulties in understanding, lack of knowledge regarding the good that is evaluated in the valuation exercise and finally, as Milon (1989:294) mentions, respondents have to deal with their strategic incentives.

Since 1993 a debate has started questioning the validity of the CVM. As a consequence of the Exxon Valdez oil spill a conference took place and brought the attention towards CVM (Haab et al., 2013:594). The conclusion of the conference was that CVM was unreliable (Carson, 2012:30). In 1993 the National Oceanic and Atmospheric Administration (NOAA) assembled a panel in order to evaluate the CVM. The NOAA panel gave a number of guidelines for the method and these guidelines triggered the debate. In 1994, many articles were published in the Journal of Economic Perspectives criticizing the method (Haab et al., 2013:594).

In 1994 Portney first criticized the guidelines that the NOAA panel had highlighted for the method. Portney mentioned that these guidelines created a lot of displeasure to many supporters of CV since the surveys became more complicated and more expensive and furthermore these guidelines may lead to underestimating lost existence values (Portney, 1994:9-10). On the other hand, CV still had supporters and Hanemann was an avid supporter of the method. Hanemann in his article “Valuing The Environment Through Contingent Valuation” (1994) summarizes that a researcher will receive reliable results by using CVM when surveys are properly designed (Hanemann, 1994:21).

Diamond and Hausman (1994) disputed the method and they stated the problem of reliability and also the existence of biases (Diamond and Hausman, 1994: 45-6). In their critique they mention as well that the CVM is unable to measure the preferences that are attempted to be measured in a particular survey (Diamond and Hausman, 1994:46). But although “Diamond and Hausman raised a number of important issues, their negative opinion has done little to quell the demand for contingent valuation research” (Haab et al., 2013:594).

The debate over the validity of the method was settled down until the BP Deepwater Horizon oil spill in 2010 took place, where the damage from the oil spill had to be evaluated. So in 2012 the debate started again. Kling, Phaneuf and Zhao (2012) mention that since the first debate, stated preference techniques, such as CV, have been enriched by new developments in theory and the contribution of the knowledge from behavioral economics, so researchers should take into account all the progress that has taken place since 1994 (Kling et al., 2012: 21-2). Carson defended the CVM and remarked that like all the economic techniques CVM is not perfect but there is no other alternative to evaluate some goods, while he also pointed out that the method had a successful progress the last twenty years (Carson, 2012:40).

On the other hand, Hausman (2012:43) stated that there is no progress in the past twenty years and that the method has to deal with serious problems (Haab et al., 2013:594-5). Hausman criticized strongly the method and he mentioned three important problems that need to be solved. The first problem is the difference between the two measures WTP and WTA, the second problem is the lack of scope effects and the third problem is Hypothetical bias (Hausman, 2012:43).

1.2.2. The problem of Hypothetical Bias

In hypothetical surveys or referenda, the participants tend to express higher values of money for goods than the participants that are dealing with a similar choice involving real money payments (Foster and Burrows, 2017:270). A main concern about the CVM results is that they are based on respondents' answers to the CVM scenario and the answers are based on the fact that the respondents know that the scenario is hypothetical. If they knew that the scenario was about to be implemented their answers would be different. As is mentioned in Haab, Interis, Petrolia and Whitehead (2013:596)

"What people say is different from what they do".

More specifically, because the supply of the good that is examined in the scenario is hypothetical, and so is the amount of payment that the respondents agree to pay, the reliability and the validity of the CVM results have been the matter of a debate (Aadland and Caplan, 2006:563).

Hypothetical bias indicates that there is an unwanted difference between what the respondents answer about the maximum they would be willing to pay and what they would actually pay. This fact leads to researchers obtaining less reliable results when applying CVM if the presence of this potential gap is ignored.

Hypothetical bias can take two different forms and it can lead to overestimation or underestimation. In the case of overestimation the hypothetical estimates are higher than the real ones. In case of WTA there are more no-saying answers (i.e respondents overstate the amount of compensation) while for WTP yea-saying answers (i.e respondents overstate how much they are willing to pay). The other case of Hypothetical Bias is underestimation where the hypothetical estimates are smaller than the real ones. In case of a WTA survey we have more yea-saying answers and in a WTP no-saying answers.

As Murphy et al. (2005:313-14) discuss there are two dominating questions regarding hypothetical bias, "what is the magnitude of hypothetical bias associated with the Stated Preference valuation approach" and "what factors are responsible for this bias".

Bateman et al. (1995:164-65) refers that one reason why respondents may understate their WTP, especially in Open-Ended elicitation format, is because of the "free ride" problem. More specifically, the respondent might pretend that his interest for the good that is evaluated is lower when he expects that the good will be provided anyway. Additionally, another reason might be that the respondents believe that the costs of the project will be shared per capita so they respond by giving the expected cost if it is below WTP or they respond zero if it is not. Another reason is unfamiliarity with the Open-Ended format questions which leads to risk-averse strategies in regard to their answers. Finally, if the good is not well described, respondents are dealing with unfamiliar situations leading their states to be biased (Bateman et al., 1995:165).

On the other hand, one reason why respondents may overstate their WTP is because respondents might answer positively in order to satisfy the interviewer. Respondents believe that the positive answer is the answer that the interviewer would like to hear so they say yes (Bateman et al., 2006:6). Additionally, yea-saying bias can be motivated by the "warm glow" effect which means that the respondents by answering positively may feel satisfaction that they have contributed for the good that is evaluated. Furthermore, overstating WTP may be triggered by the fact that respondents feel social pressure during the survey. People tend to be sensitive to public opinion in their community so they answer positively in order to be part of a community with a high public spirit which means contribution for providing public good (Chien et al.,2005:364).

An important parameter that plays a crucial role in the issue of hypothetical bias is the problem of incentives. Due to strategic behavior by respondents, the WTP amount would either be higher or lower. Furthermore researchers explored if the incentives for strategic behavior could be connected to the elicitation method, since the estimates obtained by different elicitation method had quite large differences (Carson and Hanemann, 2005:875-77).

In regard to the incentives, Carson and Groves (2007) examined the incentive properties of preference questions. More specifically they took under consideration the properties of binary discrete choice questions to determine if such question formats are incentive compatible in the sense of whether a true answer to an actual question is an optimal strategy (Carson and Groves, 2007:182-184).

The binary discrete choice elicitation methods have the property of being incentive compatible and this fact explains the reason why the NOOA Panel proposed such methods for CVM surveys (Carson and Groves, 2007:187). Haab, Interis, Petrolia and Whitehead (2013:596) mention that “Carson and Groves’ arguments regarding incentive properties open a new exciting line of research for applied, behavioral, and experimental researchers to investigate the degree to which the incentive properties of various question formats can reduce or increase hypothetical bias”.

Although there are plenty of studies referring to hypothetical bias there is no consensus about the causes or the ways to adjust survey responses in order to avoid hypothetical bias (Murphy et al., 2005:313).

1.2.3. A review of meta-analyses of hypothetical bias

Carson et al. (1996) conducted a meta-analysis of 83 studies which include 616 comparisons between CV estimates and Revealed Preference estimates for quasi-public goods. They examined these studies and found that for most cases the CV estimates where a little lower than the Revealed Preference estimates and in some cases the CV estimates distinctly exceeded the Revealed Preference estimates. Finally, they believe that their findings could play a crucial role in discussions of whether the CV estimates need to be in general, adjusted either upwards or downwards (Carson et al., 1996:93-4).

Murphy, Allen, Stevens and Weatherhead (2005) reported results of a meta-analysis of hypothetical bias taking into considerations 28 stated preference studies estimating WTP by using the same mechanism for hypothetical and actual values. The main finding in their analysis is that the basic factor that can explain hypothetical bias is the size of the hypothetical value, furthermore no clear results were found about other factors that may be associated with hypothetical bias (Murphy et al., 2005:322).

List and Gallet (2001:246) found that in hypothetical choices the estimates are about 3 times higher than the estimates that come from real choices and they also noticed that the differences depend on the elicitation method and if the scenario presents a WTP or a WTA question. The elicitation method has been also mentioned by Green et al. (1998:85), more specifically they have mention that the referendum elicitation methods tend to return higher mean estimates than from Open-Ended responses. Little and Berrens (2004:5) also have found that the difference is about 3.13 times higher for the estimates that are based on hypothetical choices.

Foster and Burrows (2017) gathered the literature on hypothetical bias and more specifically on previous meta-analysis in order to study if among the characteristics of the survey designs that contributes to overcoming hypothetical bias, there is a practical and reliable way to overcome hypothetical bias. As they mention, until the time their work was published, previous meta-analysis had confirmed that in stated preference studies hypothetical bias exists but they couldn’t offer definite guidelines that could be used in order to reduce hypothetical bias (Foster and Burrows, 2017:270).

More specifically, they examined how the “bias Ratio” (the ratio of the mean WTP from the hypothetical treatment to the WTP from the real treatment) (Foster and Burrows, 2017:271), was affected by variables representing a number of commonly used techniques and two additional variables that they added. The common techniques that they used were certainty correction, cheap-talk, the same respondent vs different respondent technique, if the observations are derived from conjoint or choice experiment, if the respondent is a student or not, if the hypothetical and the real survey instruments are implemented in a laboratory study or not and if the good that is evaluated is public or private.

Furthermore, they added two new variables indicating if the good that is evaluated is familiar or unfamiliar to the respondent and if the valuation of the good is mainly generated by non-use considerations (Foster and Burrows, 2017:273-79). Finally, after their attempt to update the existing prior meta-analyses, they conclude they their meta-analysis did not offer definite insights in order to eliminate or reduce hypothetical bias (Foster and Burrows, 2017:286).

Some empirical studies have shown that there is a possibility that estimates from real choices are higher than those from hypothetical choices and in this case hypothetical bias exists in the form of underestimation. Ehmke, Lusk and List (2008:489-90) found that hypothetical bias is not independent of location since many cultural factors exists and they could affect the existence of hypothetical bias. In their survey they used data from several countries such as, for example, China, France, Niger and a numbers of states in America and they concluded that less developing countries, like China and Niger, tend to vote “no” in hypothetical scenarios and “yes” in the real ones (Ehmke et al., 2008:497).

1.2.4. Suggested methods to overcome hypothetical bias

In general, it is difficult to measure hypothetical bias and the reason which explains this difficulty is that in order to test if hypothetical bias exists there has to be a comparison with real payments (Jakobsson and Dragun, 1996:84). Since for non-market and public goods the problem named market failure exists, “measuring hypothetical bias is difficult for non marketed resources and public goods” (Loomis, 2014:35). But it should be mentioned that Little and Berrens (2004:6) found that there is no evidence that private goods have potentially less difference between hypothetical and real payments than public goods.

One part of the literature concerned with hypothetical bias has proposed ex ante approaches whereas the problem can be treated through auxiliary mechanisms like "cheap talk" or "solemn oath" and "scenario adjustment" (Haab et al., 2013:599). On the other hand, the ex post procedure can correct hypothetical bias by using statistical techniques (Hofler and List, 2004:213).

Ex-ante Methods

The technique of solemn oath was introduced by Jacquemet, Joule, Luchini and Shogren (2013) in order to ensure respondents will be honest in their answers and thus eliminate the hypothetical bias problem. Just like courts, where witnesses take an oath "to tell the truth and nothing but the truth" in solemn oath the same procedure occurs. The respondents are asked to answer to the valuation questions after they swear a "truth-telling-commitment" that they will be honest (Jacquemet et al., 2013:111). More specifically, in a survey including the solemn oath technique probably there will be an additional statement that the respondent will be asked to agree or not. The statement might be like the following: "I swear upon my honor that, during the whole experiment, I will tell the truth and always provide honest answers" (Jacquemet et al., 2013:115).

Some researchers with non-experimental applications, have found that the oath framework "has significant effects on hypothetical response behavior across individuals and across multiple countries that is consistent with what would be expected from reduced hypothetical bias" (Haab et al., 2013:599).

Scenario adjustment plays an important role in the choice procedure. There are three types of scenario adjustments that can occur since the researcher gives additional information to the respondent that already has an idea about any aspects of the presented scenario. Firstly, the respondents may replace their prior beliefs about the scenario with the new information that the researcher provides. Secondly, the respondents may reject the additional information and finally the respondents may combine their prior beliefs with the researcher's information (Cameron et al., 2011:10).

If researchers do not include these scenario adjustments they may probably overestimate or underestimate WTP for a number of respondents. As Cameron, DeShazo and Johnson (2011:11) suggest "researchers should probably calculate and compare estimates of WTP both with and without corrections for scenario adjustment". And finally, researchers should include from the early design these possible scenario adjustments so neither overestimation nor underestimation for WTP will occur (Cameron et al., 2011:11).

Cheap talk on the other hand, is a survey design in which "the valuation context involves providing respondents with additional instructions that explicitly encourage them to treat hypothetical scenario as if an actual monetary transaction were taking place" (Haab et al., 2013:599).

In a cheap talk script the researcher explains to the respondents the problem of hypothetical bias that he has dealt with in other surveys in the past. In this way he tries to make the respondents answer sincerely. He may use either short script or long script. In the short script he will only mention the hypothetical bias that they have found in similar surveys. On the other hand, in the long cheap talk script the researcher mentions what kind of substitute goods exist and reminds the respondent that he should have in mind his household budget (Aadland and Caplan, 2006:565-67).

Loomis (2014:36-8) has analyzed as well the ex ante survey design approaches and classifies them in the following four categories: "Consequentiality Designs", "Honesty and Realism Approaches", "Cheap Talk" and "Reducing Social Desirability Bias and Cognitive Dissonance".

Consequentiality refers to the fact that the survey should have some potential effect on the respondents such as affecting the likelihood of the provision of the good and/or changes in taxes. Honestly and realism approaches is a method in which the researcher makes the respondents to give their honest answer such us the inclusion of a solemn oath statement that we discussed above. Cheap talk has been discussed above, while "reducing social desirability bias and cognitive dissonance" refers to, among others, the problem that arises when some respondents answer payment questions based on social norms rather than their own personal values and one possible solution would be by asking what the respondent think others will pay for the provision of the good (Loomis, 2014:43).

All the above techniques, as it has been mentioned, are considered as ex ante procedures and as Hofler and List (2004:213) emphasized "recent technology using ex ante procedures has produced some strong evidence that hypothetical bias can be overcome". However, in the case of cheap talk the evidence is not so positive since according to the study of Aadland and Caplan (2006:562) "a short, neutral cheap talk script appears to exacerbate rather than mitigate the bias". According to the authors, respondents hearing that hypothetical bias occurred in other similar studies might try to attempt to give the "right" answer and thus increase the bias.

Ex-post Methods

Apart from the ex ante procedures to overcome hypothetical bias another set of procedures exist, the ex post procedures which can correct hypothetical bias by using statistical techniques (Hofler and List, 2004:213).

For example, Loomis (2014:43) mentioned three ex post methods that can be included in surveys. Firstly, the preference towards the median WTP responses rather than the mean WTP responses. Secondly, including uncertainty when the respondent is not sure about his positive WTP answer and finally "relying on the degree of hypothetical bias uncovered in an experiment with a deliverable good to scale the WTP from a stated preference survey" (Loomis, 2014:43).

Since there is a gap between the hypothetical and the real estimates, this difference must be taken into account in the models in order to get more reliable and realistic evaluations. Hofler and List (2004) designed an experiment to examine if the results are different between a hypothetical and an actual auction for a baseball card and they used stochastic frontier approach.

More specifically, Hofler and List (2004:220) proposed a statistical approach in order to link the actual with the hypothetical statements from data for a baseball card auction. They conclude that people overstate their WTP and they have used calibration function derived from a stochastic frontier regression model in order to overcome the gap between the actual and the hypothetical bid.

Additionally, Chien, Huang and Shaw (2005) proposed a modeling approach for double-bounded dichotomous choice data based on the stochastic frontier model that can accommodate both yea-saying behavior and starting point bias. Furthermore, Kumbhakar, Parmeter and Tsionas (2012) consider possibility of underestimation or overestimation in first-price auctions. They developed an optimization error approach that allowed the optimal bids to differ from the bids they observed and propose the application of a stochastic frontier model (Kumbhakar et al., 2012:47-48).

A researcher has the possibility to use more than one approach to overcome hypothetical bias. Usually cheap talk is combined with another ex ante approach or even with ex post methods. It is mentioned though that researchers that choose to combine the methods should be very careful because in this way there is a possibility of correcting more than the wanted hypothetical bias and so WTP is no longer overestimated but underestimated (Loomis, 2014:43).

1.3. Stochastic frontier for CVM modeling with the presence of hypothetical bias

Hofler and List (2004) propose the use of stochastic frontier model in order to calibrate hypothetical statements to real values. A one sided error is included in the hypothetical open-ended bid function capturing the difference between the actual and the hypothetical as shown in Eq. (1.1) below, (Hofler and List, 2004:215)

$$Y_i^H = X_i\beta + v_i + u_{i,}\ i = 1, \dots, n \quad (1.1)$$

- Y_i^H represents the hypothetical bid for each individual i,
- X_i is a row of explanatory variables, the determinants of the bid for person i,
- β is a column vector of the coefficients.

The function though, has two error terms, v is the usual regression error term and u is an additional one-sided error that represents the gap between the hypothetical and the true bid for each person (Hofler and List, 2004:216).

Additionally, the model can be written in the following form (Hofler and List, 2004:216)

$$Y_i^H = Y_i^A + u_i\ ,\quad i = 1, \dots, n \quad (1.2)$$

- Y_i^A represents the actual bid for each i person.

In the case of yea-saying the analysis is based on the hypothesis that $u_i \geq 0$. In the case that the actual bid is equal to the hypothetical, the estimate of $u_i = 0$ and the respondent answered sincerely. On the other hand when $u_i > 0$ it is implied that the respondents overstate the bid since the hypothetical bid is bigger than the actual (Hofler and List, 2004:216). The method proposed by Hofler and List (2004) can be easily implemented to CV survey data that use an open-ended elicitation format.

Chien et al. (2005) on the other hand have introduced a stochastic frontier model of overestimation for Double-Bounded DC method. More analytically, they assume that WTP is given by $W^* = X\beta + V$ where W^* is the latent willingness to pay which is not observed and only the yes/no answers to the presented bids are observed. In the presence of yea-saying, the upward shift of the WTP can be captured again by a one-sided, non-negative error term as shown below,

$$W_1 = W^* + U = X\beta + \varepsilon$$

where

- β is the coefficient vector
- X are the respondent's characteristics
- V is the statistical noise where $E(V) = 0$
- U is the one-sided non-negative random error for yea-saying bias
- ε is the composite error (Chien et al., 2005:365).

On the other hand the respondents to CV surveys need not be a heterogeneous group and hypothetical bias could be present only for a subset of respondents. In this case a latent class model where some respondents answer truthfully to the payment questions while others overstate might be a promising venue. In the case of production economics Kumbhakar, Parmeter and Tsionas (2013) have proposed the zero inefficiency model where both efficient and inefficient firms can be present. In their study a latent class stochastic frontier model is proposed to analyze production inefficiency. Taking as a starting point the well known production stochastic frontier model given below,

$$y_i = x_i{}' \beta + v_i + u_i = x_i{}' \beta + \varepsilon_i \text{ for i = 1, . . . ,n}$$

they assume that some firms are fully efficient with $u_i = 0$ and some other firms are inefficient $u_i > 0$ .Their zero inefficiency stochastic frontier model is:

- $y_i = x_i{}' \beta + v_i$ with probability p, where p is the probability of the firm to be fully efficient and
- $y_i = x_i{}' \beta + (v_i - u_i)$ with probability $(1 - p)$ (Kumbhakar et.al, 2013:67-68).

The present thesis builds up on the results presented above and proposes a latent class stochastic frontier model for hypothetical bias for both open-ended and dichotomous choice formats. In our models we will have a latent class model where both hypothetical bias and sincerely answering may occur with a probability. More specifically, in this thesis we take into account the work of Chien, Huang and Shaw (2005) and of Kumbhakar, Parmeter and Tsionas (2013) in order to propose a mixture stochastic frontier model for CV data both under the open-ended and double-bounded format.

1.4. Thesis overview

The contents of this thesis are arranged as follows. In Chapter 2 a mixture model is proposed for estimating CV survey data under an open-ended format in the presence of yea-saying and simulations of 1000 replications have been conducted in order to investigate the performance of the model for several different cases. Two different cases are analyzed for the probability of class membership, namely the probability is constant over individuals and the probability depends on some regressor and therefore varies over individuals.

Chapter 3 proposes a mixture model when the double-bounded elicitation format is used and respondents overstate their willingness to pay. The performance of the model is evaluated with simulations of 1000 replications under different scenarios. Moreover the method is applied to an empirical study about the valuation of the Kakadu Conservation Zone, which is based on the CV survey that took place in Australia and was published by Carson, Wilks and Imber (1994).

In Chapter 4 the importance of different strategies for selecting starting values is analyzed in detail and the proposed model of overcoming hypothetical bias has been tested for several different clustering methods in order to investigate the performance of the model for several different starting values methods.

Finally, concluding remarks are presented since all the analysis of the previous chapters have led to the optimal proposed model which responds better in order to overcome Hypothetical Bias.

References of Chapter 1

Aadland, D. and Caplan, A.J., 2006, "Cheap talk reconsidered: New evidence from CVM",*Journal of Economic Behavior & Organization,* vol.60, p.562-578

Ajzen, I., Brown T.C. and Rosenthal, 1996, "Information Bias in Contingent Valuation: Effects of Personal Relevance, Quality of Information, and Motivational Orientation", *Journal of Environmental Economics and Management,* vol.30, p.43-57

Arrow, K., Solow, R., Portney P.R., Leamer, E.E., Radner, R., and Schuman, H., 1993, *Report of NOAA Panel on Contingent Valuation,* National Oceanic and Atmospheric Administration

Bonnieux, F. and Rainelli, P, 1999, Contingent Valuation Methodology and the EU Institutional Framework. In: Bateman, I.J. and Willis K.G., eds., 1999, *Valuing Environmental Preferences Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries,* New York, Oxford University Press, chapter 17, p.585-612

Bateman, I.J., Langford, I.H. and Rasbash, J., 1999, Willingness-to-Pay Question Format Effects in Contingent Valuation Studies. In: Bateman, I.J. and Willis K.G., eds., 999, *Valuing Environmental Preferences Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries,* New York, Oxford University Press, chapter 15, p.511-539

Bateman, I.J., Langford, I.H., Turner, R.K., Willis, K.G. and Garrod, G.D., 1995, "Elicitation and truncation effects in contingent valuation studies", *Ecological Economics,* vol. 12, p.161-179

Bateman, I., Murno,A., Rhodes., B.,Starmer, C.V. and Sugden, R., 2006, Anchoring and yea-saying with private goods: An experiment. In:List, J. Ed., *Using Experimental Methods in Environmental and Resource Economics*, Edward Elgar, Northampton,pp. 1–19 available at:<http://www3.grips.ac.jp/~munro/ELICIT2.pdf> [access March 31, 2022]

Berg, N., 2005, Non-Response Bias, In: Kempf-Leonard, K., ed., 2005, *encyclopedia of social measurement,*vol.2, London, Academic Press, pp. 865-873

Bockstael, N.E. and McConell, K.E., 2007, Environmental and Resource Valuation with Revealed Preferences, A theoretical Guide to Empirical Studies. In: Bateman,I.J, ed.,2007, *The Economics of Non-Market Goods and Resources*,vol.7,Spinger

Boyle, K.J., Bishop, R.C. and Welsh, M.P.,1985, "Starting Point Bias in Contingent Valuation Bidding Games" *Land Economics,* vol.61, no.2, p.188-194

Boyle, K.J. and Bergstrom, J.C., 1999, Doubt, Doubts, and Doubters: The Genesis of a New Reasearch Agenda? In: Bateman, I.J. and Willis K.G., eds., 1999, *Valuing Environmental Preferences Theory and Practiceof the Contingent Valuation Method in the US, EU, and Developing Countries,* New York, Oxford University Press, chapter 7, p.183-207

Calson, D. E., Ripberger, J. T., Jenkins-Smith, H. C., Silva, C. L., Gupta, K., Berrens, R. P. and Jones, B. A., 2016, "Contingent Valuation and the Policymaking Process: An Application to

Used Nuclear Fuel in the United States", *Journal Benefit Cost Analysis,* vol. 7, no.3, p.459-487

Cameron, T.A., DeShazo, J.R. and Johnson, E.H., 2011, "Scenario adjustment in stated preference research", *Journal of Choice Modelling,*vol.4, p 9-43

Carson, R.T., 2012, "Contingent Valuation:A Practical Alternative whenPrices Aren't Available", *Journal of Economic Perspectives,* vol.26,no.4,p.27-42

Carson, R.T., Flores, N.E., Martin, K.M. and Wright, J.L., 1996, "Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods", *Land Economics,* vol.72, no.1, p.80-99

Carson, R.T.,and Groves, T., 2007, "Incentive and Informational properties of preference questions", *Environmental and Resource Economics,* vol. 37, p.181-210

Carson, R.T. and Hanemann, W.M., 2005, Contingent Valuation. In: Mäler K.G. and Vincent J.R., eds., 2005, *Handbook of Environmental Economics,* vol.2, Elsevier B.V.,chapter 17, p.821-936

Carson, R.T.,Wilks, L. and Imber, D., 1994, "Valuing the Preservation of Australia's Kakadu Conservation Zone",*Oxford Economic Papers,* vol.46, p.727-749

Chien, Y-L., Huang, C. J. and Shaw, D., 2005, "A general model of starting point bias in double-bounded dichotomous contingent valuation surveys", *Journal of Environmental Economics and Management,*vol.50, p.362-377

Cooper,J. C., Hanemann,M. and Signorello,G.,2002, "One-And-One-Half-Bound Dichotomous Choice Contingent Valuation", *The Review of Economics and Statistics,* vol.84, no.4, p.742:750

Diamond, P.A. and Hausman, J.A., 1994, "Contingent Valuation: Is Some Number better than No Number?", *The Journal of Economic Perspectives,* vol.8,no.4, p.45-64

Du Preez, M., Menzies, G., Sale, M. and Hosking, S., 2012, "Measuring the indirect costassociated with the establishment of wind farm: An application of thecontingent valuation method",*Journal of Energy in Southern Africa,* vol. 23,no.1, p. 2-7

Ehmke, M.D., Lusk, J.L. and List, J.A., 2008, "Is Hypothetical Bias a Universal Phenomenon? A Multinational Investigation", *Land Economics,* vol.84, n.3, p.489- 500

Foster,H. and Burrows, J., 2017, Hypothetical bias: a new meta-analysis. In: McFadden, D. and Train, K., eds., 2017, *Contingent Valuation of Environmental Goods: A Comprehensive Critique,* United Kingdom, Edward Elgar Publishing, chapter 10, p.270-291

Green, D.,Jacowitz, K.E., Kahneman, D. and McFadden D., 1998, "Referendum contingent valuation, anchoring, and willingness to pay for public goods", *Resource and Energy Economics,* vol. 20, p.85-116

Green, C. and Tunstall, S., 1999, A Psychological Perspective. In: Bateman, I.J. and Willis K.G., eds., 1999, *Valuing Environmental Preferences Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries,* New York, Oxford University Press, chapter 8, p.207-257

Haab, T.C., Interis M.G., Petrolia, D.R. and Whitehead, J.C., 2013, "From Hopeless toCurious? Thoughts on Hausman's "Dubious to Hopeless" Critique of Contingent Valuation",*Applied Economic Perspectives and Policy*, vol. 35, n.4, p.593-612

Hanemann, M.W., 1994, "Valuing The Environment Through Contingent Valuation", *Journal of Economic Perspectives,* vol.8, no.4, p.19-43

Hanemann, M., Loomis, J. and Kanninen,B., 1991, "Statistical Efficiency of Double Bound Dichotomous Choice Contingent Valuation", *Americal Journal of Agricultural Economics,* vol.73, no.4, p.1255-1263

Hausman, L., 2012, "Contingent Valuation: From Dubious to Hopeless", *Journal of Economic Perspectives,* vol.26, no.4, p.43-56

Hofler, R.A. and List, J.A., 2004, "Valuation on the frontier: Calibrating actual and hypothetical statements of value", *American Journal of Agricultural Economics,*vol.86, no.1, p.213-221

Jacquemet, N., Joule, R.V., Luchini, S. and Shogren, J.F., 2013, "Preference Elicitationunder Oath",[pdf], *Journal of Environmental Economics and Management,* vol.65, p.110-132,

Jakobsson, K.J. and Dragun, A.K., 1996, *Contingent Valuation and Endangered Species Methodological issues and application*, (New Horizons in environmental economics series),general editor Oates, W.E., Great Britain, Edward Elgar Publishing Limited

Karlis, D. and Xekalaki, E., 2003, "Choosing initial values for the EM algorithm for finite mixtures", *Computational Statistics and Data Analysis,* vol. 41, p. 577-590

Kartman, B., Stälhammar, N-O. and Johannesson, M., 1996, "Valuation of health changes with the contingent valuation method: a test of scope and question order effects", *Health Economics,* vol.5,p.531-541

Kling, C.L.,Phaneuf, D.J. and Zhao, j., 2012, "From Exxon to BP: Has Some Number Become Better than No Number?", *Journal of Economic Perspectives,* vol.26,no.4, p.3-26

Kumbhakar, S.C., Parmeter, C.F. and Tsionas, E.G., 2012, "Bayesian estimation approaches to first- price auctions", *Journal of Econometrics,* vol.168, p.47-59

Kumbhakar, S.C., Parmeter C.F. and Tsionas E.G., 2013, "A zero inefficiency stochastic frontier model", *Journal of Econometrics,* vol.172, p.66-76

Leroux, B.G., 1992, «Consistent Estimation of a Mixing Distribution", *The Annals of Statistics,* vol. 20, no.3, p.1350–1360

Little, J. and Berrens, R., 2004, "Explaining Disparities between Actual and Hypothetical Stated Values:Further Investigation Using Meta−Analysis", [pdf], *Economics Bulletin,* vol. 3, no. 6, p. 1−13, available at: <http://www.accessecon.com/pubs/eb/2004/volume3/EB03C90005A.pdf> [assess March 31,2022]

List, J.A. and Gallet, C.A., 2001, "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics,* vol.20, p.241-254

Loomis, J.B., 1999, Contingent Valuation Methodology and the US Institutional Framework In: Bateman, I.J. and Willis K.G., eds.,1999, *Valuing Environmental Preferences Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries,* New York, Oxford University Press, chapter 18, p.613-629

Loomis, J.B., 2014, “2013 WAEA Keynote Address: Strategies for Overcoming Hypothetical Bias in Stated Preference Surveys”, *Journal of Agricultural and Resource Economics,* vol.39, no.1, p.34-46

Mann, A.K. and Kaur, N., 2013, “Review Paper on Clustering Techniques”, *Global Journal of Computer Science and Technology,* vol.13,no.5-C,[pdf], available at: <https://computerresearch.org/index.php/computer/article/view/353>[access 31 March 2022]

McLachlan, G.J., 1988,” On the Choice of Starting values for the EM Algorithm in Fitting Mixture Models”, *Journal of theRoyalStatistical Society, Series D (* The Statistician), vol. 37, no.4/5, p 417–425

Mc Lachlan G. and Pell D.,2000, *Finite Mixture Models*, United States of America, A Wiley Interscience Publication

McLachlan, G.J., Lee, S.X. and Rathnayake, S.I., 2019, “Finite Mixtures”, ,*Annual Review of Statistics and Its Application,* vol.6, p. 355-378

Milon, W., 1989, “Contingent ValuationExperiments for Strategic Behavior”, *Journal of Environmental Economics and Management,* vol.17,p.293-308

Mitchell, R.C. and Carson, R.T., 1989, *Using Surveys to Value Public Goods: The Contingent Valuatiuon Method,* United States of America, Resources for the Future

Murphy, J.J., Allen, P.G., Stevens, T.H. and Weatherhead, D., 2005, “A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation”, *Environmental and Resource Economics,* vol. 30, p.313-325

Portney, P.R., 1994, “The Contingent Valuation Debate:Why Economists Should Care”, *Journal of Economic Perspectives*, vol. 8, no. 4, p.2-17

Randall, A., Ives, B. and Eastman, C., 1974, “Bidding Games for Valuation of Aesthetic Environmental Improvements”, *Journal of Environmental Economics and Management,* p.132-149

Sugden, R., 1999, Public Goods and Contingent Valuation. In: Bateman, I.J. and Willis K.G., eds., 1999, *Valuing Environmental Preferences Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*, New York, Oxford University Press, chapter 5, p.131-151

Tsay, W-J.,Huang C.J., Fu, T-T. and Ho, I-L., 2013, “a simple closed –form approximation for the cumulative distribution function of the composite error of stochastic frontier model” *, J Prod Anal*,vol.39, p.259-269

Vatn, A., 2005, *Institutions and the Environment*, Great Britain, Edward Elgar Publishing

Woodward, W.A., Parr, W.C., Schucany, W.R., Lindsey, H., 1984, “A comparison of minimum distance and maximum likelihood estimation of a mixture proportion”, *Journal of the American Statistical Association*, vol. 79, p. 590–598

Chapter 2
Applying stochastic frontier and mixture models to contingent valuation under the open-ended format

Introduction

In Chapter 2 a stochastic frontier mixture model is applied to CV under the open-ended format aiming to overcome hypothetical bias. The chapter comprises of four main parts.

The first section presents the theoretical background of the CVM under the open-ended elicitation format. The second section analyzes the stochastic frontier model applied to open-ended format data under the presence of hypothetical bias while the third part introduces and analyzes the mixture model for open-ended data that it is suggested in order to overcome hypothetical bias.

Finally the last part of the chapter presents the results of simulations that took place in order to test the proposed model. More analytically, the last section describes in detail the data generation process, the initialization strategy that was followed and all the simulation results for a number of different cases. Furthermore, the simulations took place for two different scenarios for class probability determination, in the first case the class probability is a constant and consequently all respondents have the same probability of overstating their willingness to pay (WTP hereafter) and in the other case the probability of overstating WTP differs among respondents since the probability depends on a variable z.

2.1. Contingent valuation and open-ended elicitation format

A CV study aims to measure for each individual his/hers monetary value for an item denoted as q or for a change in its provision. More analytically, each individual has a direct utility function $u(x, q)$ which is defined by a number of commodities x and also q, and an indirect utility function $\upsilon(p, q, y)$ where p contains the prices of the market commodities and y is the person's income. Furthermore, the assumption that is made is that $u(x, q)$ is increasing and quasi-concave in x thus $v(p, q, y)$ satisfies the standard properties for p and y. Furthermore, each agent regards q as "good" or "bad".

If q is considered as "good": $u(x, q)$ and $v(p, q, y)$ will be increasing in q.

If q is considered as "bad": $u(x, q)$ and $v(p, q, y)$ will be decreasing in q.

If the agent is indifferent to q: $u(x, q)$ and $v(p, q, y)$ will be independent of q.

In the valuation process there must be a comparison between two situations with respect to the provision of q. In a few words, if the changes in q is $q^0 \to q^1$, each person will have different utility functions, the utility function before the change will be $u^0 \equiv v(p, q^0, y)$ and after the change $u^1 \equiv v(p, q^1, y)$.

If the change represents an improvement $u^1 > u^0$.

If the change represents a worsening $u^1 < u^0$.

If the change is indifferent for the agent $u^1 = u^0$.

The change in q $(q^0 \to q^1)$ which leads to a change in the utility $(u^0 \to u^1)$, in monetary terms is represented by the compensating variation C which satisfies Eq. (2.1) and the equivalent variation E which satisfies Eq. (2.2).

$$v(p, q^1, y - C) = v(p, q^0, y) \qquad (2.1)$$

$$v(p, q^1, y) = v(p, q^0, y + E) \qquad (2.2)$$

In the case where the change is an improvement, $C > 0$ and $E > 0$. Additionally, C is the person's maximum WTP in order to ensure that the change will be implemented and additionally E is the minimum willingness to accept (WTA hereafter). On the other hand, in the case where the change is regarded as being worse, $C < 0$ and $E < 0$. In this case, C is the person's WTA and respectively, E measures the WTP in order to avoid the change (Carson and Hanemann, 2005:844-45).

The CVM uses a survey in order to evaluate the WTP or the WTA that people have for a change in q, one elicitation method that is used in surveys in order to derive the WTP/WTA is the open-ended question format. The open-ended format reveals directly the respondents' WTP, more analytically, the open-ended question given to respondents is “How much are you willing to pay for the change from q^0 to q^1?” (Carson and Hanemann, 2005:848). Suppose that the answer is A this means that the respondents compensating variation C (or his /hers WTP) is equal to A.

The open-ended elicitation format has a number of advantages and disadvantages that have been reported in the literature. To begin with, one advantage is that the WTP is elicited directly and no further inference is needed (Loomis, 1990:79). As Ahmed and Gotoh (2006:16) state, the major advantage of the open-ended elicitation method is that it “provides straightforward actual valuation of amenities”.

Furthermore, open-ended questions provide a richer set of information in regards to respondents’ preferences and additionally open-ended format is more suitable when a survey takes place in more than one country (Håkansson, 2008:186). For example, Istamto et al., (2014) implement a multi-country study and state that they applied the open-ended elicitation method because it is referred as stable over time and furthermore it is considered as free of anchoring effects and starting point bias (Istamto et al., 2014:11).

On the other hand, one disadvantage is that the respondents consider the procedure of stating a specific amount for WTP as a difficult mental task thus many respondents don’t answer at all or answer by understating their WTP (Loomis, 1990:79). Furthermore, Carson (2000:1416) states that open-ended survey questions typically elicit a large number of protest zeros and a small number of very large responses and that even this small number can influence in a dramatic way the mean WTP. In a few words a major disadvantage of the open-ended format is that it may provide unrealistic responses (Ahmed and Gotoh, 2006:16).

Additionally, the NOOA Panel (Arrow et al., 1993:20-1) criticized the open-ended elicitation format and stated that the open-ended questions are not providing the most reliable valuations for two reasons. The first reason is that the scenario lacks of realism since respondents in their everyday lives rarely are asked to pay for a particular public good. The second reason is because open-ended questions lead to strategic overstatement.

Kealy and Turner (1993:326-7) tested the equality of open-ended and closed-ended CV results and they found that in the case of a public good there was a significant difference in the results obtained by the two elicitation methods. Furthermore, for the case of a private good there were no differences in the estimates of WTP because no incentives for strategic behavior exist for the private good and additionally the respondents were more familiar with the good.

Bateman et al. (1995:161) stated that in their survey they applied three different WTP elicitation methods. One elicitation method was the open-ended elicitation format and their results indicate that the respondents had to deal with significant uncertainty in regards to answering the open-ended questions and furthermore they may reveal free riding tendency or strategic overbidding.

More analytically, some respondents may understate their WTP because they may adopt risk-averse strategies that place downwards the stated WTP due to unfamiliarity with the open-ended format questions. On the other hand, some respondents tend to overstate their WTP in the case where the respondents have realized that the decision in regards the provision of the good depends upon mean WTP, in such case they overstate their WTP in order to increase the mean WTP and therefore improve the chance of provision (Bateman et al., 1995:164-5).

Assuming a linear model, the maximum willingness to pay for individual i is given below

$$WTP_i = \beta' x_i + v_i \quad i = 1, \cdots, n \qquad (2.3)$$

$$where\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_\kappa \end{bmatrix}, x_i = [x_{i1} \quad \cdots \quad x_{i\kappa}]',\ v_i \sim N(0, {\sigma_v}^2)$$

and x_i is a vector of observed explanatory variables that affect WTP with $x_{i1} \equiv 1 \forall i$. If respondents answer truthfully to the open-ended question, i.e under the absence of perception and strategic errors, then the mean willingness to pay (MWTP) is given by $\beta' \bar{x}$ and an estimate is given by $\hat{\beta}' \bar{x}$.

Under the presence of hypothetical bias whereas respondents might overstate their bids (yea-saying behavior), model (2.3) does not hold anymore and a stochastic frontier model can be used to reflect this behavior. In the next section the stochastic frontier model for the open-ended elicitation format under the presence of hypothetical bias is going to be analyzed in detail.

2.2. Stochastic frontier model for the open-ended method under the presence of hypothetical bias

Hofler and List (2004) proposed the use of a stochastic frontier model in order to take into consideration the difference between a real and a hypothetical auction bid. In the present section the same methodology is going to be applied in order to include hypothetical bias in Eq. (2.3).

More analytically, for the case where hypothetical bias exists in the form of overstatement of WTP Eq. (2.3) becomes

$$WTP_i{}^* = WTP_i + u_i \qquad (2.4)$$

where $u_i \sim iid\ N^+(0, \sigma_u^2)$ nonnegative Half Normal which in the stochastic frontier literature is known as The Normal-Half Normal Model.

More analytically, Eq. (2.4) can be written as

$$WTP_i^* = \beta_1 + \beta_2 x_{i2} + \ldots\ldots + \beta_\kappa x_{i\kappa} + v_i + u_i \qquad (2.5)$$

where

- v_i is the two sided "noise" component and
- u_i is the one-sided error term (Kumbhakar and Lovell, 2000:140)

It should be noted that in the case where hypothetical bias exists in the form of understatement of WTP the stochastic frontier model becomes

$$WTP_i{}^* = WTP_i - u_i \quad (2.6) \quad \text{(Kumbhakar and Lovell, 2000:74)}$$

Furthermore in the case of overstatement, from Eq. (2.4) it follows that

$$WTP^*{}_i - WTP_i = u_i \qquad (2.7)$$

Consequently, when the error term u_i approaches zero, the gap between the real and hypothetical values is decreased and the hypothetical values → real values (Hofler and List, 2004:216).

Furthermore the composed error is given by

$$\varepsilon_i = v_i + u_i \qquad (2.8)$$

So Eq. (2.5) can be written

$$WTP_i^{\ *} = \beta^{'} x_i + \varepsilon_i \qquad (2.9)$$

Since $E(\varepsilon_i) = E(u_i) = \sigma_u \sqrt{\frac{2}{\pi}}$, (Kumbhakar and Lovell, 2000), ignoring the presence of hypothetical bias will lead to overestimation of the constant term of equation (2.3).

Very often the model is parameterized in terms of the two parameters defined below

$$\sigma^2 = \sigma_v^2 + \sigma_u^2 \qquad (2.10)$$

and

$$\lambda = \frac{\sigma_u}{\sigma_v} \qquad (2.11)$$

If overestimation occurs then the parameter λ should be statistically significant and greater than zero. If λ approaches values close to zero, σ_u approaches values close to zero as well and the composed error tends to be v.

2.3. Open-ended mixture model

2.3.1. Mixture models theory

Finite Mixtures of distributions is an approach for modeling many kinds of random phenomena. Due to their flexibility, mixture models are increasingly used because of their convenience to model unknown distributional shapes. Mixture models are applied in a variety of fields, such as biology, agriculture, marketing, engineering, medicine, economics, social sciences and many more, and furthermore finite mixture models support many statistical techniques such as cluster and latent class analysis (McLachlan et al., 2019:355-6).

By definition a g-component finite mixture density $f(y_i; \theta_j)$ is given by

$$f(y_i; \Psi) = \sum_{j=1}^{g} \pi_j \ f_j(y_i; \theta_j) \quad (2.12)$$

Where

- $f_j\,(y_i)$ are the densities and they are called the components densities of the mixture
- π_j nonnegative quantities $0 \leq \pi_j \leq 1 \;\; j = 1, \ldots, g)$ and

$$\sum_{j=1}^{g} \pi_j = 1$$

$\pi_1, \ldots, \pi_g$ are called the mixing proportions or weights

- θ_j the vector of unknown parameters
- Ψ is the vector with all the unknown parameters $\Psi = (\pi_1, \ldots., \pi_g, \xi^T)^T$ and
- ξ is the vector including all the parameters in $\theta_1, \ldots, \theta_g$ known a priori to be distinct (McLachlan and Peel, 2000:6-22).

In a few words taking into consideration Eq. (2.12), $f(y_i; \Psi)$ is a linear combination of densities $f_j\,(y_i)$ and the weights π_j are the class probabilities.

2.3.2. Open-ended mixture model

Hypothetical bias may occur if the respondents do not answer sincerely and as a result a gap is created between their real WTP and the WTP they state. Although it is possible that some respondents may state the wrong WTP it is possible that a number of respondents may answer sincerely.

In such cases a latent class model or a mixture model could capture this heterogeneity in the response behavior of individuals. In a few words it can't be considered that all responders are overstating their WTP because some respondents might actually answer sincerely. The present consideration follows the same notion that Kumbhakar, Parmeter and Tsionas (2013:67) followed in their paper related to the productive inefficiency of firms. In a few words, they stated that in a sample both efficient and inefficient firms can exist with a given probability.

Taking into account the finite mixture models theory, the model for WTP will be considered as a mixture of two classes. Class 1 has no hypothetical bias and respondents answer sincerely so a model with a Normal error holds and class 2 overstates WTP so a model with a composed Normal-Half-normal error term exists.

In a few words it is assumed that a number of respondents answer truthfully according to their real WTP and a number of other respondents overstate their WTP. The two classes are:

Class 1: people that answer sincerely and WTP is given by Eq. (2.3), and

Class 2: people that overstate their WTP and therefore Eq. (2.9) holds.

The probability of belonging to class 1 and class 2 is given by p_1 and $p_2 = (1 - p_1)$ respectively.

In the present case with the two classes described above, the model becomes

$$WTP_i^* = \begin{cases} \beta' x_i + v_i & with\ probability\ p_1 \\ \beta' x_i + v_i + u_i & with\ probability\ p_2 \end{cases} \qquad (2.13)$$

<u>The density functions for each case for the error are the following:</u>

No Hypothetical Bias:

$$f_1(v_i) = \frac{1}{\sigma_v \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v_i}{\sigma_v}\right)^2} = \frac{1}{\sigma_v \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{wt\,p_i - \beta' x_i}{\sigma_v}\right)^2} \qquad (2.14)$$

which is the density of the $N(0, \sigma_v^2)$.

Hypothetical Bias:

For Normal ($v_i \sim iid\ N(0, \sigma_v^2)$) and Half-Normal $\left(u_i \sim iid\ N^+(0, \sigma_u^2)\right)$ distributions we have the composed error density distribution:

$$f_2(\varepsilon_i) = \frac{2}{\sigma} \varphi\left(\frac{\varepsilon_i}{\sigma}\right) \Phi\left(\frac{\lambda}{\sigma} \varepsilon_i\right) \qquad (2.15)$$

(Kumbhakar and Lovell, 2000)

where φ(.)/$\Phi(.)$ are the density/cumulative distribution of the $N(0,1)$ and ε_i, σ and λ are given by Eq. (2.8), the square root of Eq. (2.10) and (2.11) respectively.

The density of the mixture model is

$$f(WTP^*;\theta) = \sum_{J=1}^{2} p_j f_j(WTP^*;\theta_j) = p_1 f_1(WTP^*;\theta_1) + p_2 f_2(WTP^*;\theta_2) \quad (2.16)$$

Where $\theta = (\alpha, \beta, \sigma_v^2, \sigma_u^2)$ and α refers to the constant term.

For a sample of n observations the likelihood function is

$$L = \prod_{i=1}^{n} f(WTP_i^*;\theta) = \prod_{i=1}^{n}\sum_{j=1}^{2} p_j f_j\left(WTP^*;\theta_j\right) \qquad (2.17)$$

and the log-likelihood function is given by

$$logL = \sum_{i=1}^{n} \log\left(\sum_{j=1}^{2} p_j f_j\left(WTP^*;\theta_j\right)\right) =$$

$$= \sum_{i=1}^{n} \log(p_1 f_1(WTP^*;\theta_1) + p_2 f_2(WTP^*;\theta_2)) \qquad (2.18)$$

The log-likelihood function is going to be maximized with respect to the unknown parameters, θ and p_1.

Furthermore, because mixture models present difficulties in the maximization process Eq. (2.18) is going to be estimated with the EM algorithm (Dempster et. al., 1977). The Maximum Likelihood Estimates for the mixing proportion for mixtures of Normals cannot be written in closed form. As a consequence these MLEs have to be computed iteratively while the EM algorithm greatly facilitates their computation (McLachlan and Peel, 2000:25).

In order to estimate a Mixture model the EM algorithm is going to be applied in the present thesis. The EM algorithm treats the estimation problem as a missing data problem, where the missing data is the information about class membership. Moreover, it consists of two steps, the E-step (expectation) and the M-step (maximization). Applying the EM algorithm to the mixture problem ensures monotonic increases of the likelihood values (McLachlan and Peel, 2000:48). Appendix A describes in detail how the EM algorithm works.

It should be noted that fitting mixture models has to deal with a number of issues such as the presence of multiple maxima in the mixture likelihood function and therefore the choice of starting values plays a very important role.

2.4. Simulations for open-ended mixture model

Simulations were conducted in order to test the validity of the proposed model for a number of different cases. For each different case 1000 replications were considered with samples of 1000 observations.

Additionally, with respect to the probability of class membership, the simulations consider two alternative scenarios. The first scenario considers the probability p_1 to be constant over respondents while the second scenario assumes that class membership depends on a variable z.

2.4.1. Data generation

In the data generation process of each case, the model that is going to be used is a simple regression model of the form,

$$WTP_i^* = a + \beta x_i + \omega_i \qquad (2.19)$$

given by one explanatory-independent variable $x_i \sim N(4,1)$ where the coefficient of x_i β is equal to 2 and the constant term α is equal to 5. Taking into consideration Eq. (2.13) and Eq. (2.19) the model becomes

$$WTP_i{}^* = \begin{cases} 5 + 2x_i + v_i & with\, probability\, p_1 \\ 5 + 2x_i + v_i + u_i & with\, probability\; p_2 \end{cases} \qquad (2.20)$$

Where $v_i \sim iid\, N(0, \sigma_v^2)$ and $u_i \sim iid\, N^+(0, \sigma_u^2)$.

For each case that was examined, different values were given to σ_v and σ_u. More analytically, Table 2.1 illustrates the values of σ_v, σ_u is given as a function of σ_v in Eq. (2.21) and λ is determined by Eq. (2.11).

$$\sigma_u = \begin{cases} 10\sigma_v \\ 5\sigma_v \\ 2\sigma_v \end{cases} \qquad (2.21)$$

Table 2.1: Values for σ_v

Values for σ_v		
$\sigma_v = 0.5$	$\sigma_v = 0.7$	$\sigma_v = 0.8$
$\sigma_v = 1$	$\sigma_v = 1.2$	$\sigma_v = 1.5$

Additionally the mean WTP at the mean value of x_i is given by Eq. (2.22) below

$$mean\, WTP = (5 + 2\bar{x}) = 12.994 \quad (2.22)$$

Finally, for the class membership probability p_1 (probability to belong in class 1 were respondents answer sincerely) and p_2 (probability to belong in class 2 were overstatement occurs) two cases were considered.

Case A: p_1 is a constant, equal to 0.75 and p_2 is equal to 0.25. In this case all respondents that belong in the same class have the same probability.

Case B: p_1 is no longer a constant, each respondent has a different probability to belong to class 1 since p_1 depends on a variable z.

More analytically, denoting by p^{**} an unobserved latent variable

$$p^{**} = d_1 + d_2 z_i + w_i \Rightarrow$$

$$p^{**} = 2 + 2z_i + w_i \qquad (2.23)$$

where $w_i \sim Logistic(0, 1)$ or standard logistic and $z_i \sim Normal(1, 4.84)$.

The probability that respondent i belongs in class 1 ($p^{**} > 0$) is given by

$$p_{1i} = \frac{1}{1 + e^{-(d_1 + d_2 z_i)}} \qquad (2.24)$$

Furthermore, in case B, since there is different probability p_{1i} for each respondent, the probability of class 1 can't be illustrated since it is practically very difficult to illustrate 1000 different probabilities, consequently in this case, the mean probability is computed and its value is given by

$$p_1^* = \frac{\sum_{i=1}^{1000} p_{1i}}{1000} = 0.80 \quad (2.25).$$

2.4.2. Starting values and Estimation Strategy

In the case of the open-ended method model we followed a number of steps and estimations in order to get the starting values for the model and especially for the EM algorithm. The estimates obtained by the EM algorithm were then used as starting values for the ML estimation of the mixture model, a similar procedure is followed for instance in Stata (StataCorp, 2021).

The above mentioned procedure can be decomposed in a number of steps that are described below, namely determining the starting values of the EM algorithm (Steps 1-3) and application of the EM algorithm and subsequent estimation by ML (Step 4).

Step 1: Random assignment of observations to two classes

The first step in obtaining starting values for the EM algorithm consisted in randomly assigning observations to the two classes. For this purpose, random draws from a Uniform(0,1) were generated and the observation was classified in the first class whenever the draw was below 0.5.

Step 2: Assigning an error distribution (normal/composed) to each class

In order to determine which model is represented from each group, ordinary least squares was applied to each group separately. Taking into account our previous observation in section 2.2 about the positive mean of the composed error term, the group with the bigger estimate of the constant term is assumed to be the class with overstatement (composed error model), while the other class is assumed to be the class where respondents answer sincerely.

More analytically, since the two classes have been determined, from each team's OLS (Ordinary Least Squares) estimation estimates for a and β were received. As it has been already mentioned, since $E(\varepsilon_i) = E(u_i) = \sigma_u\sqrt{\frac{2}{\pi}}$, ignoring the presence of hypothetical bias will lead to overestimation of the constant term of the equation. Taking under consideration this fact, it is expected that the class with the bigger estimate of a will represent the class with respondents that overstate their WTP.

From the OLS regression starting values for α, β and σ_v have been determined and consequently a starting value for σ_u is needed. More specifically, σ_u was computed by

$$\hat{\sigma}_u = \frac{\hat{a}_2 - \hat{\alpha}_1}{\sqrt{\frac{2}{\pi}}}$$

where $\hat{a}_2$ is the estimate of the constant term from OLS for the class where overestimation presumably occurs and $\hat{\alpha}_1$ is the estimate obtained by OLS regression for the class where no overestimation occurs.

From the clustering procedure the proportion of the number of the observations-respondents that belong in each class was calculated and this proportion was used as a starting value of the class membership probability.

Step 3: Estimating Eq. (2.19) with a composed error by ML

After the OLS estimation procedure was completed and the classes were determined, a ML estimation followed for the composed error model (class with overstatements of WTP) assuming that all respondents have overstate their WTP. This step provided starting values for EM mainly for the parameter λ.

More specifically the starting values of α and β were defined as the mean of the values of the OLS of both classes. The starting value of σ_v^2 was determined as the σ^2 estimate obtained from the OLS for the normal error model class and finally the starting value for λ was determined to be equal with the estimate of λ obtained by the ML estimation of the composed model.

Step 4: Application of EM algorithm

In this step, the EM algorithm was run for the mixture model until a tolerance criterion was reached. The tolerance criterion was until the log-likelihood, obtained for iteration k, satisfies Eq. (2.26).

$$|loglikelihood(k) - loglikelihood(k-1)| < 0.001 \qquad (2.26)$$

where k is the number of the iteration.

The estimates produced from the EM algorithm were used as starting values to maximize the log-likelihood of the proposed mixture model. This strategy is similar to the one implemented by Stata (StataCorp, 2021) whereas a few iterations of the EM algorithm are used to refine starting values for maximum likelihood.

Additionally, for case B where the probability is different for each respondent, the parameters d_1 and d_2 need starting values as well thus an additional procedure was added in step 4. More analytically, the extra procedure that was added was the estimation of a "logit" model for the class probability and the results were used as starting values for the parameters d_1 and d_2.

2.4.3. Simulation Results for the open-ended method model

For all cases the simulation results were obtained after 1000 replications. However, in some cases, there were problems for a number of replications since standard errors could not be computed. Therefore the tables that follow present the results for the successful replications and the number of failed replications are reported as well. More specifically the estimation results that are illustrated are obtained after the replications with the problem were removed.

For each parameter the tables report the mean estimates and the standard deviation. Furthermore, apart from the parameter estimates, we illustrate the bias of the mean WTP estimate for each case. More specifically the bias is given by the following equation:

$$Bias = expected\ value\ of\ estimated\ mean\ WTP - mean\ WTP \Rightarrow$$
$$Bias = \left(\hat{a} + \hat{\beta}\bar{x}\right) - (5 + 2\bar{x}) \qquad (2.27)$$

Where

$\hat{a} = \frac{\sum_{r=1}^{R} \hat{a}^r}{R}$, $\hat{\beta} = \frac{\sum_{r=1}^{R} \hat{\beta}^r}{R}$ and R is the number of replications.

The closer to zero the bias is for each case, it means that the estimates of the parameters are closer to the real values. It is very important to obtain very small values of bias since the major goal is to overcome hypothetical bias. The smaller the bias is, the more appropriate the model is in order to overcome hypothetical bias.

Case A. Fixed probability of class membership for all respondents (p_1=0.75)

In this subsection the results of 1000 replications are presented for the case where the class membership probability for class 1 (no overstatement occurs) is constant and equal to 0.75 and consequently 25% of the respondents overstate their WTP.

Table 2.2 illustrates the mean estimates of a, β, σ_v^2 and λ, the class probability was parametrized as in Eq. (2.28) below in order to ensure that the estimate lies in the open unit interval and in the estimation process the parameter $kappa$ is the parameter that was estimated.

$$\hat{p}_1 = \frac{1}{(1 + e^{\widehat{kappa}})} \qquad (2.28)$$

Furthermore, the bias of the mean WTP estimate of each case is illustrated in Table 2.4 and Table 2.3 illustrates the number of replications with breaking down issues that have been removed. More specifically, if in a specific replication a parameter's standard deviation was infinite or appeared as NaN (not a number), this replication was removed and the mean estimates were calculated from the remaining replications.

Table 2.2: Simulation results for open-ended and class probability p_1=0.75

Estimation Results for open-ended p_1=0.75					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^2$	$\hat{\lambda}$	$\hat{p}_1$
$\sigma_v = 0.5$ *and* $\lambda = 10$					
Mean	4.9888	2.0027	0.2483	10.0155	0.7496
Standard deviation	0.0783	0.019	0.0152	0.5424	
$\sigma_v = 0.5$ *and* $\lambda = 5$					
Mean	4.9814	2.0044	0.2479	5.0108	0.7491
Standard deviation	0.0778	0.0187	0.0164	0.2947	
$\sigma_v = 0.7$ *and* $\lambda = 10$					
Mean	4.9845	2.0037	0.4867	10.0173	0.7497
Standard deviation	0.1098	0.0267	0.0299	0.5416	
$\sigma_v = 0.7$ *and* $\lambda = 5$					
Mean	4.9742	2.0061	0.4859	5.011	0.7491
Standard deviation	0.1088	0.0262	0.0322	0.2947	
$\sigma_v = 0.7$ *and* $\lambda = 2$					
Mean	4.9733	2.008	0.502	1.1305	0.6974
Standard deviation	0.1478	0.0243	0.1085	0.9877	
$\sigma_v = 0.8$ *and* $\lambda = 10$					
Mean	4.9823	2.0043	0.6357	10.0172	0.7497
Standard deviation	0.1254	0.0305	0.039	0.5416	
$\sigma_v = 0.8$ *and* $\lambda = 5$					
Mean	4.9703	2.007	0.6346	5.0107	0.7491
Standard deviation	0.1245	0.03	0.0421	0.2947	

Table 2.2: (continued)

$\sigma_v = 0.8$ *and* $\lambda = 2$					
Mean	4.957	2.0087	0.6304	2.0179	0.7455
Standard deviation	0.13	0.028	0.0541	0.169	
$\sigma_v = 1$ *and* $\lambda = 10$					
Mean	4.9775	2.0054	0.9951	10.0315	0.7495
Standard deviation	0.1569	0.038	0.061	0.5417	
$\sigma_v = 1$ *and* $\lambda = 5$					
Mean	4.9628	2.0088	0.9916	5.0107	0.7491
Standard deviation	0.1556	0.0375	0.0657	0.2947	
$\sigma_v = 1$ *and* $\lambda = 2$					
Mean	4.9459	2.0109	0.9849	2.0173	0.7453
Standard deviation	0.1625	0.035	0.0844	0.1691	
$\sigma_v = 1.2$ *and* $\lambda = 10$					
Mean	4.9731	2.0065	1.4304	10.0171	0.7497
Standard deviation	0.1881	0.0457	0.0879	0.5416	
$\sigma_v = 1.2$ *and* $\lambda = 5$					
Mean	4.9554	2.0105	1.4278	5.0107	0.7491
Standard deviation	0.1868	0.0449	0.0946	0.2947	
$\sigma_v = 1.2$ *and* $\lambda = 2$					
Mean	4.9202	2.0131	1.4052	2.0280	0.719
Standard deviation	0.2336	0.0422	0.1198	0.1625	

Table 2.2: (continued)

$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 10}$					
Mean	4.9663	2.0081	2.2353	10.0167	0.7497
Standard deviation	0.2352	0.0571	0.1373	0.5418	
$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 5}$					
Mean	4.9442	2.0132	2.231	5.0107	0.7491
Standard deviation	0.2335	0.0562	0.1478	0.2947	

Table 2.2 shows that in all cases the mean estimates for all the parameters are very close to the real values (with one exception for the case where $\lambda = 2$ and $\sigma_v = 0.7$). Over all cases considered the bias for the estimate of α is less than 1.6% of the parameter value while for the estimate of β it is less than 0.7%.

More specifically, in cases where $\lambda = 10$ and 5 the mean estimates are very close to the real values for all parameters. Additionally in the case where $\lambda = 2$ and $\sigma_v = 0.7$ the mean estimate for p_1 and λ are not as close to the real values as in other cases. So does in the case where $\sigma_v = 1.2$ and $\lambda = 2$, the mean estimate for p_1 is not as close to the real value as in the other cases. To sum up, the probability estimate in the rest of the cases is very close to the real probability.

In general and as expected, for a given value of λ, both the bias and standard deviation of the estimates of α and β increase as the variance of the two-sided error term σ_v increases. On the other hand, for a given value of σ_v the bias and standard deviation of the mentioned parameter estimates decreases with λ.

Moreover, from the results presented in Table 2.3, in cases where $\lambda = 10$ and $\lambda = 5$ the replications that had to be removed were very few, especially for the cases where $\lambda = 5$ almost all of them had no replications with standard error issues. On the other hand though, in cases where $\lambda = 2$ for some cases the number of the replications removed were more than 500 replications thus for such cases the simulation results are not illustrated in Table 2.2 and 2.4. Such cases were for example the cases with $\sigma_v = 0.5$ and $\sigma_v = 1.5$ where in both cases $\lambda = 2$.

Kumbhakar et al. (2013:68) have pointed out that when $\lambda \rightarrow 0$ the identification of the model breaks down. This can explain the fact why in many cases where $\lambda = 2$ the program returned many replications with standard error issues.

Table 2.3: Replication removed for each case due to standard error issues

Number of replications removed			
	$\boldsymbol{\lambda = 10}$	$\boldsymbol{\lambda = 5}$	$\boldsymbol{\lambda = 2}$
$\boldsymbol{\sigma_v}$**=0.5**	2	0	-
$\boldsymbol{\sigma_v}$**=0.7**	2	1	374
$\boldsymbol{\sigma_v}$ **=0.8**	2	0	3
$\boldsymbol{\sigma_v}$**=1.0**	4	0	0
$\boldsymbol{\sigma_v}$**=1.2**	4	0	28
$\boldsymbol{\sigma_v}$**=1.5**	4	0	-

Finally Table 2.4 illustrates the bias of the estimated mean WTP as given by Eq. (2.27). Furthermore, Table 2.4 illustrates as well the bias of the mean WTP as it is given by Eq. (2.27) in the case where the model is estimated without considering the case of overestimation.

Table 2.4: Bias of Mean WTP for open-ended mixture model vs Bias of the Mean WTP for open-ended normal model (no Hypothetical bias considered) with class probability fixed

Bias of Mean WTP for Open-ended model (probability p_1=0.75)				
	$\sigma_v = 0.5$ & $\lambda = 10$	$\sigma_v = 0.5$ & $\lambda = 5$	$\sigma_v = 0.7$ & $\lambda = 10$	$\sigma_v = 0.7$& $\lambda = 5$
Mean $\widehat{WTP}_{mixture\,model}$	12.9936	12.993	12.9933	12.9926
$Bias_{mixture\,model}$	-0.0004	-0.001	-0.0007	-0.0014
Mean $\widehat{WTP}_{normal\,model}$	13.9922	13.4933	14.3915	13.693
$Bias_{normal\,model}$	0.9982	0.4993	1.3975	0.699
	$\sigma_v = 0.7$ & $\lambda = 2$	$\sigma_v = 0.8$ & $\lambda = 10$	$\sigma_v = 0.8$ & $\lambda = 5$	$\sigma_v = 0.8$ & $\lambda = 2$
Mean $\widehat{WTP}_{mixture\,model}$	12.9992	12.9935	12.9923	12.9858
$Bias_{mixture\,model}$	0.0053	-0.0005	-0.0017	-0.0082
Mean $\widehat{WTP}_{normal\,model}$	13.2739	14.5912	13.7928	13.3139
$Bias_{normal\,model}$	0.2799	1.5971	0.7988	0.3199
	$\sigma_v = 1$ & $\lambda = 10$	$\sigma_v = 1$ & $\lambda = 5$	$\sigma_v = 1$ & $\lambda = 2$	$\sigma_v = 1.2$ & $\lambda = 10$
Mean $\widehat{WTP}_{mixture\,model}$	12.993	12.992	12.9835	12.9931
$Bias_{mixture\,model}$	-0.0009	-0.002	-0.0105	-0.0009

Table 2.4: (continued)

Mean $\boldsymbol{W\hat{T}P_{normal\,model}}$	14.9904	13.9926	13.3938	15.3897
$\boldsymbol{Bias_{normal\,model}}$	1.9964	0.9985	0.3998	2.3957
	$\boldsymbol{\sigma_v = 1.2}$ & $\boldsymbol{\lambda = 5}$	$\boldsymbol{\sigma_v = 1.2}$ & $\boldsymbol{\lambda = 2}$	$\boldsymbol{\sigma_v = 1.5}$ & $\boldsymbol{\lambda = 10}$	$\boldsymbol{\sigma_v = 1.5}$ & $\boldsymbol{\lambda = 5}$
Mean $\boldsymbol{W\hat{T}P_{mixture\,model}}$	12.9914	12.9666	12.9927	12.991
$\boldsymbol{Bias_{mixture\,model}}$	-0.0026	-0.0274	-0.0013	-0.003
Mean $\boldsymbol{W\hat{T}P_{normal\,model}}$	14.1923	13.4738	15.9887	14.4918
$\boldsymbol{Bias_{normal\,model}}$	1.1983	0.4798	2.9946	1.4978

The bias obtained by the proposed model in all cases is very small and close to zero, with the maximum value of the bias being around 0.21% of the true value of mean WTP in only one case. The majority of the cases have bias less than the 0.1% of the true value of mean WTP. More analytically, in cases where $\sigma_v = 0.5$ and $\lambda = 10$ or 5, $\sigma_v = 0.7$, 0.8 or 1.2 and $\lambda = 10$ the bias is below 0.001. The biggest bias is in the case where $\sigma_v = 1.2$ and $\lambda = 2$ which is -0.0274 and on the other hand the smallest bias in the case where $\sigma_v = 0.5$ and $\lambda = 10$ which is -0.0004.

On the other hand, the bias obtained by the estimation of the normal model that doesn't take into account the possible existence of hypothetical bias is large in all cases. More specifically, as Table 2.4 shows, in the cases where $\lambda = 10$ or 5 the bias was larger since in these cases the overstatement was bigger. Furthermore, since the potential bias was not considered during the estimation process it was expected that the constant will be bigger and consequently the estimate of the mean WTP will be large. Overall the maximum value of the bias is 2.9946 and the smaller bias is 0.2799, obtained in the case where $\lambda = 2$ and $\sigma_v = 0.7$.

Overall, taking into consideration the above simulation results for open-ended format CV model with latent classes, with class membership probability p_1 equal to 0.75, it can be concluded that the model that is proposed in order to overcome hypothetical bias was able to fulfill its main goal.

Additionally, comparing the mean WTP estimate of the proposed model and the normal model where overstatement is ignored the gain of the proposed model is clear since the proposed model reduces the bias of the mean WTP estimate. The mean estimates of the parameters for almost all cases were very close to the real values, the class membership probability estimate was very close to the real probability and the bias of the mean WTP estimate was almost zero. However it should be noted that for the cases in which the parameter λ is 2, many replications would break down and the bias was considerably larger.

Case B. Probability depending on variable z

In case B the real class membership probability p_1 is no longer a fixed constant but varies between respondents. More analytically, the probability is determined by a variable z and is given by Eq. (2.24). It is quite possible that some underlying characteristics of respondents could affect their likelihood of overstating or understating their WTP, therefore it is more realistic to assume that the class membership probability varies over individuals.

In the present subsection, several cases are going to be analyzed in order to test how the model responds when each respondent has different class probability. Regarding the model structure two extra parameters (d_1 and d_2) have been included in the estimation process. Tables 2.5 and 2.6 report the results for the above scenario. Additionally, since the class membership probability differs among respondents, it is impossible for each case to illustrate 1000 mean probabilities, so Table 2.6 shows the mean estimates of parameters d_1, d_2 and the mean probability of all class membership probabilities given by Eq. (2.25). Furthermore Table 2.7 illustrates the number of replications that have been removed due to standard error issues and finally Table 2.8 presents the bias of the mean WTP estimate.

Table 2.5: Open ended model simulation results for probability depending on variable z (1): parameters $\alpha, \beta, \sigma_v^2$ and λ

Estimation Results for open-ended model and p_1 depending on z (1)				
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v{}^2$	$\hat{\lambda}$
$\sigma_v = 0.5$ *and* $\lambda = 10$				
Mean	5.0045	1.9989	0.2494	10.0288
Standard deviation	0.0773	0.0189	0.0131	0.5675
$\sigma_v = 0.5$ *and* $\lambda = 5$				
Mean	5.0014	1.9997	0.2492	5.0067
Standard deviation	0.075	0.0182	0.013	0.2999
$\sigma_v = 0.7$ *and* $\lambda = 10$				
Mean	5.0037	1.9991	0.4885	10.0265
Standard deviation	0.1055	0.0258	0.0261	0.5916
$\sigma_v = 0.7$ *and* $\lambda = 5$				
Mean	5.0058	1.9985	0.4886	5.0121
Standard deviation	0.1078	0.0263	0.0258	0.2992
$\sigma_v = 0.8$ *and* $\lambda = 10$				
Mean	5.005	1.9993	0.6389	10.03
Standard deviation	0.1114	0.0273	0.032	0.5649
$\sigma_v = 0.8$ *and* $\lambda = 5$				
Mean	5.0029	1.9995	0.6386	5.0123
Standard deviation	0.1199	0.028723	0.0325	0.2972
$\sigma_v = 1$ *and* $\lambda = 10$				
Mean	5.0051	1.9988	0.997	10.0265
Standard deviation	0.1508	0.0369	0.0533	0.592

Table 2.5: (continued)

		$\sigma_v = 1$ *and* $\lambda = 5$		
Mean	5.0065	1.9984	0.9968	5.0166
Standard deviation	0.15	0.0366	0.0545	0.3142
		$\sigma_v = 1$ *and* $\lambda = 2$		
Mean	5.0063	1.999	0.9994	2.0037
Standard deviation	0.1454	0.0351	0.0557	0.1642
		$\sigma_v = 1.2$ *and* $\lambda = 10$		
Mean	5.0027	1.9995	1.4363	10.0228
Standard deviation	0.1798	0.0431	0.0726	0.5637
		$\sigma_v = 1.2$ *and* $\lambda = 5$		
Mean	5.0077	1.9981	1.4355	5.0165
Standard deviation	0.18	0.0439	0.0785	0.3142
		$\sigma_v = 1.2$ *and* $\lambda = 2$		
Mean	5.0042	1.9992	1.4359	2.0083
Standard deviation	0.1755	0.0417	0.0754	0.1594
		$\sigma_v = 1.5$ *and* $\lambda = 10$		
Mean	5.0079	1.9981	2.2431	10.0265
Standard deviation	0.2262	0.0553	0.12	0.5916
		$\sigma_v = 1.5$ *and* $\lambda = 5$		
Mean	5.0097	1.9976	2.2429	5.0165
Standard deviation	0.225	0.0549	0.1227	0.3142
		$\sigma_v = 1.5$ *and* $\lambda = 5$		
Mean	5.0103	1.9984	2.2473	2.005
Standard deviation	0.2191	0.0529	0.1251	0.1629

Table 2.6: Open ended- probability depending on variable z (2): probability estimates

Estimation Results for open-ended and p_1 depending on z (2)			
	$\hat{d}_1$	$\hat{d}_2$	Mean $\hat{p}_1$
$\sigma_v = 0.5$ *and* $\lambda = 10$			
Mean	2.0074	2.0161	0.7968
Standard deviation	0.2038	0.1966	
$\sigma_v = 0.5$ *and* $\lambda = 5$			
Mean	2.0313	2.0355	0.7974
Standard deviation	0.247	0.2464	
$\sigma_v = 0.7$ *and* $\lambda = 10$			
Mean	2.0245	2.028	0.7973
Standard deviation	0.2083	0.2168	
$\sigma_v = 0.7$ *and* $\lambda = 5$			
Mean	2.0113	2.0262	0.7966
Standard deviation	0.2446	0.2422	
$\sigma_v = 0.8$ *and* $\lambda = 10$			
Mean	2.0171	2.0278	0.7969
Standard deviation	0.2094	0.207	
$\sigma_v = 0.8$ *and* $\lambda = 5$			
Mean	2.039	2.0466	0.7974
Standard deviation	0.2478	0.2668	
$\sigma_v = 1$ *and* $\lambda = 10$			
Mean	2.0249	2.0281	0.7974
Standard deviation	0.2081	0.2168	

Table 2.6: (Continued)

$\sigma_v = 1\ and\ \lambda = 5$			
Mean	2.0395	2.0442	0.7975
Standard deviation	0.2592	0.2739	
$\sigma_v = 1\ and\ \lambda = 2$			
Mean	2.1311	2.1564	0.7975
Standard deviation	0.4873	0.485458	
$\sigma_v = 1.2\ and\ \lambda = 10$			
Mean	2.025	2.0287	0.7974
Standard deviation	0.205	0.2148	
$\sigma_v = 1.2\ and\ \lambda = 5$			
Mean	2.0394	2.0443	0.7975
Standard deviation	0.2593	0.2739	
$\sigma_v = 1.2\ and\ \lambda = 2$			
Mean	2.1098	2.1395	0.7973
Standard deviation	0.4677	0.5214	
$\sigma_v = 1.5\ and\ \lambda = 10$			
Mean	2.0245	2.0279	0.7973
Standard deviation	0.2083	0.2168	
$\sigma_v = 1.5\ and\ \lambda = 5$			
Mean	2.0395	2.0444	0.7975
Standard deviation	0.2594	0.2739	
$\sigma_v = 1.5\ and\ \lambda = 2$			
Mean	2.1257	2.1505	0.7974
Standard deviation	0.4747	0.475	

As Table 2.5 illustrates, the mean estimates, after removing the replications with standard error issues, are very close to the real values for all cases. Additionally, Table 2.6 illustrates the mean estimates for the class membership parameters and the mean estimated probability. In all cases the mean estimates for d_1 and d_2 are very close to the given values. Each respondent has a different probability in each replication thus Eq. (2.24) and Eq. (2.25) were applied in order to receive an indicative p_1 from each replication. Table 2.6 shows the mean probability of all the indicative probabilities of all the replications.

Table 2.7 presents the number of replications that have been removed for each case. It can be seen that in cases where $\lambda = 10$ and $\lambda = 5$ the replications that had to be removed were very few, just like in the case where the class membership probability was the same for all respondents. On the other hand though, in cases where $\lambda = 2$ the majority of the cases are not presented at all due to many standard error issues.

Table 2.7: Number of replications with standard error problems

	Number of replications removed		
	$\boldsymbol{\lambda = 10}$	$\boldsymbol{\lambda = 5}$	$\boldsymbol{\lambda = 2}$
$\boldsymbol{\sigma_v}$=0.5	0	14	-
$\boldsymbol{\sigma_v}$=0.7	0	1	-
$\boldsymbol{\sigma_v}$=0.8	0	0	-
$\boldsymbol{\sigma_v}$=1.0	1	0	5
$\boldsymbol{\sigma_v}$=1.2	0	0	4
$\boldsymbol{\sigma_v}$=1.5	0	0	5

In order to explain the identification problem of the two classes in cases where $\lambda = 2$, the following Figures 2.1-2.3 illustrate graphically the densities for $\sigma_v^2 = 0.49$ and $\lambda = 10, 5$ and 2.

Figure 2.1: Density illustration of Normal error N(0, 0.49) and composed error ($\lambda = 10$)

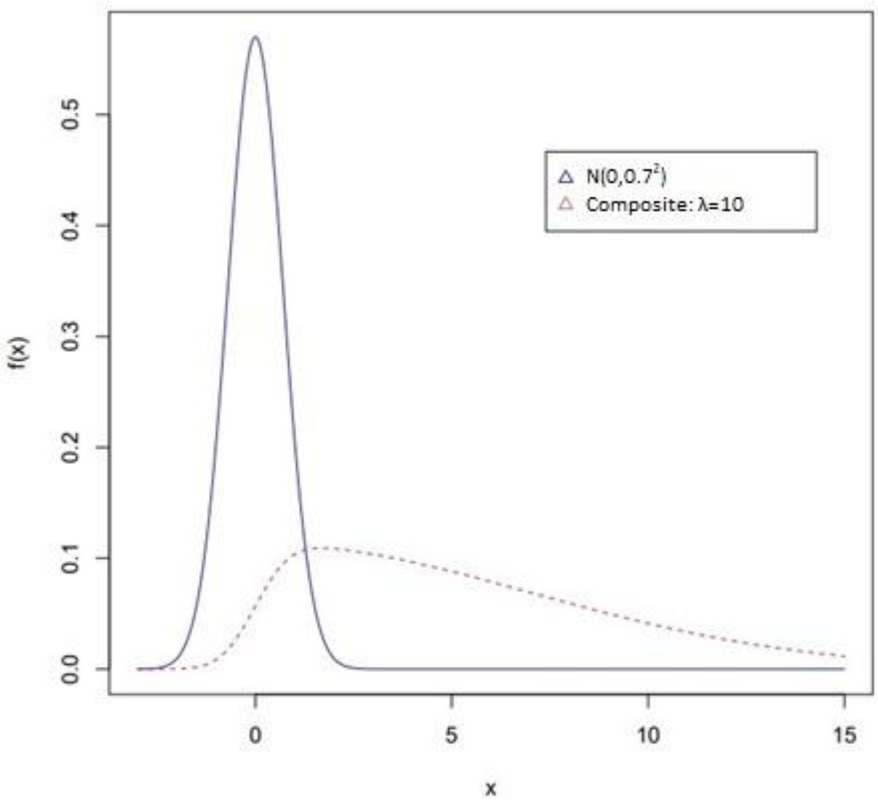
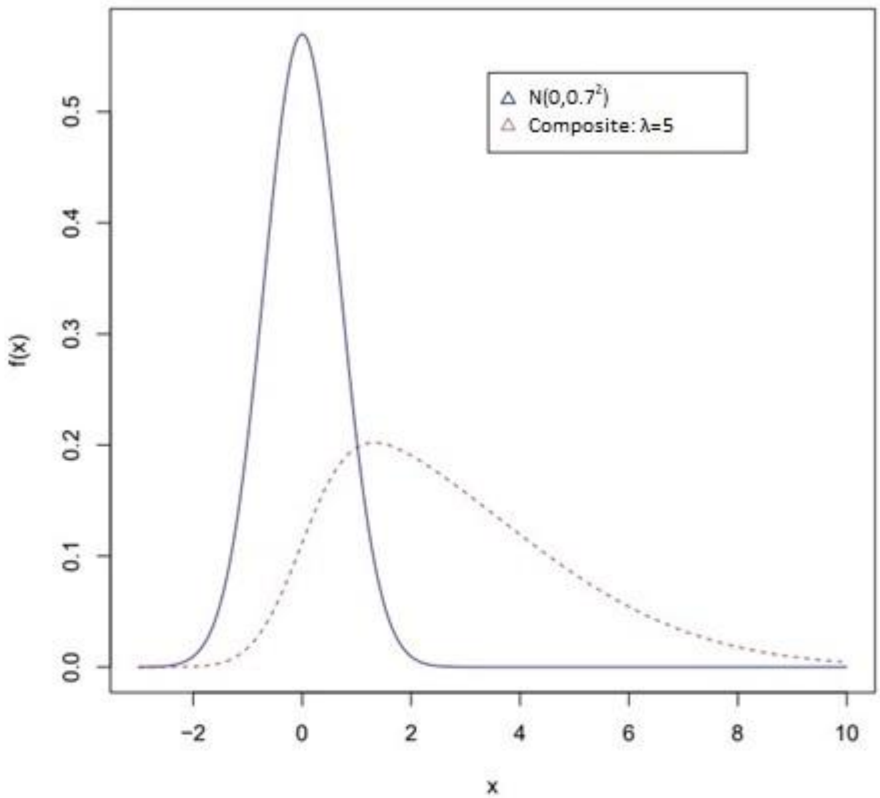


Figure 2.1, presents the illustration of the Normal error density and the Composed error density in the case where $\lambda = 10$. As it is shown the two densities in this case are easily distinguishable and thus the program was able to identify the two classes and return estimates close to the real values, without breaking down problems.

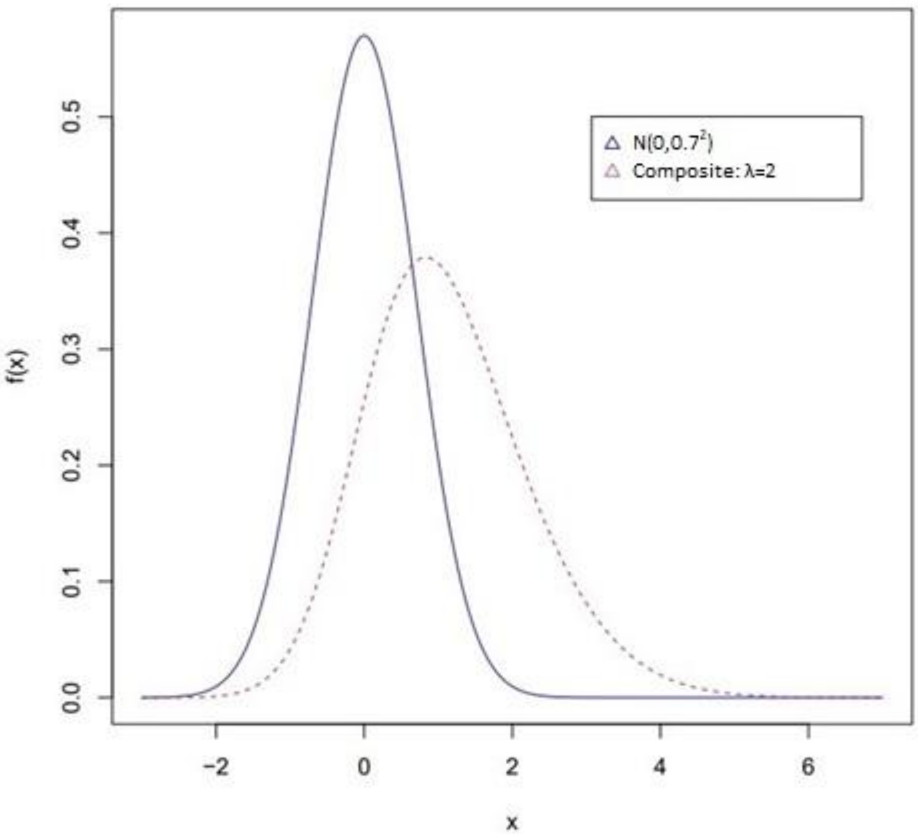
In Figure 2.2, the density of the normal error has remained the same but the density of the composed error is illustrated for $\lambda = 5$. It can be noticed that the two densities are still easily distinguishable but the observations of the composed density are gathered closer to the center. In this case the density of the composed error, compared to the previous case where $\lambda = 10$, is still easily distinct from the normal error density, but as it is shown, the shape of the composed error density is more symmetric than before.

Figure 2.2: Density illustration of Normal error N(0,0.49) and composed error (λ=5)



Additionally, Figure 2.3, shows the normal error density compared to the composed error in the case where $\lambda = 2$. As λ decreases the overlap of the two densities in the region where the normal density is not close to zero increases and the composed error density in less skewed. In this case the two densities are getting to look very similar so in such cases the program deals with serious issues in order to identify for each respondent in which class he/she belongs.

Figure 2.3: Density illustration of Normal error N(0, 0.49) and composed error (λ=2)



Overall it can be concluded that as σ_u is getting bigger, so does the mean of the composed error as well, $E(v+u) = 0 + \sigma_u \frac{\sqrt{2}}{\sqrt{\pi}}$ and skewness increases while the center of the density moves to the right and possibly the two densities are more easily distinguished. On the other hand, in cases where σ_u doesn't have an outstanding difference compared to σ_v, the model won't have the ability to detect which class refers to each respondents so the algorithm has identification issues. Such cases are the ones where $\lambda = 2$ and this explains why many cases with $\lambda = 2$ had identification problems.

Finally, Table 2.8 illustrates the biases for the mean WTP estimate for each case as given by Eq. (2.27).

.

Table 2.8: Bias of Mean WTP for open-ended and probability depending on variable z

Bias of Mean WTP for open-ended elicitation format model (class probability depending on variable z)			
	$\sigma_v = 0.5\ and\ \lambda = 10$	$\sigma_v = 0.5\ and\ \lambda = 5$	$\sigma_v = 0.7\ and\ \lambda = 10$
Mean $\widehat{WTP}_{mixture\ model}$	12.9941	12.9942	12.9941
$Bias_{mixrure\ model}$	0.0001	0.0002	0.0001
Mean $\widehat{WTP}_{normal\ model}$	13.9922	13.4933	14.3915
$Bias_{normal\ model}$	0.9982	0.4993	1.3975
	$\sigma_v = 0.7\ and\ \lambda = 5$	$\sigma_v = 0.8\ and\ \lambda = 10$	$\sigma_v = 0.8\ and\ \lambda = 5$
Mean $\widehat{WTP}_{mixture\ model}$	12.9938	12.9962	12.9949
$Bias_{mixture\ model}$	-0.0002	0.0022	0.0009
Mean $\widehat{WTP}_{normal\ model}$	13.693	14.5912	13.7928
$Bias_{normal\ model}$	0.699	1.5971	0.7988
	$\sigma_v = 1\ and\ \lambda = 10$	$\sigma_v = 1\ and\ \lambda = 5$	$\sigma_v = 1\ and\ \lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	12.9943	12.9941	12.9963
$Bias_{mixture\ model}$	0.0003	0.0001	0.0023
Mean $\widehat{WTP}_{normal\ model}$	14.9904	13.9926	13.3938
$Bias_{normal\ model}$	1.9964	0.9985	0.3998

Table 2.8: (continued)

	$\sigma_v = 1.2$ *and* $\lambda = 10$	$\sigma_v = 1.2$ *and* $\lambda = 5$	$\sigma_v = 1.2$ *and* $\lambda = 2$
Mean $W\hat{T}P_{mixture\ model}$	12.9947	12.9941	12.995
$Bias_{mixture\ model}$	0.0007	0.0001	0.001
Mean $W\hat{T}P_{normal\ model}$	15.3897	14.1923	13.4738
$Bias_{normal\ model}$	2.3957	1.1983	0.4798
	$\sigma_v = 1.5$ *and* $\lambda = 10$	$\sigma_v = 1.5$ *and* $\lambda = 5$	$\sigma_v = 1.5$ *and* $\lambda = 2$
Mean $W\hat{T}P_{mixture\ model}$	12.9943	12.9941	12.9979
$Bias_{mixture\ model}$	0.0003	0.0001	0.0039
Mean $W\hat{T}P_{normal\ model}$	15.9887	14.4918	13.5937
$Bias_{normal\ model}$	2.9946	1.4978	0.5997

As is illustrated in Table 2.8, the bias in all cases is very small, below 0.1 and more specifically the bias is very close to 0 for all cases which means that the mean estimates are almost the same with the true value.

More analytically, in the majority of the cases the bias is below 0.001which represents a very small percentage of mean WTP value. The biggest bias is in case where $\sigma_v = 1.5$ and $\lambda = 2$ which is 0.0039 and on the other hand the smallest bias is 0.0001 which this value is the bias for several cases such as $\sigma_v = 0.5$ and $\lambda = 10$ and more.

Additionally, the bias of the mean WTP for the model where overstatement is ignored is larger. More specifically, the biggest bias is in the case where $\sigma_v = 1.5$ and $\lambda = 10$

which is 2.9946, in this case the bias from the proposed model is 0.0003 thus it is concluded that the gain of the proposed model is clear.

Overall, taking into consideration the above simulation results for a composed open-ended elicitation format model with class membership probability p_1 different for each respondent, it can be concluded that the model that is proposed in order to overcome hypothetical bias was able to fulfill its main goal in most of the experiments.

Conclusions

In this chapter a mixture open-ended model is proposed as a way to tackle the problem caused by yea-saying behavior. The simulation results are quite encouraging under both scenarios of either constant mixture weights or individual dependent mixture weights. The proposed model is effective in dealing with Hypothetical Bias and can provide unbiased estimates of mean WTP when CV data include yea-saying behavior.

Additionally, comparing the simulation results when OLS is applied to all responses when hypothetical bias is ignored to the simulation results obtained by the proposed model it can be noticed that the proposed model reduces significantly the bias of mean WTP. It can be noticed that when hypothetical bias is not taken into account the mean WTP estimate has higher values and thus there is a problem of overestimation. By applying the proposed model the mean WTP is no longer overestimated and the estimates are more reliable since hypothetical bias has been taken into account in the econometric model. Furthermore, it should be noted that as the parameter λ gets smaller the program had problems identifying the two classes and consequently more breaking down issues occurred.

At this point, the first evidence show how the mixture model works in order to overcome hypothetical bias have been gathered and the next step in the following chapter is to apply the stochastic frontier model in a double-bounded DC model in order to expand the model to a more popular elicitation method.

Appendix A

EM algorithm

The EM algorithm treats the estimation problem as a missing data problem, where the missing data is the information about class membership. More analytically and following McLachlan and Peel (2000), the EM algorithm treats the observed data vector $y_1, \dots, y_n$ as incomplete since the component-label vectors $z_1, \dots, z_n$ are not available. The component-label vectors are taken to be realized values of the random vectors $Z_1, \dots, Z_n$ where we assume they are distributed unconditionally as $Z_1, \dots, Z_n \overset{iid}{\sim} Mult_g(1, \pi)$. The j-th mixing proportion π_j is the prior probability that the entity belongs to the j-th component of the mixture. On the other hand, the posterior probability that the entity belongs to the j-th component with y_i having been observed is given by

$$\tau_j(y_i) = pr(entity \in jth\ component\ |y_i) = pr(Z_{ji} = 1 | y_i) = \pi_j f_j(y_i)/f(y_i)$$

$(j = 1, \dots, g; i = 1, \dots, n)$(McLachlan and Peel, 2000:19-20).

The E-step handles the posterior probability and in the M-step since E-step replaces the unobservable part z_{ji} with the current conditional expectation $\tau_j(y_i)$ we can obtain the updated estimate of π_j (McLachlan and Peel, 2000:49-50).

In a few words, a finite mixture maximizes the likelihood (McLachlan et al., 2019:360) and the ML estimate of Ψ, $(\hat{\Psi})$ is given by a proper root of the likelihood equation,

$$\partial \log L(\Psi)/\partial \Psi = 0$$

$$\log L(\Psi) = \sum_{i=1}^{n} log\ f y_i; \Psi) =$$

$$\sum_{i=1}^{n} \log\{\sum_{j=1}^{g} \pi_j\ f_j(y_i; \theta_j)\}$$

Additionally, a g-dimensional vector z_i is being fitted in the likelihood function, where $z_{ji} = (z_i)_j = 1\ or\ 0$ according to if y_i is or isn't arisen from the j-th component of the mixture, $1, \dots, g\ and\ i = 1, \dots, n$.

The complete data log likelihood function for Ψ is

$$\log L_c(\Psi) = \sum_{j=1}^{g}\sum_{i=1}^{n} z_{ji}\,[\log \pi_j + \log f_j(y_i;\theta_j)]$$

(McLachlan and Peel, 2000:48).

In the E-step in the EM algorithm we take the conditional expectation of $\log L_c(\Psi)$ given y and $\Psi^{(0)}$ is used for Ψ.

$$Q\left(\Psi;\Psi^{(0)}\right) = E_{\Psi(0)}[log\ L_c\ (\Psi)|y]$$

On the $(k+1)$ iteration $Q\left(\Psi;\Psi^{(k)}\right)$ is required on the E-step and $\Psi^{(k)}$ is the value of Ψ after the k-th iteration of the EM. The E-step provides the quantity $\tau_j\left(y_i;\Psi^{(k)}\right)$

$$\tau_j\left(y_i;\Psi^{(k)}\right) = {\pi_j}^k f_j(y_i;{\theta_j}^{(k)}) / \sum_{h=1}^{g} {\pi^{(k)}}_h f_h(y_i;{\theta_h}^{(k)})$$

which is the posterior probability that the i-th part of the sample with value y_i belongs to the j-th component of the mixture.

In the M-step the equation that is maximized is

$$Q\left(\Psi;\Psi^{(k)}\right) = \sum_{j=1}^{g}\sum_{i=1}^{n} \tau_j\left(y_i;\Psi^{(k)}\right)[log\,\pi_j + log\,f_j(y_i;\theta_j)]$$

Where $\widehat{\pi_j}$ will be given by

$$\widehat{\pi_j} = \sum_{i=1}^{n} z_{ji}/n$$

and by replacing z_{ji} by $\tau_j\left(y_i;\Psi^{(k)}\right)$ we have

$\widehat{\pi_j^{(k+1)}} = \sum_{i=1}^{n} \tau_j\left(y_i;\Psi^{(k)}\right)/n$ (McLachlan and Peel, 2000:48-50).

References of Chapter 2

Ahmed, S. U. and Gotoh, K., 2006, *Cost-Benefit Analysis of Environmental Goods by Applying the Contingent Valuation Method, Some Japanese Case Studies,* Japan, Springer-Verlag Tokyo

Arrow, K., Solow, R., Portney P.R., Leamer, E.E., Radner, R., and Schuman, H., 1993, *Report of NOAA Panel on Contingent Valuation,* National Oceanic and Atmospheric Administration

Bateman, I.J., Langford, I.H., Turner, R.K., Willis, K.G. and Garrod, G.D., 1995, "Elicitation and truncation effects in contingent valuation studies", *Ecological Economics,* vol. 12, p.161-179

Carson, R.T., 2000, "Contingent Valuation: A User's Guide", *Environmental Science and Technology,* vol. 34, no. 8, p. 1413-1418

Carson, R.T. and Hanemann, W.M., 2005, Contingent Valuation. In: Mäler K.G. and Vincent J.R., eds., 2005, *Handbook of Environmental Economics,* vol.2, Elsevier B.V., chapter 17, p.821-936

Dempster, A.P., Laird, N.M. and Rubin, D.B., 1977, "Maximun Likelihood from Incomplete Data via the EM Algorythm", *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 39, no 1, p.1-38

Håkansson, C., 2008, "Anew valuation question: analysis of and insights from interval open-ended data in contingent valuation", *Environmental Resource Economics,* vol. 39, p. 175-188

Hofler, R.A. and List, J.A., 2004, "Valuation on the frontier: Calibrating actual and hypothetical statements of value", *American Journal of Agricultural Economics,*vol.86, no.1, p.213-221

Istamto, T.,Houthuijs, D. and Lebret,E., 2014, "Multi-country willingness to pay study on road traffic environmental health effects: are people willing and able to provide a number?", *Environmental Health,* p.13:35, available at: https://ehjournal.biomedcentral.com/articles/10.1186/1476-069X-13-35 [access 10/5/2022]

Kealy, M. J. and Turner, R. W., 1993, "A Test of The Equalityof Closed-Ended and Open-Ended Contingent Valuation", *American Agricultural Economics Association,* vol. 75, p.321-331

Kumbhakar, S.C. and Lovell, K.C.A., 2000, *Stochastic Frontier Analysis*, United States of America, Cambridge University Press

Kumbhakar, S.C., Parmeter C.F. and Tsionas E.G., 2013, "A zero inefficiency stochastic frontier model", *Journal of Econometrics,* vol.172, p.66-76

Loomis, J. B., 1990, “Comparative Reliability of the Dichotomous Choice and Open ended Contingent Valuation Techniques”, *Journal of Environmental Economics and Management,* vol. 18, p.78-85

Mc Lachlan G. and Peel D.,2000, *Finite Mixture Models*, United States of America, A Wiley-Interscience Publication

McLachlan, G.J., Lee, S.X. and Rathnayake, S.I., 2019, “Finite Mixtures”, ,*Annual Review of Statistics and Its Application,* vol.6, p. 355-378

Chapter 3
Applying stochastic frontier and mixture models to contingent valuation under the double-bounded dichotomous choice format

Introduction

In Chapter 3 a stochastic frontier CV under the double-bounded dichotomous choice format is going to be applied in order to overcome hypothetical bias. The main difference from the open-ended format is that a number of bids are given to respondents and they have to answer with a Yes or No regarding their WTP.

The chapter is organized as follows. The first section presents the theoretical background of the CV method under the double-bounded elicitation format. The second section analyses the stochastic frontier model of the double-bounded format with the presence of hypothetical bias and the third part analyses the double-bounded mixture model that it is suggested in order to overcome hypothetical bias. The fourth section of the chapter presents simulations that took place intending to test the proposed model and finally the last section of this chapter, presents an empirical application with real CV data, in order to investigate if hypothetical bias exists and how the proposed model can overcome the bias.

Considering the fourth section, it describes in detail the data generation process, the initialization strategy that it was followed and the simulation results for a number of different cases. Additionally, the simulations took place for two different probability determination cases. In the first case the class probability is a constant and consequently all respondents have the same probability for overstating their WTP and in the other case the probability of overstating WTP can differ among respondents since the probability depends on a variable z.

3.1. Contingent valuation and double-bounded dichotomous choice elicitation method

In a CV study the main goal is to measure in monetary terms for each individual an item q. Each person has a utility function $u(x,q)$which is defined by a number of commodities x and the item q. Each person has an indirect utility function $v(p,q,y)$ where p are the market prices of the commodities and y is the income. It is important to compare if after a change in q each individual will be better off or worse off. For example:

If the change in q represents an improvement $u^1 > u^0$.

If the change represents a worsening $u^1 < u^0$.

If the change if indifferent for the agent $u^1 = u^0$.

Where u^0, u^1 are the utility levels associated to q^0 and q^1 for a given x. The change in monetary terms is represented by the Compensating variation C which satisfies Eq. (3.1) and the equivalent variation E which satisfies Eq. (3.2).

$$v(p,q^1,y-C) = v(p,q^0,y) \qquad (3.1)$$

$$v(p,q^1,y) = v(p,q^0,y+E) \qquad (3.2)$$

In the case where the change is an improvement $C > 0$ and $E > 0$, where C is the person's maximum WTP in order to ensure that the change will be implemented, additionally E is the minimum WTA. On the other hand, in the case where the change leads to being worse off $C < 0$ and $E < 0$. In this case, C is the person's WTA and respectively E measures the WTP in order to avoid the change (Carson and Hanemann, 2005:844-45).

One elicitation method in order to evaluate the WTP or WTA that individuals have for a change in q is the double-bounded dichotomous choice format. The double-bounded question format is a closed-ended format. More analytically, the respondents are asked, "Would you vote to support the change from q^0 to q^1 if it would cost you \$A?" The respondent will answer "yes" if his value C (his WTP) is at least A (Carson and Hanemann, 2005:848).

The double-bounded format is a referendum format. The NOOA Panel (Arrow et al., 1993:21) proposed that researchers should use a referendum format because it has many advantages. First of all it is more realistic because in the provision of public goods referenda are common. Another advantage of the double-bounded elicitation format is that respondents can answer without high mental demands thus there are less non-responses by the end of the survey.

Furthermore, the question format matches the market setting where the price is stated and the individuals are price-takers and consequently they decide if they are going to buy or not at the given price. Additionally, the double-bounded format is referred to as an "incentive compatible device" in the sense that respondents would reveal their true preferences in regard to the provision of a good (Loomis, 1990:79).

On the other hand, the double-bounded format has a number of disadvantages, one disadvantage is that the estimates could be sensitive to distributional assumptions and to the functional form of the utility function (Loomis, 1990:79). Furthermore, a major disadvantage of the double-bounded format is that there is a possibility that the respondents are influenced from the first offer and consequently they tend to accept the follow-up offer (Ahmed and Gotoh, 2006:16).

In CV surveys the respondents are asked to reveal their preference about a given scenario by answering with a Yes or No to the question of whether they are willing to pay a certain amount (the bid). The double-bounded format asks the respondents twice and the second question-bid depends on the answer of the first question.

By applying the double-bounded approach is implicitly assumed that "the respondent's answers to both of the payment questions are driven by one underlying WTP value" (Alberini, 1995:297). Additionally, if this assumption holds it means that the information regarding the true WTP is increased by the second discrete choice question because a tighter interval around the true WTP has been created and therefore there is a gain in efficiency with respect to the single-bound elicitation format.

The maximum willingness to pay for individual i is given below

$$WTP_i = \beta' x_i + v_i \quad i = 1, \cdots, n \qquad (3.3)$$

$$where\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_\kappa \end{bmatrix}, x_i = [x_{i1} \quad \cdots \quad x_{i\kappa}]' \text{ and } v_i \sim N(0, \sigma_v^2)$$

and x_i is a vector of observed explanatory variables that affect WTP and $x_{i1} \equiv 1 \forall i$.

The density function for the error is the following:

$$f_1(v_i) = \frac{1}{\sigma_v \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v_i}{\sigma_v}\right)^2} = \frac{1}{\sigma_v \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{wtp_i - \beta' x_i}{\sigma_v}\right)^2} \qquad (3.4)$$

Each respondent has to answer two successive bids whereas the second bid depends on the answer that is given for the first bid. Firstly, the respondents have to answer the first bid with a yes or no. Denoting as y_1 and y_2 the responses to the two bids respectively

$$y_{1i} = \begin{cases} 1 & yes\ to\ bid_{1i} \\ 0 & no\ to\ bid_{1i} \end{cases}, \qquad i = 1, \cdots, n \qquad (3.5)$$

If yes to $bid_{1i}, y_{1i} = 1$ and $bid_{2i} > bid_{1i},\ y_{2i} = \begin{cases} 1\ for\ yes\ to\ bid_{2i} \\ 0\ for\ no\ to\ bid_{2i} \end{cases}$ (3.6)

If no to $bid_{1i}, y_{1i} = 0$ and $bid_{2i} < bid_{1i}, y_{2i} = \begin{cases} 1\ for\ yes\ to\ bid_{2i} \\ 0\ for\ no\ to\ bid_{2i} \end{cases}$ (3.7)

In order to receive yes-yes as an answer for both bids the individual's WTP must satisfy the that $WTP_i > bid_{1i}$ and $WTP_i > bid_{2i}$ so we have that $WTP_i > bid_{2i} > bid_{1i}$. A yes answer to the first bid and no to the second bid, means that the individual's WTP is greater than bid_{1i} but smaller than bid_{2i}, $WTP_i > bid_{1i}$ and $WTP_i < bid_{2i}$ so we have that $bid_{1i} < WTP_i < bid_{2i}$.

On the other hand, for a no-yes answer to bid_{1i} and bid_{2i} respectively, the respondent's WTP must me smaller than bid_{1i} and greater than bid_{2i}, $WTP_i < bid_{1i}$ and $WTP_i > bid_{2i}$ which means that $bid_{2i} < WTP_i < bid_{1i}$. Finally, in order to receive a no answer in both bids, the respondent's WTP is smaller than both bid_{1i} and bid_{2i}, $WTP_i < bid_{1i}$ and $WTP_i < bid_{2i}$ so we have that $WTP_i < bid_{2i} < bid_{1i}$.

For yes-yes

$WTP_i > bid_{1i}\ and\ WTP_i > bid_{2i}$ so we have that $WTP_i > bid_{2i} \ > bid_{1i}$

$$P(WTP_i \geq bid_{2i}) = P(v_i \geq bid_{2i} - \beta' x_i) =$$

$$P\left(\frac{v_i}{\sigma_v} \geq \frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) = 1 - P\left(\frac{v_i}{\sigma_v} \leq \frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)$$

in terms of the standard normal cumulative distribution it can be written as

$$1 - \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) \qquad (3.8)$$

For yes-no

$WTP_i > bid_{1i}\ and\ WTP_i < bidi_2$ so we have that $bid_{1i} < WTP_i < bid_{2i}$

$$P(bid_{1i} \leq WTP_i \leq bid_{2i}) =$$

$$P((bid_{1i} - \beta' x_i) \leq v_i \leq (bid_{2i} - \beta' x_i))$$

$$P\left(\frac{v_i}{\sigma_v} \leq \frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) - P\left(\frac{v_i}{\sigma_v} \geq \frac{bid_{1i} - \beta' x_i}{\sigma_v}\right)$$

in terms of the standard normal cumulative distribution it can be written as

$$\Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) - \Phi\left(\frac{bid_{1i} - \beta' x_i}{\sigma_v}\right) \qquad (3.9)$$

For no-yes

$WTP_i < bid_{1i}\ and\ WTP_i > bid_{2i}$ so we have that $bid_{2i} < WTP_i < bid_{1i}$

$$P(bid_{2i} \leq WTP_i \leq bid_{1i}) =$$

$$P((bid_{2i} - \beta' x_i) \leq v_i \leq (bid_{1i} - \beta' x_i))=$$

$$P\left(\frac{v_i}{\sigma_v} \leq \frac{bid_{1i} - \beta' x_i}{\sigma_v}\right) - P\left(\frac{v_i}{\sigma_v} \geq \frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)$$

in terms of the standard normal cumulative distribution it can be written as

$$\Phi\left(\frac{bid_{1i} - \beta' x_i}{\sigma_v}\right) - \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) \qquad (3.10)$$

<u>For no-no</u>

$WTP_i < bid_{1i}$ and $WTP_i < bid_{2i}$ so we have that $WTP_i < bid_{2i} < bid_{1i}$

$$P(WTP_i \leq bid_{2i}) = P(v_i \leq bid_{2i} - \beta' x_i) = P\left(\frac{v_i}{\sigma_v} \leq \frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)$$

in terms of the standard normal cumulative distribution it can be written as

$$\Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) \qquad (3.11)$$

The log-likelihood function is given by:

$$\ln L = \sum_{i=1}^{n} y_{1i} y_{2i} \ln\left(1 - \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)\right) + (1 - y_{1i})(1 - y_{2i}) \ln \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)$$
$$+ y_{1i}(1 - y_{2i}) \ln\left(\Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right) - \Phi\left(\frac{bid_{1i} - \beta' x_i}{\sigma_v}\right)\right)$$
$$+ (1 - y_{1i}) y_{2i}\, ln\left(\Phi\left(\frac{bid_{1i} - \beta' x_i}{\sigma_v}\right) - \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)\right) \qquad (3.12)$$

If respondents answer truthfully to the double-bounded questions, i.e under the absence of perception and strategic errors, then the mean willingness to pay (MWTP) is given by $\beta' \bar{x}$ and an estimate is given by $\hat{\beta}' \bar{x}$. Under the presence of hypothetical bias, respondents might overstate their bids (yea-saying behavior). In such case the model given by Eq. (3.3) does not hold anymore and a stochastic frontier model can be used to reflect this behavior. The next section analyses the stochastic frontier model for the double-bounded DC elicitation format under the presence of hypothetical bias.

3.2. Stochastic frontier model for the double-bounded DC method with the presence of hypothetical bias

Hofler and List (2004) proposed the use of the stochastic frontier model as a way to take into consideration the difference between the real and the hypothetical auction bid. Additionally, Chien et al. (2005:362) proposed “a general model that addresses the starting point bias in the dichotomous choice evaluation data by incorporating both the anchoring effect and yes-saying bias". More specifically, Chien et al. (2005:365) applied stochastic frontier model by adding a composed error in their model in order to include the fact that yea-saying tendency will increase a respondent's WTP. In a few words, in the presence of yea-saying behavior, stated WTP can be modeled as "true" WTP augmented with a one-sided nonnegative error term as below

$$WTP_i^* = WTP_i + u_i \qquad (3.13)$$

or

$$WTP_i^* = \beta' x_i + \varepsilon_i \qquad (3.14)$$

where $u_i \sim iid\ N^+(0, \sigma_u^2)$ nonnegative Half Normal which gives rise to the so called Normal-Half Normal Model in the stochastic frontier literature.

More analytically, Eq. (3.13) is

$$WTP^*{}_i = \beta_1 + \beta_2 x_{i2} + \ldots\ldots + \beta_\kappa x_{i\kappa} + v_i + u_i \qquad (3.15)$$

and the composed error is given by

$$\varepsilon_i = v_i + u_i \qquad (3.16)$$

Note that the composed error doesn't have a zero mean, since $E(\varepsilon_i) = E(u_i) = \sigma_u \sqrt{\frac{2}{\pi}}$, (Kumbhakar and Lovell, 2000), thus ignoring the presence of hypothetical bias will lead to overestimation of the constant term of equation (3.3). In the case where hypothetical bias exists in the form of understatement of WTP the stochastic frontier model becomes

$$WTP_i^* = WTP_i - u_i \qquad (3.17)$$

From Eq. (3.3) and Eq. (3.13) follows that

$$WTP_i^* - WTP_i = u_i \qquad (3.18)$$

When the error term u_i approaches zero, the gap between the real and hypothetical values is decreased and the hypothetical values → true values (Hofler and List, 2004:216). Additionally, the model is parameterized very often in terms of the two parameters defined below

$$\sigma^2 = \sigma_v^2 + \sigma_u^2 \qquad (3.19)$$

and

$$\lambda = \frac{\sigma_u}{\sigma_v} \qquad (3.20)$$

If overestimation occurs then the parameter λ should be statistically significant and greater than zero. If λ approaches values close to zero, σ_u approaches values close to zero as well and the composed error tends to be equal to v . For Normal $(v_i \sim iid\ N(0, \sigma_v^2))$ and Half-Normal $\left(u_i \sim iid\ N^+(0, \sigma_u^2)\right)$ distributions we have that the composed error density distribution:

$$f_2(\varepsilon_i) = \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right) \qquad (3.21)$$

(Kumbhakar and Lovell, 2000; Chien et al., 2005:366)

where φ(.)/$\Phi(.)$ are the density/cumulative distribution of the N(0,1) and ε_i, σ and λ are given by Eq. (3.18), the square root of Eq.(3.21) and (3.22) respectively.

In the case where hypothetical bias exists, Eq. (3.5), Eq. (3.6) and Eq. (3.7) hold but Eqs. (3.8)-(3.11) need to be modified as now the model includes a composed error term and its cumulative distribution needs to be used instead of the cumulative normal. The cumulative distribution of the composed error term involves integrating the expression in Eq. (3.21) which is non-trivial. Tsay et al., (2013) proposed a closed-form approximation for the cumulative distribution function of a normal-half normal composed error that represents the integral

$$F = \int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right) d\varepsilon_i \qquad (3.22)$$

For any number Q, which in our case is $Q = bid_{ji} - \beta' x_i$. Therefore $\Phi(\,)$ is replaced by the closed form approximation denoted by F_a proposed by Tsay et al., (2013:261) when deriving Eqs. (3.8)-(3.11).

Additionally, Amsler et al. (2019) present another approximation for the cumulative distribution function of a normal-half normal composed error. They proposed a new simulation based method and they compare their method with Tsay's et al. (2013) approximation. They found that in non-extreme values the two approximations are quite close but for extreme values they claim that their approximation is more accurate (Amsler et al., 2019:32). The new simulation based method for the cdf is not applied in the present thesis because it was not published yet by the time the present dissertation was starting.

The log-likelihood given in Eq. (3.12) is modified and is given as

$$\ln L = \sum_{i=1}^{n} y_{1i} y_{2i} \ln\big(1 - F_a(bid_{2i} \ - \beta' x_i)\big) + (1 - y_{1i})(1 - y_{2i}) \ln F_a(bid_{2i} \ - \beta' x_i)$$
$$+ y_{1i}(1 - y_{2i}) \ln\big(F_a(bid_{2i} \ - \beta' x_i) - F_a(bid_{1i} \ - \beta' x_i)\big)$$
$$+ \ (1 - y_{1i}) y_{2i}\, ln\big(F_a(bid_{1i} \ - \beta' x_i) - F_a(bid_{2i} \ - \beta' x_i)\big) \qquad (3.23)$$

In the case where hypothetical bias exists in the form of underestimation, which means that individuals respond more no-no, the approximation of the cumulative distribution function of a normal-half normal composed error differs.

More analytically, in the case where underestimation occurs, Eq. (3.17) holds and the density of the composed error term is not given by Eq. (3.22) but by Eq. (3.24) below (Kumbhakar and Lovell, 2000).

$$f_2'(\varepsilon_i) = \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(-\frac{\lambda}{\sigma}\varepsilon_i\right) \qquad (3.24)$$

The integral of Eq. (3.24) represents the cdf of the composed error in the case of underestimation. In order to present the cumulative distribution of the composed error term in this case, some algebra shows it can be expressed as a function of *F* in Eq. (3.22) as follows

$$\int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(-\frac{\lambda}{\sigma}\varepsilon_i\right) d\varepsilon_i = \int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\left(1 - \Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right)\right) d\varepsilon_i$$

$$= \ \int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right) d\varepsilon_i - \int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right) d\varepsilon_i$$

$$= \int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right) d\varepsilon_i - F = 2\Phi\left(\frac{Q}{\sigma}\right) - F$$

The last equality is obtained by performing a change of variable in the integral of the left hand side. The closed-form approximation of the composed error in the case of underestimation is given by the expression below,

$$2\Phi\left(\frac{Q}{\sigma}\right) - F_a$$

3.3. Double-bounded mixture model

Hypothetical bias may occur if the respondents do not answer honestly, consequently a gap between the real WTP and the estimated hypothetical WTP is created. Although it is possible that a number of respondents might answer insincerely, a number of respondents might reveal their true WTP. It cannot be considered that all responders are overstating their WTP because some responders might actually answer sincerely. In cases where this heterogeneity occurs in regard to individuals behavior a latent class model or a mixture model is proposed in order to estimate the WTP.

The present consideration follows the same idea that Kumbhakar, Parmeter and Tsionas (2013:67) followed in their paper related to the productive inefficiency of firms. In a few words, they stated that in a sample both efficient and inefficient firms can exist with a probability. Taking into account the finite mixture models theory, the model for WTP will be considered as a mixture of two classes. Class 1 has no hypothetical bias and respondents answer sincerely thus a model with a Normal distributed error holds and class 2 overstates WTP so a model with a composed Normal-Half-normal error term exists. The two classes are:

Class 1: respondents answer sincerely and WTP is given by Eq. (3.3) and

Class 2: respondents overstate their WTP and therefore Eq. (3.15) holds.

The probability of belonging to class 1 and class 2 is noted as p_1 and $p_2 = (1 - p_1)$ respectively.

For the present case with the two classes described above, the model becomes

$$WTP_i^* = \begin{cases} \beta' x_i + v_i & with\ probability\ p_1 \\ \beta' x_i + v_i + u_i & with\ probability\ p_2 \end{cases} \qquad (3.25)$$

And the log-likelihood function is given by

$$\ln L = \sum_{i=1}^{n} ln\big[y_{1i}y_{2i}\big(p_1(1-f_{12}) + p_2(1-f_{22})\big)$$
$$+ y_{1i}(1-y_{2i})\big(p_1(f_{12}-f_{11}) + p_2(f_{22}-f_{21})\big)$$
$$+ (1-y_{1i})y_{2i}\big(p_1(f_{11}-f_{12}) + p_2(f_{21}-f_{22})\big)$$
$$+ (1-y_{1i})(1-y_{2i})\ (p_1 f_{12} + p_2 f_{22})\big] \qquad (3.26)$$

Where $f_{11} = \Phi\left(\frac{bid_{1i} - \beta' x_i}{\sigma_v}\right), f_{12} = \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)$,

$f_{21} = F_a(bid_{1i} - \beta' x_i)$ and $f_{22} = F_a(bid_{2i} - \beta' x_i)$.

The log-likelihood function is maximized with respect to the unknown parameters, $\alpha, \beta, \sigma_v^2, \lambda$ and p_1. Furthermore, because mixture models have difficulties in the maximization process Eq. (3.26) is going to be estimated with EM algorithm.

3.4. Simulations for the double-bounded DC mixture model

In order to test the validity of the model a number of simulations are presented for several cases. For each case 1000 observations were generated and 1000 replications took place. The procedure was followed for two different cases according to the determination of the class membership probability. The first case assumes that the class membership probability p_1 is a constant while the second scenario assumes that the class membership probability differs among respondents since it depends on a variable z.

3.4.1. Data generation

The model that is going to be employed is a simple regression model given in Eq. (3.27) with one explanatory-independent variable $x_i \sim N$ (4,1) where the coefficient of x_i, β is equal to 2 and the constant term α is equal to 5.

$$WTP_i^* = a + \beta x_i + \omega_i \qquad (3.27)$$

Combining Eq. (3.25) and Eq. (3.27) the model becomes

$$WTP_i{}^* = \begin{cases} 5 + 2x_i + v_i & with\, probability\, p_1 \\ 5 + 2x_i + v_i + u_i & with\, probability\ \ p_2 \end{cases} \qquad (3.28)$$

Where $v_i \sim iid\ N(0, \sigma_v^2)$ and $u_i \sim iid\ N^+(0, \sigma_u^2)$.

For each case σ_v and σ_u is determined as follows, the values of σ_v are 0.7 0.9, 1, 1.25 and 1.5 and σ_u is defined by a function of σ_v given in Eq. (3.29).

$$\sigma_u = \begin{cases} 10\sigma_v \\ 5\sigma_v \\ 2\sigma_v \end{cases} \qquad (3.29)$$

The parameter λ is determined by Eq. (3.20) and the bids that are given to respondents are 11, 13, 14 and 15. Table 3.1 shows analytically the structure of the bids, bid_1 and bid_2 which are determined according Eq. (3.30).

$$bid_2 = \begin{cases} 1.25bid_1 & if\ yes\ in\ bid_1 \\ 0.75bid_1 & if\ \ no\ in\ bid_1 \end{cases} \qquad (3.30)$$

Table 3.1: bid_1 and bid_2

Bids given to respondents							
$\boldsymbol{bid_1}$ **=11**		$\boldsymbol{bid_1}$ **=12**		$\boldsymbol{bid_1}$ **=13**		$\boldsymbol{bid_1}$ **=15**	
Yes	No	Yes	No	Yes	No	Yes	No
$\boldsymbol{bid_2}$ **=14**	$\boldsymbol{bid_2}$ **=9**	$\boldsymbol{bid_2}$ **=15**	$\boldsymbol{bid_2}$ **=9**	$\boldsymbol{bid_2}$ **=17**	$\boldsymbol{bid_2}$ **=10**	$\boldsymbol{bid_2}$ **=19**	$\boldsymbol{bid_2}$ **=12**

The mean WTP at the mean value of x_i is given by Eq. (3.31) below

$$mean\ WTP = (2 + 5\bar{x}) = 12.994 \qquad (3.31)$$

and finally two cases were considered in regard to the class membership probability p_1 (probability to belong in class 1 were respondents answer sincerely) and p_2 (probability to belong in class 2 were overstatement occurs).

Case A: the probability is constant $p_1 = 0.75$ and $p_2 = 0.25$. In this case all respondents that belong to the same class have the same probability.

Case B: p_1 is no longer a constant, but depends on a variable z that varies across respondents. Therefore each respondent has a different probability to belong to a given class. More analytically, denoting by p^{**} an unobserved latent variable

$$p^{**} = d_1 + d_2 z_i + w_i \Rightarrow$$

$$p^{**} = 2 + 2z_i + w_i \qquad (3.32)$$

where $w_i \sim Logistic(0{,}1)$ or standard logistic and $z_i \sim Normal(1, 4.84)$.

The probability that respondent i belongs to class 1 ($p^{**} > 0$) is given by

$$p_{1i} = 1/(1 + e^{-(d_1 + d_2 z_i)}) \qquad (3.33)$$

In case B, since each respondent has different probability p_{1i}, the probability of each class cannot be illustrated since it is practically very difficult to illustrate 1000 different probabilities. In this case, the mean probability is given by Eq. (3.34).

$$p_1^* = \frac{\sum_{i=1}^{1000} p_{1i}}{1000} = 0.80 \qquad (3.34)$$

3.4.2. Starting values and estimation strategy

Choosing initial values is a very crucial step in the estimation of mixture models, nevertheless some preliminary results were obtained using the EM algorithm with arbitrary initial values. This procedure aimed to investigate how the model responds and it was very helpful in order to define the procedure of estimating the starting values.

Given starting values

The mixture model was estimated by the EM algorithm with starting values defined manually for the parameters $\alpha, \beta, \sigma_v^2$ and λ for the case where the class membership p_1 is a constant and equal to 0.75. The starting value for σ_v^2 was chosen to be around the double of the real value and values for σ_{u}^2 were chosen so that the resulting value of λ deviated from the real value by no more than 60% for most cases. Table 3.2 illustrates the starting values that were given for the estimation process.

Table 3.2: Starting values determined manually

Starting values						
	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\sigma_v^2}$	$\boldsymbol{\sigma_u^2}$	λ	p_1
$\boldsymbol{\sigma_v = 1\ and\ \lambda = 10}$						
Starting values	4	3	2	120	7.74	0.5
$\boldsymbol{\sigma_v = 1\ and\ \lambda = 5}$						
Starting values	3	3	2	100	7.071	0.5
$\boldsymbol{\sigma_v = 1\ and\ \lambda = 2}$						
Starting values	3	3	2	80	6.32455	0.5
$\boldsymbol{\sigma_v = 3\ and\ \lambda = 10}$						
Starting values	3	3	20	1200	7.745967	0.5
$\boldsymbol{\sigma_v = 3\ and\ \lambda = 5}$						
Starting values	3	3	25	1000	6.32455	0.5
$\boldsymbol{\sigma_v = 3\ and\ \lambda = 2}$						
Starting values	3	3	20	1000	3.162278	0.5
$\boldsymbol{\sigma_v = 5\ and\ \lambda = 10}$						
Starting values	3	3	50	3000	7.7459	0.5
$\boldsymbol{\sigma_v = 5\ and\ \lambda = 5}$						
Starting values	3	3	50	2000	6.32455	0.5

Starting values from random assignment

The steps for determining starting values for the double-bounded method model under random assignment are very similar as the steps described is section 2.4.2 of Chapter 2. Because of the nature of the double-bounded method, a few changes were implemented to the procedure. The starting values were determined by steps described below:

Step 1: Random assignment of observations to two classes

The first step to determine starting values for the EM algorithm includes random assignment of the observations into two classes. To achieve this task, random draws from a Uniform(0,1) were generated and each observation was classified in the first class whenever the draw was below 0.5.

Step 2: Assigning an error distribution (normal/composed) to each class

Next for each class a "probit" model was estimated and the class with the biggest constant was assigned the model with overestimation. These estimates of both probits were used as starting values in order to run a ML estimation as if all respondents had answered honestly (Normal error model) and afterwards as if all respondents had overstated (Composed error model). By the end of this stage were determined starting values for the EM estimation.

Step 3: Application of EM algorithm

The EM algorithm was run for the mixture model until a tolerance criterion was reached. The tolerance criterion was that the EM algorithm would continue iterating until the log-likelihood obtained at iteration k satisfied Eq. (3.35).

$$|loglikelihood(k) - loglikelihood(k-1)| < 0.001 \qquad (3.35)$$

where k is the number of the iteration.

As soon as the EM algorithm estimation was completed, the estimates were used as starting values for the ML estimation of the proposed model. The idea of using EM algorithm estimates as starting values of ML estimation is a procedure that Stata implements (StataCorp, 2021).

Additionally in the case where the probability differs among respondents, it is necessary to determine starting values for the parameters d_1 and d_2.To this effect a "logit" model for the probability was added in Step 3.

3.4.3. Simulation results for the double-bounded method model

This subsection presents the simulation results for the double-bounded elicitation method mixture model for several cases. At first are presented the EM algorithm estimates with starting values defined in Table 3.2. Secondly, are illustrated the simulation results of the final proposed mixture model using random assignment to determine initial values. As mentioned before 1000 replications were applied for a number of cases, however in some cases some iterations had problems with the computation of standard errors. Therefore the tables that follow present the results after the replications with such problems were removed. Additionally, the number of the replications that failed in computing standard errors is reported as well.

The tables below report for each parameter the mean estimates and their standard deviation, as well as the bias of the mean WTP estimate for each case. More specifically the bias is given by

$$Bias = expected\ value\ of\ estimated\ mean\ WTP - mean\ WTP \Rightarrow$$

$$Bias = \left(\hat{a} + \hat{\beta}\bar{x}\right) - (5 + 2\bar{x}) \qquad (3.36)$$

where

$\hat{a} = \frac{\sum_{r=1}^{R} \hat{\alpha}^r}{R}$, $\hat{\beta} = \frac{\sum_{r=1}^{R} \hat{\beta}^r}{R}$ and R is the number of replications.

The desired outcome is to receive values for bias close to zero. The closest to zero the bias is the closest to the real values are the estimates of the parameters and consequently the proposed model is more suitable in order to overcome hypothetical bias.

Results when starting values are chosen arbitrarily

Table 3.3: Double-Bounded DC Simulation results for EM algorithm and starting values given manually

EM Estimation Results for Double-Bounded DC1000 replications						
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^2$	$\hat{\lambda}$	Mean $\widehat{WTP}$	$\hat{p}_{1_{EM}}$
$\sigma_v = 1$ *and* $\lambda = 10$						
Mean	4.9771	2.012	1.0241	10.5977	13.0193	0.7621
Standard deviation	0.2446	0.0628	0.1302	0.9448		
$\sigma_v = 1$ *and* $\lambda = 5$						
Mean	5.1366	2.0143	1.2751	7.1476	13.1878	0.8489
Standard deviation	0.2396	0.0606	0.1599	0.6192		
$\sigma_v = 1$ *and* $\lambda = 2$						
Mean	5.2679	2.0114	1.4312	5.3147	13.3077	0.9778
Standard deviation	0.2202	0.0555	0.1397	0.5777		
$\sigma_v = 3$ *and* $\lambda = 10$						
Mean	4.8991	2.0392	9.1876	11.4233	13.0496	0.7637
Standard deviation	0.5499	0.1357	1.0133	0.6946		
$\sigma_v = 3$ *and* $\lambda = 5$						
Mean	5.0888	2.0359	9.7301	10.0233	13.2263	0.8146
Standard deviation	0.5337	0.1315	1.0777	0.6009		
$\sigma_v = 3$ *and* $\lambda = 2$						
Mean	5.3157	2.0315	10.2207	4.3492	13.4354	0.9045
Standard deviation	0.5154	0.123	1.0709	0.2496		
$\sigma_v = 5$ *and* $\lambda = 10$						
Mean	4.7964	2.0633	24.9755	10.9938	13.0433	0.7599
Standard deviation	0.9202	0.2152	4.19	0.8793		

Table 3.3: (continued)

$\sigma_v = 5$ *and* $\lambda = 5$						
Mean	5.0033	2.0572	25.695	8.8546	13.2257	0.7973
Standard deviation	0.9339	0.2113	4.4129	0.7142		

Table 3.3 shows that in almost all cases the mean estimate of the constant is close to the real value, so does the mean estimates for $\hat{\beta}$ and $\hat{\sigma}_v^2$. For the rest of the parameters such as $\hat{\lambda}$ the mean estimates when $\lambda = 10$ are very close to the true values but in the cases where $\lambda = 5$ and $\lambda = 2$ the mean estimates are not very close. Additionally in the cases where $\sigma_v = 1$ and $\lambda = 10$ or 5, $\sigma_v = 3$ and $\lambda = 10$ or 5 and finally in the case where $\sigma_v = 5$ and $\lambda = 5$, no standard error issues occurred in none of the 1000 replications.

In the case where $\sigma_v = 1$ and $\lambda = 2$, $\sigma_v = 3$ and $\lambda = 2$ and in case where $\sigma_v = 5$ and $\lambda = 10$ a number of replications were removed. The number of replications removed was 16, 6 and 304 respectively. Furthermore, in all cases the standard deviation of the parameters is small except for $\hat{\sigma}_v^2$'s standard deviation in the cases where $\sigma_v = 5$. Additionally, Table 3.4 presents the mean Bias of the estimated mean WTP for each case given by Eq. (3.36). In all cases the bias is very small and especially in cases where $\lambda = 10$ the bias is even smaller than 0.1. Overall the case where $\sigma_v = 1$ and $\lambda = 10$ has the smallest bias and the case where $\sigma_v = 5$ and $\lambda = 10$ has the second smallest bias.

Table 3.4: Mean Bias of WTP for EM algorithm estimation and given starting values

Bias of Mean WTP for EM algorithm estimation and given starting values			
$\sigma_v = 1$ *and* $\lambda = 10$	$\sigma_v = 1$ *and* $\lambda = 5$	$\sigma_v = 1$ *and* $\lambda = 2$	$\sigma_v = 3$ *and* $\lambda = 10$
0.0253	0.1938	0.3136	0.0556
$\sigma_v = 3$ *and* $\lambda = 5$	$\sigma_v = 3$ *and* $\lambda = 2$	$\sigma_v = 5$ *and* $\lambda = 10$	$\sigma_v = 5$ *and* $\lambda = 5$
0.2323	0.4414	0.0493	0.2317

Finally, taking into consideration results in Table 3.3 and Table 3.4 it can be concluded that cases where λ is not large, the mean estimates aren't very close to the given starting values.

Summarizing, these first findings of how the double-bounded model responds when it is estimated with the EM algorithm with arbitrary starting values led the present research towards two directions. Firstly, the implementation of the STATA estimation procedure with the additional ML estimation by using EM estimates as starting values and secondly the determination of proper starting values became a very important task.

Results when starting values are obtained from random assignment

The following subsection presents the simulation results for the proposed model with starting values determined within the estimation process after random assignment. The simulation results are presented for the two different cases in regards of the probability determination.

Case A: Probability p_1 is fixed and equal to 0.75

Tables 3.5, 3.6 and 3.7 illustrate the simulation results for the case where the probability of overstating their WTP is equal for all respondents. More analytically Table 3.5 presents the mean estimates and the standard deviation of the parameters a, β, σ_v^2 and λ after removing any replications with standard error issues in a total of 1000 replications. Additionally, the class probability was parameterized as in Eq. (3.37) below in order to ensure that the estimate lies in the open unit interval and the parameter that was estimated during the estimation process was *kappa*.

$$\hat{p}_1 = \frac{1}{(1 + e^{\widehat{kappa}})} \qquad (3.37)$$

Table 3.6 illustrates the numbers of replications that have been removed in each case. More specifically, the replications that have been removed had for a parameter infinite standard deviation or it appeared as NaN (not a number). Finally Table 3.7 shows the bias of the mean WTP.

Table 3.5: Simulation results for Double-Bounded DC model and constant probability

Estimation Results for Double-Bounded DC and p_1=0.75 1000 replications					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^2$	$\hat{\lambda}$	$\hat{p}_1$
$\sigma_v = 0.7$ and $\lambda = 10$					
Mean	4.9785	2.0077	0.5117	10.2506	0.751
Standard deviation	0.209	0.0505	0.3055	1.7784	
$\sigma_v = 0.7$ and $\lambda = 5$					
Mean	4.967	2.0082	0.4899	5.0805	0.7479
Standard deviation	0.2017	0.0488	0.1331	0.6738	
$\sigma_v = 0.7$ and $\lambda = 2$					
Mean	4.9198	2.0071	0.4642	2.2268	0.6951
Standard deviation	0.2672	0.045	0.1239	0.5134	
$\sigma_v = 0.9$ and $\lambda = 10$					
Mean	4.9641	2.0113	0.8286	9.8666	0.7519
Standard deviation	0.2407	0.0593	0.3583	2.0512	
$\sigma_v = 0.9$ and $\lambda = 5$					
Mean	4.9582	2.0119	0.818	4.9026	0.7514
Standard deviation	0.2319	0.0569	0.2258	0.808	
$\sigma_v = 0.9$ and $\lambda = 2$					
Mean	4.9033	2.0102	0.7732	2.0605	0.7214
Standard deviation	0.3015	0.0517	0.1841	0.4829	
$\sigma_v = 1$ and $\lambda = 10$					
Mean	4.9636	2.0112	1.0144	10.3519	0.7508
Standard deviation	0.2574	0.0631	0.3377	2.0818	
$\sigma_v = 1 and\ \lambda = 5$					

Table 3.5: (continued)

Mean	4.9534	2.0126	0.9956	5.0857	0.7537
Standard deviation	0.2435	0.0604	0.1473	0.7168	
$\sigma_v = 1\ and\ \lambda = 2$					
Mean	4.8937	2.0108	0.9545	2.1075	0.7213
Standard deviation	0.3224	0.0553	0.2131	0.4388	
$\sigma_v = 1.25\ and\ \lambda = 10$					
Mean	4.9445	2.0146	1.5588	10.4779	0.7507
Standard deviation	0.2829	0.0712	0.2774	2.5891	
$\sigma_v = 1.25\ and\ \lambda = 5$					
Mean	4.9517	2.0145	1.5801	5.1063	0.7522
Standard deviation	0.2897	0.0705	0.3711	1.1316	
$\sigma_v = 1.25\ and\ \lambda = 2$					
Mean	4.8811	2.0121	1.5025	2.0698	0.7198
Standard deviation	0.3798	0.0648	0.2965	0.4539	
$\sigma_v = 1.5\ and\ \lambda = 10$					
Mean	4.9436	2.0156	2.2519	10.8362	0.75071
Standard deviation	0.3266	0.0799	0.3976	3.5602	
$\sigma_v = 1.5\ and\ \lambda = 5$					
Mean	4.9309	2.0177	2.2518	5.2014	0.7472
Standard deviation	0.33	0.0785	0.4282	1.2087	
$\sigma_v = 1.5\ and\ \lambda = 2$					
Mean	4.8572	2.0158	2.1703	2.1716	0.7032
Standard deviation	0.4271	0.0728	0.411	0.568	

Table 3.6: Number of replication removed due to standard error issues for the DB DC model for p_1 fixed

Number of replications removed for p_1 fixed			
	$\lambda = 10$	$\lambda = 5$	$\lambda = 2$
$\sigma_v = 0.7$	33	29	73
$\sigma_v = 0.9$	24	18	58
$\sigma_v = 1$	47	18	59
$\sigma_v = 1.25$	67	31	47
$\sigma_v = 1.5$	67	40	80

Table 3.5 shows that in the majority of the cases the estimates are very close to the true values. More analytically, in the cases where $\sigma_v = 0.9$, 1 and 1.25 the mean estimates are very close to the true values for all parameters, including the class membership probability. The standard deviation of all parameters is small except for parameter λ in the case where $\sigma_v = 1.25$ and $\lambda = 10$. Considering Table 3.6 for these cases the number of replications removed because of standard error issues are from 1.8%-8% . Additionally, it can be noticed that when $\lambda = 2$, $\hat{\alpha}$ and $\hat{p}_1$ although they are close to the true values, their estimates are not as close as in cases with $\lambda = 10$ or 5.

To continue with the analysis, Table 3.7 illustrates the bias of the estimate of mean WTP for each case, given by Eq. (3.36). As it can be noticed, the bias is very small and very close to 0 for all cases, which means that the mean estimates are almost the same as the true values. The biggest bias is in the case where $\sigma_v = 1.5$ and $\lambda = 2$ which is -0.0796 and the smallest bias is in the case where $\sigma_v = 0.7$ and $\lambda = 5$ which is -0.0002.

Additionally, Table 3.7 presents the simulation results of the double-bounded model for the case where hypothetical bias is ignored. In all cases the estimate of the mean WTP is bigger than the mean estimate of WTP of the proposed model. As it was expected, when hypothetical bias is not taken into consideration during the estimation process the results reveal that the bias is positive in the presence of yea-saying.

Table 3.7: Bias of mean WTP for DB DC proposed model vs Bias of the mean WTP for DB model (no Hypothetical bias considered) (fixed probability)

Bias of Mean WTP same probability for overestimation for all respondents			
	$\sigma_v = 0.7$ *and* $\lambda = 10$	$\sigma_v = 0.7$ *and* $\lambda = 5$	$\sigma_v = 0.7$ *and* $\lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	13.0033	12.9938	12.9422
$Bias_{mixture\ model}$	0.0093	-0.0002	-0.0518
Mean $\widehat{WTP}_{normal\ model}$	13.9789	13.645	13.2841
$Bias_{normal\ model}$	0.9849	0.651	0.29
	$\sigma_v = 0.9$ *and* $\lambda = 10$	$\sigma_v = 0.9$ *and* $\lambda = 5$	$\sigma_v = 0.9$ *and* $\lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	13.0033	12.9998	12.9381
$Bias_{mixture\ model}$	0.0093	-0.0058	-0.05593
Mean $\widehat{WTP}_{normal\ model}$	14.0848	13.7612	13.3569
$Bias_{normal\ model}$	1.0908	0.7672	0.3629
	$\sigma_v = 1$ *and* $\lambda = 10$	$\sigma_v = 1$ *and* $\lambda = 5$	$\sigma_v = 1$ *and* $\lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	13.0024	12.9978	12.9309
$Bias_{mixture\ model}$	0.0084	0.0038	-0.0631
Mean $\widehat{WTP}_{normal\ model}$	14.1264	13.8106	13.3906

Table 3.7: (continued)

$\boldsymbol{Bias_{normal\,model}}$	1.1324	0.8166	0.3966
	$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 10}$	$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 5}$	$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 2}$
Mean $\boldsymbol{\widehat{WTP}_{mixture\,model}}$	12.9969	13.0037	12.9235
$\boldsymbol{Bias_{mixture\,model}}$	0.0029	0.0097	-0.0705
Mean $\boldsymbol{\widehat{WTP}_{normal\,model}}$	14.211	13.9165	13.4705
$\boldsymbol{Bias_{normal\,model}}$	1.2169	0.9224	0.4765
	$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 10}$	$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 5}$	$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 2}$
Mean $\boldsymbol{\widehat{WTP}_{mixture\,model}}$	13	12.9957	12.9144
$\boldsymbol{Bias_{mixture\,model}}$	0.006	0.0017	-0.0796
Mean $\boldsymbol{\widehat{WTP}_{normal\,model}}$	14.2802	14.005	13.5459
$\boldsymbol{Bias_{normal\,model}}$	1.2861	1.011	0.5519

Taking into consideration the simulation results illustrated in Table 3.5, 3.6 and 3.7 it can be concluded that the model is effective in accomplishing its main goal to overcome hypothetical bias and return unbiased estimates.

Case B: Probability p_1 is different for each respondent

The present subsection analyses the simulation results for the double-bounded DC model in the case where the probability is not fixed but each respondent has a different probability of overstating WTP. Since in this case each respondent's probability is determined by a variable z, two additional parameters are added within the estimation process.

The Tables 3.8-3.11 illustrate the simulation results for a number of cases. More analytically, Table 3.8 and 3.9 present the mean estimates of the parameters, including the parameters related to the class probability, Table 3.10 contains the number of replications that were removed from each case and finally, Table 3.11 presents the bias of the mean WTP for each case.

Table 3.8: Double-Bounded DC model simulation results for probability depending on variable z (1): parameters $\alpha, \beta, \sigma_v^2$ and λ

Estimation Results for Double-Bounded DC model and p_1 depending on z(1)				
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^{2}$	$\hat{\lambda}$
$\sigma_v = 0.95$ and $\lambda = 10$				
Mean	4.9906	2.0024	0.9014	10.307
Standard deviation	0.2185	0.0551	0.0881	1.7271
$\sigma_v = 0.95$ and $\lambda = 5$				
Mean	4.9967	2.0008	0.8981	5.054
Standard deviation	0.213	0.0538	0.1101	0.634
$\sigma_v = 0.95$ and $\lambda = 2$				
Mean	5.006	1.9985	0.8937	2.0303
Standard deviation	0.208	0.0526	0.0885	0.2753
$\sigma_v = 1$ and $\lambda = 10$				
Mean	4.9875	2.0032	0.9962	10.3362
Standard deviation	0.228	0.0576	0.0979	1.7702
$\sigma_v = 1$ and $\lambda = 5$				

Table 3.8: (continued)

Mean	4.9983	2.0004	0.9924	5.0751
Standard deviation	0.2235	0.0563	0.0976	0.6278
$\sigma_v = 1$ and $\lambda = 2$				
Mean	5.0085	1.998	0.9913	2.0237
Standard deviation	0.2156	0.0547	0.0954	0.2682
$\sigma_v = 1.25$ and $\lambda = 10$				
Mean	4.9925	2.0022	1.5652	10.4184
Standard deviation	0.2623	0.066	0.137	2.1611
$\sigma_v = 1.25$ and $\lambda = 5$				
Mean	5.0027	1.9992	1.5551	5.0919
Standard deviation	0.2565	0.0642	0.1357	0.7094
$\sigma_v = 1.25$ and $\lambda = 2$				
Mean	5.0018	1.9989	1.5508	2.0194
Standard deviation	0.248	0.062	0.1398	0.2779
$\sigma_v = 1.5$ and $\lambda = 10$				
Mean	5.0011	1.9998	2.2403	10.4926
Standard deviation	0.2967	0.0743	0.1771	2.4955
$\sigma_v = 1.5$ and $\lambda = 5$				
Mean	5.0016	1.9997	2.2387	5.1095
Standard deviation	0.2976	0.0743	0.179	0.8007
$\sigma_v = 1.5$ and $\lambda = 2$				
Mean	5.0001	1.9996	2.2345	2.0184
Standard deviation	0.2889	0.0715	0.1855	0.2861

Table 3.9: Double-Bounded DC model simulation results for probability depending on variable z (2): probability estimates

Estimation Results for Double-Bounded DC model and p_1 depending on z (2)			
	$\hat{d}_1$	$\hat{d}_2$	**Mean** $\hat{p}_1$
$\sigma_v = 0.95$ and $\lambda = 10$			
Mean	2.0281	2.0393	0.7952
Standard deviation	0.2605	0.2585	
$\sigma_v = 0.95$ and $\lambda = 5$			
Mean	2.0536	2.0808	0.796
Standard deviation	0.4123	0.3677	
$\sigma_v = 0.95$ and $\lambda = 2$			
Mean	2.2271	2.3132	0.7931
Standard deviation	0.9041	0.9918	
$\sigma_v = 1$ and $\lambda = 10$			
Mean	2.0309	2.0414	0.797
Standard deviation	0.2634	0.2578	
$\sigma_v = 1$ and $\lambda = 5$			
Mean	2.0618	2.081	0.7969
Standard deviation	0.367	0.3686	
$\sigma_v = 1$ and $\lambda = 2$			
Mean	2.257	2.3289	0.7946

Table 3.9: (continued)

Standard deviation	0.9952	0.961	
$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 10}$			
Mean	2.033	2.0456	0.797
Standard deviation	0.2692	0.2657	
$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 5}$			
Mean	2.0664	2.0871	0.7969
Standard deviation	0.3684	0.362	
$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 2}$			
Mean	2.2807	2.5523	0.7942
Standard deviation	1.9851	5.366	
$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 10}$			
Mean	2.0519	2.066	0.793
Standard deviation	0.3041	0.3045	
$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 5}$			
Mean	2.0668	2.0873	0.7969
Standard deviation	0.3799	0.3762	
$\boldsymbol{\sigma_v = 1.5\ and\ \lambda = 2}$			
Mean	2.2801	2.4021	0.7946
Standard deviation	2.1368	3.1363	

Table 3.10: Number of replication removed for Double-Bounded DC model with probability depending on variable z

Number of replications removed for p1 depends on z			
	$\boldsymbol{\lambda = 10}$	$\boldsymbol{\lambda = 5}$	$\boldsymbol{\lambda = 2}$
$\boldsymbol{\sigma_v}$**=0.95**	5	4	44
$\boldsymbol{\sigma_v}$**=1**	5	6	35
$\boldsymbol{\sigma_v}$**=1.25**	4	3	11
$\boldsymbol{\sigma_v}$**=1.5**	13	6	25

As Table 3.8 and 3.9 illustrate, for the majority of the cases, the mean estimates of all parameters are close to the true values. More analytically, for the cases with $\sigma_v \leq 1.25$ it is noticed that the mean estimates of the parameters are very close to the given values and the standard deviation for most of the parameters is small.

Additionally, Table 3.9 presents all the information concerning the estimate of the class membership probability. In the present case, each respondent has a different probability of overestimating his/hers WTP thus the mean probability of all respondents is calculated for each replication by using Eq. (3.34). Table 3.9 shows the mean estimates for the parameters $d_1$and d_2 which are very close to the true values. Additionally, Table 3.9 contains a representative probability which is the mean probability of the 1000 class membership probabilities.

To continue with the analysis, Table 3.10 illustrates the number of replications removed from each case. As it can be seen, the replications removed for all cases were less than 45. More specifically the number of replications removed ranges between 3 and 44 replications. Additionally it can be noticed that as λ gets smaller the number of replications removed increases compared to the cases where λ is 10 or 5.

Finally, Table 3.11 presents the bias between the mean WTP and the estimated mean WTP for each case, given by Eq. (3.36). Firstly, the bias of Mean WTP is close to 0 for all cases which means that the estimation procedure returned estimates very close to the true mean WTP. The bias in absolute values is very small since it is even below 0.01 for the majority of the cases. More analytically, in all cases the bias is smaller than 0.01 and the smallest bias is in the case where $\sigma_v = 0.95$ and $\lambda = 2$ which is almost zero. Additionally, the biggest bias is in the case where $\sigma_v = 1.25$ and $\lambda = 2$ which the bias is equal to -0.026.

Furthermore, Table 3.11 presents as well the mean WTP estimate and the bias of the mean WTP for the case where hypothetical bias is ignored during the estimation process. As it can be noticed in all cases the mean WTP for the normal model displays much higher biases than the ones in the mixture case. In a few words, comparing the two models it can be concluded that if overestimation is not considered during the estimation process, the estimates are biased and consequently not reliable. The application of the proposed model has a clear gain since unbiased estimates of the mean WTP are obtained and the proposed model is effective in dealing with Hypothetical Bias.

Table 3.11: Bias of Mean WTP for Double-Bounded DC model and probability different for each respondent

Bias of Mean WTP for Double-Bounded DC model and probability of overstatement different for each respondent			
	$\sigma_v = 0.95\ and\ \lambda = 10$	$\sigma_v = 0.95\ and\ \lambda = 5$	$\sigma_v = 0.95\ and\ \lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	12.9942	12.9939	12.994
$Bias_{mixture\ model}$	0.0002	-0.0001	0(0.00001)
Mean $\widehat{WTP}_{normal\ model}$	14.1062	13.7866	13.3739
$Bias_{normal\ model}$	1.1122	0.7926	0.3799
	$\sigma_v = 1\ and\ \lambda = 10$	$\sigma_v = 1\ and\ \lambda = 5$	$\sigma_v = 1\ and\ \lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	12.9943	12.9939	12.9945

Table 3.11: (continued)

$Bias_{mixture\ model}$	0.0003	-0.0001	0.0005
Mean $\widehat{WTP}_{normal\ model}$	14.1264	13.8106	13.3906
$Bias_{normal\ model}$	1.1324	0.8166	0.3966
	$\sigma_v = 1.25\ and\ \lambda = 10$	$\sigma_v = 1.25\ and\ \lambda = 5$	$\sigma_v = 1.25\ and\ \lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	12.9953	12.9935	12.9914
$Bias_{mixture\ model}$	0.0013	-0.0005	-0.0026
Mean $\widehat{WTP}_{normal\ model}$	14.211	13.9165	13.4705
$Bias_{normal\ model}$	1.2169	0.9224	0.4765
	$\sigma_v = 1.5\ and\ \lambda = 10$	$\sigma_v = 1.5\ and\ \lambda = 5$	$\sigma_v = 1.5\ and\ \lambda = 2$
Mean $\widehat{WTP}_{mixture\ model}$	12.9943	12.9944	12.9925
$Bias_{mixture\ model}$	0.0003	0.0004	-0.0015
Mean $\widehat{WTP}_{normal\ model}$	14.2802	14.005	13.5459
$Bias_{normal\ model}$	1.2861	1.011	0.5519

Considering the findings above, for the case where the probability of class membership is different for each respondent, it can be concluded that the proposed model works effectively in returning estimates close to the true values and furthermore to minimize the bias.

3.5. The double-bounded DC mixture model applied to empirical data

The simulation results have shown that the proposed mixture model can succeed in overcoming hypothetical bias. In the present section an empirical examination of the proposed model with real CV data is presented.

3.5.1. The dataset

In the present empirical application the data that are used refer to a sample of 1827 observations and 22 variables, 3 of the variables require information about the yes-no answers regarding the WTP scenario and the rest of the variables are additional information obtained from the questionnaires. The dataset is available in the R package "Ecdat" under the name "kakadu" obtained from Werner (1999), (Croissant and Graves, 2020:84-85). The origin of the dataset comes from the paper "Valuing the Preservation of Australia's Kakadu Conservation Zone" published by Richard T. Carson, Leanne Wilks and David Imber in 1994.

The main issue of the CV scenario was whether the mining industry should proceed in the Kakadu Conservation Zone or instead the Kakadu Conservation Zone should be added to the Kakadu National Park (Carson et al., 1994: 727). The survey took place in Australia and in the Northern Territories of the Kakadu National Park for two different impact scenarios. The major impact scenario contained an analytical description about the chemicals that are going to be used and furthermore it was mentioned that there might be a water shortage and there could be losses in wildlife. On the other hand the minor impact scenario mentioned only that toxic chemicals will be used and that the wildlife will be disturbed without mentioning the possible water shortage or that there would be losses in wildlife (Carson et al., 1994:730).

Carson et al. (1994) applied the double-bounded DC elicitation method and Table 3.12 presents the four sets of dollar amounts that were used. Additionally, the questionnaire aimed to elicit information about recycling, watching nature shows on television, membership in environmental organizations and demographic information such as income, age, education etc. (Carson et al., 1994: 732-733). The survey took place in

September 1990 and 2,034 respondents were interviewed. The respondents were given randomly one out of the eight different versions of the questionnaire, the two impact scenarios and the four different sets of amounts (Carson et al., 1994: 732).

Table 3.12: Empirical application-the 4 bid sets given to respondents

Bids given to respondents							
$bid_1 = 100$		$bid_1 = 50$		$bid_1 = 20$		$bid_1 = 5$	
yes	no	yes	No	Yes	No	Yes	no
$bid_2 = 250$	$bid_2 = 50$	$bid_2 = 100$	$bid_2 = 20$	$bid_2 = 50$	$bid_2 = 5$	$bid_2 = 20$	$bid_2 = 2$

Werner (1999:479) used the Kakadu dataset in order to use a mixture distribution allowing respondents in the lowest WTP category to be classified in two groups, those who have zero WTP and those that have a non zero WTP but smaller than the amount that they were asked. For her research she used the Kakadu dataset for only the Australian sample and only for the major impact scenario. In a few words, Werner assumed that there is an unknown part of the sample where respondents that answered no-no to the double-bounded WTP questions actually have a zero WTP. Our difference from Werner is that the present study assumes that there is an unknown number of respondents that when they respond no to the first bid, irrelevant from the answer in the second bid, there might be a chance that the respondent understates his/hers WTP.

Following Werner's work the model given in Eq. (3.38) is going to be estimated in the case where the class membership probability is no longer the same for all respondents but is determined by a number of variables and the data that were used refer only to the Australian sample.

The WTP was determined by:

$$\log(WTP_i) = a + \beta_1 JOBS + \beta_2 FINBEN + \beta_3 MINEPARKS + \beta_4 MOREPARKS + \beta_5 ENVCOV + \beta_6 AGE + \beta_7 INCOME + v_i \quad (3.38)$$

And the probability of understating WTP is determined by the following variables and it is modeled by using a logistic:

$RECPARKS, LOWRISK, ABORIGINAL, FINBEN, MINEPARKS, MOREPARKS, AGE$ and $MAJOR$.

The variables that were used as explanatory variables in Eq. (3.38) are the explanatory variables that Werner (1999) has used and are statistically significant. Table 3.13 describes in detail the definition of each variable from Eq. (3.38) and the determinants of the probability as they are defined by Carson et al. (1994:742). To sum up, Eq. (3.38) and the probability were estimated with the double-bounded DC format mixture model for the case where probability is no longer a constant but differs among respondents. Table 3.14 illustrates the estimation results.

Table 3.13: Variable definition (Carson et al., 1994:742)

The variables used in the estimation model	
Variable name	**Variable definition**
RECPARKS	Measures the agreement by respondent that the greatest value of national parks and nature reserves is in recreational activities such as camping, bushwalking, photography (1-5).
JOBS	How important the respondent feels jobs are in making resource decisions (forest and mineral resources).High values indicate jobs are an important factor.
LOWRISK	Measures acceptance of low risk mining activities. High values indicate greater acceptance.
ABORIGINAL	Measures the importance of Kakadu to Aboriginals should be taken into account as a factor in making decisions concerning Kakadu. High values indicate that this factor should be taken into account.
FINBEN	Measures the importance of financial benefits in making natural resources decisions. High values indicate great importance.
MINEPARKS	Measures how strongly the respondents feel that mining within national parks reduces the value of the parks. High values indicate mining reduces the value of the parks
MOREPARKS	Measures how strongly the respondents feel more national parks should be created from state forests. High values indicate that respondents favor more parks.

Table 3.13: (continued)

ENVCON	Measures if the respondents are environmentally minded consumers (1: if respondents are recycling and buy environmentally friendly products, 0: otherwise).
AGE	The age of the respondent.
INCOME	The yearly income that the respondent report.
MAJOR	Indicates if the respondent received the major impact scenario.

3.5.2. Estimation Results

As a first step, Eq. (3.38) was estimated for the case where no overestimation or underestimation occurs in the sample. The ML estimates of the parameters and the median WTP are illustrated in the second half of Table 3.14. The main reason of presenting these results is to enable the comparison with the estimates obtained by the mixture model estimation.

As a next step, it was necessary to evaluate if hypothetical bias occurs and in which form, in the form of overestimation or underestimation. Firstly, Eq. (3.38) was estimated for the case where hypothetical bias exists in the form of overestimation by using the closed-form approximation for the cumulative distribution function of a normal-half normal composed error that Tsay et al. (2013) proposed, given by Eq. (3.22). The estimate of λ was almost zero, $\hat{\lambda} = 0.0088$ and a procedure was followed in order to test if overestimation occurs. More specifically, we used the pseudo-likelihood ratio test - PLR test (Kumbhakar et al., 2013:69) and the null hypothesis was that $p = 1$ which means that no overestimation occurs in the sample. The pseudo-likelihood ratio test is given by Eq. (3.39),

$$PLR = -2(L_{Normal} - L_{Mixture}) \quad (3.39)$$

where L_{Normal} is the log-likelihood of the normal linear model and $L_{Mixture}$ is the log-likelihood of the mixture model and it is distributed as a mixture of χ^2distributions. By applying Eq. (3.39), $PLR = -0.012$ and compared to the value of $\chi^2_{1,0.05} = 2.706$, given by Kodde and Palm (1986:1246), it can be concluded that the null hypothesis was not rejected since $PLR = -0.012 < \chi^2_{1,0.05} = 2.706$ and no overestimation occurs. Since no overestimation occurs in the data the estimates of the mixture model in this case are not presented.

Secondly, the data were estimated by the mixture model that takes into consideration hypothetical bias in the form of underestimation. For the case of underestimation Eq. (3.38) was again estimated but instead of using the closed-form approximation for the cumulative distribution function of a normal-half normal composed error that Tsay et al. (2013) proposed, given by Eq. (3.22), was used the closed-form approximation for the cumulative distribution function of a normal-half normal composed error in the case where underestimation occurs, given by Eq. (3.24).

Applying the mixture model in order to overcome hypothetical bias in the form of underestimation, it was found that indeed WTP was underestimated. More specifically, applying the pseudo-likelihood ratio test given by Eq. (3.39), $PLR = 238.264$ compared to $\chi^2_{1,0.05} = 2.706$ (Kodde and Palm, 1986:1246) results that the null hypothesis, $p = 1$ which means that no underestimation occurs in the sample, is rejected and hypothetical bias occurs in the form of underestimation. In Table 3.14, the first part illustrates the estimates of the parameters and the median WTP.

Table 3.14: Hypothetical bias vs No hypothetical bias

Estimates for the Kakadu National Park hypothetical bias taking into account during estimation and probability determined by variables						
	Parameter estimates (Mixture Model: Hypothetical bias is taken into account)			Parameter estimates (Normal Model: No Hypothetical bias taken into account)		
WTP parameters	Estimates	Standard error	t value	Estimates	Standard error	t value
Constant	6.4453	0.6244	10.323	2.7203	0.6422	4.236
JOBS	-0.4926	0.0798	-6.170	-0.6144	0.0973	-6.312
FINBEN	-0.2929	0.081	-3.617	-0.8035	0.0969	-8.295
MINEPARKS	0.1141	0.0918	1.243	1.1485	0.0965	11.902
MOREPARKS	0.1202	0.0889	1.353	0.6309	0.0937	6.731
ENVCON	0.4077	0.1705	2.391	0.5469	0.2091	2.616
AGE	-0.018	0.0054	-3.301	-0.041	0.0062	-6.568
INCOME	0.0249	0.0066	3.769	0.0136	0.0065	2.090
σ_v^2	4.4923	0.5635	7.972	3.2713	0.16	20.443
λ	8.8033	2.224[a]	3.9583			
Probability Parameters	Estimates	Standard error	t value			
Constant	-0.8231	1.3555	-0.607			
RECPARKS	0.5717	0.1776	3.219			
LOWRISK	-1.2435	0.214	-5.811			
ABORIGINAL	0.411	0.1468	2.8			
FINBEN	-0.8114	0.1616	-5.020			
MINEPARKS	1.3412	0.2326	5.765			

Table 3.14: (continued)

MOREPARKS	0.4637	0.1668	2.780	
AGE	-0.0317	0.0094	-3.383	
MAJOR	1.0217	0.3039	3.362	
Mean probability	0.7145			
Median WTP	178.4478			69.2936
Log-likelihood	-1420.56			-1539.692
BIC	2983.818			3146.978
AIC	2879.12			3097.384

[a] the standard error for parameter λ was computed by applying delta method since λ was parametrised as $\lambda = e^{\gamma}$.

Table 3.14 shows that the median WTP for the case where no hypothetical bias is taken into account is smaller than in the case where hypothetical bias is taken into consideration during the estimation. The main reason why the median WTP was determined instead of the mean WTP was because the distribution of WTP is log normal which is not symmetric consequently the median WTP is determined by Eq. (3.40)

$$median\ WTP = e^{mean\ \ log\,(WTP)} \qquad (3.40)$$

In order to compare the two models the BIC (Schwarz Bayesian Information Criterion) and the AIC (Akaike Information Criterion) were computed for both models that take into account the log-likelihoods and the number of parameters. The mixture model has ten more parameters than the normal model and the BIC for the mixture model is equal to 2903.093 and for the normal model is equal to 3108.7396. The BIC for the mixture model is smaller which means that it has a better fit. The same conclusion is made by applying the AIC, more specifically, the AIC for the mixture model is 2879.12 and for the normal model is 3097.384.

As it is expected, in the case where underestimation occurs and it is not taken into account during the estimation process, the constant term would be affected and the estimate is expected to be smaller compared to the case where hypothetical bias is considered.

Conclusions

The main goal of this thesis is to propose a statistical model that can be applied to CV survey data in order overcome hypothetical bias, i.e the fact that some respondents might not answer truthfully to the valuation question. The sections above analyze in detail the theoretical framework, the methodology that was used and all the steps that were followed in order to build the proposed mixture model. In order to test the validity of the model and check whether it is successful in overcoming hypothetical bias several simulations took place and we showed the results of 1000 replications.

The findings confirm that the model is effective in overcoming hypothetical bias since the mean estimates are very close to the generated values for all the parameters. More analytically, the Mean WTP, the parameter λ which is the overestimation indicator and the probability of the class membership were estimated and their values were close to the "true" values. Overall, these results are able to confirm that our model will overcome one of the basic critiques that the CVM has received and that questions the validity of the valuation estimates obtained from the application of this method.

References of Chapter 3

Ahmed, S. U. and Gotoh, K., 2006, *Cost-Benefit Analysis of Environmental Goods by Applying the Contingent Valuation Method, Some Japanese Case Studies,* Japan, Springer-Verlag Tokyo

Alberini, A., 1995, "Optimal Designs for Discrete Choice Contingent Valuation Surveys: Single Bound.Double-Bound and Bivariate Models", *Journal of Environmental Economics and Management,* vol.28, p.287-306

Amsler, C., Schmidt,P. and Tsay W-J, 2019, "Evaluating the CDF of the distribution of the stochastic frontier composed error", *Journal of Productivity Analysis,* vol. 52, p.29-35

Arrow, K., Solow, R., Portney P.R., Leamer, E.E., Radner, R., and Schuman, H., 1993,*Report of NOAA Panel on Contingent Valuation,*National Oceanic and Atmospheric Administration

Carson, R.T. and Hanemann, W.M., 2005, Contingent Valuation. In: Mäler K.G. and Vincent J.R., eds., 2005, *Handbook of Environmental Economics,* vol.2, Elsevier B.V., chapter 17, p.821-936

Carson, R.T., Wilks, L. and Imber, D., 1994, "Valuing the Preservation of Australia's Kakadu Conservation Zone", *Oxford Economic Papers,* vol.46, (special issue on Environmental Economics), p.727-749

Chien, Y-L., Huang, C. J. and Shaw, D., 2005, "A general model of starting point bias in double bounded dichotomous contingent valuation surveys", *Journal of Environmental Economics and Management,*vol.50, p.362-377

Croissant, Y.and Graves, S., 2002, Ecdat: Data Sets for Econometrics, R package version 0.4-1 Available at: https://cran.r-project.org/web/packages/Ecdat/Ecdat.pdf (last access 14-04-2022)

Hofler, R.A. and List, J.A., 2004, "Valuation on the frontier: Calibrating actual and hypothetical statements of value", *American Journal of Agricultural Economics,*vol.86, no.1, p.213-221

Kodde, D.A. and Palm, F.C., 1986, "Wald Criteria for Jointly Testing Equality and Inequality Restrictions", *Econometrica*, vol. 54, n.5, p.1243-1248

Kumbhakar, S.C. and Lovell, K.C.A., 2000, *Stochastic Frontier Analysis*, United States of America, Cambridge University Press

Kumbhakar, S.C., Parmeter C.F. and Tsionas E.G., 2013, "A zero inefficiency stochastic frontier model", *Journal of Econometrics,* vol.172, p.66-76

Loomis, J. B., 1990, "Comparative Reliability of the Dichotomous Choice and Open ended Contingent Valuation Techniques", *Journal of Environmental Economics and Management,* vol. 18, p.78-85

Tsay, W-J.,Huang C.J., Fu, T-T. and Ho, I-L., 2013, “a simple closed –form approximation for the cumulative distribution function of the composite error of stochastic frontier model” , *J Prod Anal*,vol.39, p.259-269

Werner, M., 1999, “Allowing for zeros in Dichotomous-Choice Contingent-Valuation Models”, *Journal of Business & Economic Statistics,* vol.17, no. 4, p. 479-48

Chapter 4
Comparing different starting values techniques for the double-bounded DC format mixture model

Introduction

In the proposed mixture model, the EM algorithm is an important component of the estimation process. The EM algorithm is an iterative algorithm to find maximum likelihood estimates that starts from an initial point for the parameters and proceeds to iteratively update the parameters estimates until convergence is obtained.

The EM algorithm though, deals with a number of drawbacks. The first drawback is the need for good initial values and the second refers to the possibility that the algorithm might get trapped in local optima (Panić et al., 2020:1). The determination of the initial-starting point for the EM algorithm is a very important task thus in the following sections this topic is going to be examined in detail. The main goal is to examine different initialization strategies in regard to the class assignment. More specifically, the initial class assignment was obtained by different methods in order to obtain starting values for the EM algorithm. Afterwards a comparison of the results took place in order to conclude which initialization technique comes up with better starting values in order to obtain estimates close to the real values and consequently to eliminate hypothetical bias.

The present chapter is organized as follows: the first section presents the double-bounded mixture model that is going to be examined in regard to the initialization techniques. Additionally, the first section discusses the importance of determining proper starting values. The second section presents the initialization techniques that are used in the comparison while the third section contains the simulation results for each initialization technique. Finally the last section presents a description of the comparison criteria that are used and the related literature as well as the comparison results for the conducted experiments.

4.1. Initializing the double-bounded mixture model

4.1.1. The double-bounded mixture model

The double-bounded mixture model allows for heterogeneity in the response behavior of individuals to WTP questions whereas part of the sample answers truthfully to the valuation exercise while the rest of the sample overstates their WTP. As will be shown below this heterogeneity can be modeled by using a conventional two-sided error term for the first group and a composed error term for the second group.

This strategy is the one followed by Kumbhakar et al. (2013:67) in their paper related to the productive inefficiency of firms, in which they stated that in a sample both efficient and inefficient firms can exist with a probability. In the present case the data consists of discrete responses to different bids presented and the underlying model for WTP is as a mixture of two classes. The two classes are:

Class 1: respondents that answer sincerely and WTP is given by

$$WTP_i = \beta' x_i + v_i \quad i = 1, \cdots, n \qquad (4.1)$$

$$where\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_\kappa \end{bmatrix}, x_i = [x_{i1} \quad \cdots \quad x_{i\kappa}]', x_{i1} \equiv 1 \forall i \text{ and } v_i \sim N(0, \sigma_v^2)$$

Class 2: respondents that overstate their WTP and therefore Eq. (4.2) holds.

$$WTP_i^* = \beta' x_i + \varepsilon_i \qquad (4.2)$$

where

$$\varepsilon_i = v_i + u_i \qquad (4.3)$$

$u_i \sim iid\ N^+(0, \sigma_u^2)$ non-negative Half Normal. Additionally, the probabilities of belonging to class 1 and class 2 are given by p_1 and $p_2 = (1 - p_1)$ respectively. For class 2 it is considered that hypothetical bias exists in the form of overstatement of WTP, thus in Eq. (4.3) where the composed error ε_i is determined, u_i is the one-sided non-negative error and it reflects that the yes-saying tendency will raise each respondents elicited WTP (Chien et al., 2005:364-65).

Furthermore when the error term u_i approaches zero, the gap between the real and hypothetical values is decreased and the hypothetical values → real values (Hofler and List, 2004:216). In a few words, in cases where the error term u_i approaches zero, the hypothetical WTP tends to be equal to the actual WTP and consequently the existence of hypothetical bias tends to disappear.

Additionally, the model is parameterized very often in terms of the two parameters defined below

$$\sigma^2 = \sigma_v^2 + \sigma_u^2 \qquad (4.4)$$

and

$$\lambda = \frac{\sigma_u}{\sigma_v} \qquad (4.5)$$

Summing up the model becomes

$$WTP_i^{\ *} = \begin{cases} \beta' x_i + v_i & with\ probability\ p_1 \\ \beta' x_i + v_i + u_i & with\ probability\ p_2 \end{cases} \qquad (4.6)$$

Each respondent has to answer two successive bid questions whereas the second bid depends on the answer that is given to the first bid. Firstly, the respondents have to answer to the first bid with a yes or no. Denoting as y_{1i} and y_{2i} the responses to the two bids respectively, y_{1i} and y_{2i} are presented in Eq. (4.7), Eq. (4.8) and Eq. (4.9).

$$y_{1i} = \begin{cases} 1 & yes\ to\ bid_{1i} \\ 0 & no\ to\ bid_{1i} \end{cases}, \qquad i = 1, \cdots, n \qquad (4.7)$$

If yes to bid_{1i} ,$y_{1i} = 1$ and $bid_{2i} > bid_{1i},\ y_{2i} = \begin{cases} 1\ for\ yes\ to\ bid_{2i} \\ 0\ for\ no\ to\ bid_{2i} \end{cases}$ (4.8)

If no to bid_{1i}, $y_{1i} = 0$ and $bid_{2i} < bid_{1i},\ y_{2i} = \begin{cases} 1\ for\ yes\ to\ bid_{2i} \\ 0\ for\ no\ to\ bid_{2i} \end{cases}$ (4.9)

In order to receive yes-yes as an answer for both bids the individual's WTP^* must satisfy that $WTP_i^* > bid_{1i}$ and $WTP_i^* > bid_{2i}$ so we have that $WTP_i^* > bid_{2i} > bid_{1i}$. A yes answer to the first bid followed by no to the second bid, means that the individual's WTP^* is greater than bid_{1i} but smaller than bid_{2i}, $WTP_i^* > bid_{1i}$ and $WTP_i^* < bid_{2i}$ so we have that $bid_{1i} < WTP_i^* < bid_{2i}$.

On the other hand, for a no answer to bid_{1i} and a yes answer to bid_{2i} the respondent's WTP^* must me smaller than bid_{1i} and greater than bid_{2i} , $WTP_i^* < bid_{1i}$ and $WTP_i^* > bid_{2i}$ so we have that $bid_{2i} < WTP_i^* < bid_{1i}$. Finally, in order to receive a no answer to both bids, the respondent's WTP^* is smaller than bid_{1i} and bid_{2i} , $WTP_i^* < bid_{1i}\ and\ WTP_i^* < bid_{2i}$ so we have that $WTP_i^* < bid_{2i}<bid_{1i}$.

Additionally, the density functions for each case for the error are the following:

No Hypothetical Bias:

$$f_1(v_i) = \frac{1}{\sigma_v\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v_i}{\sigma_v}\right)^2} = \frac{1}{\sigma_v\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{wtp_i-\beta' x_i}{\sigma_v}\right)^2} \qquad (4.10)$$

Hypothetical Bias:

For Normal ($v_i \sim iid\ N(0,\sigma_v^2)$) and Half-Normal $\left(u_i \sim iid\ N^+(0,\sigma_u^2)\right)$ distributions we have the composed error density distribution:

$$f_2(\varepsilon_i) = \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right) \qquad (4.11)$$

(Kumbhakar and Lovell, 2000)

where $\varphi(.)/\Phi(.)$ are the density/cumulative distribution of the $N(0,1)$ and ε_i, σ and λ are given by Eq. (4.3), the square root of Eq. (4.4) and (4.5) respectively.

As explained in Chapter 3 (p. 96-98), we will use the closed form approximation of Tsay et al. (2013) for the cumulative distribution of the composed error term where the latter is given by

$$F = \int_{-\infty}^{Q} \frac{2}{\sigma}\varphi\left(\frac{\varepsilon_i}{\sigma}\right)\Phi\left(\frac{\lambda}{\sigma}\varepsilon_i\right)\, d\varepsilon_i \qquad (4.12)$$

where $Q = bid_i - \beta' x_i$.

For a sample of n observations the log-likelihood function for the proposed mixture model is given by

$$\ln L = \sum_{i=1}^{n} ln\big[y_{1i}y_{2i}\big(p_1(1-f_{12}) + p_2(1-f_{22})\big)$$
$$+ y_{1i}(1-y_{2i})(p_1(f_{12}-f_{11}) + p_2(f_{22}-f_{21}))$$
$$+ (1-y_{1i})y_{2i}(p_1(f_{11}-f_{12}) + p_2(f_{21}-f_{22}))$$
$$+ (1-y_{1i})(1-y_{2i})\,(p_1 f_{12} + p_2 f_{22})\big] \qquad (4.13)$$

where $f_{11} = \Phi\left(\frac{bid_{1i} - \beta' x_i}{\sigma_v}\right)$, $f_{12} = \Phi\left(\frac{bid_{2i} - \beta' x_i}{\sigma_v}\right)$,

$f_{21} = F'(bid_{1i} - \beta' x_i)$ and $f_{22} = F'(bid_{2i} - \beta' x_i)$.

In order to estimate a Mixture model we apply the EM algorithm which was introduced by Dempster, Laird and Rubin in the late 1970s (McLachlan and Peel, 2000:4). The EM algorithm consists of two steps, the E-step (expectation) and the M-step (maximization) and treats the estimation problem as a missing data problem, where the missing data is the information about class membership. Appendix A of Chapter 2 presents in detail a description of how the EM algorithm works. Applying EM to the mixture problem ensures monotonic increases of the likelihood values (McLachlan and Peel, 2000:48).

4.1.2. EM algorithm and the importance of initial values

For estimating finite mixture models, the ML estimation via the EM algorithm has dominated the field for several reasons. Firstly, it is simple compared to other methods, secondly, it can exhibit monotonic convergence, thirdly, its' statistical interpretation is simple and finally, if the starting values are within admissible range so do the estimates (Karlis and Xekalaki, 2003:577-78).

On the other hand, there are several difficulties regarding the application of the EM algorithm. Namely a few drawbacks of the EM algorithm is that it has slow convergence, another drawback is the need of proper choice of stopping rule in order to detect if the maximum is reached and additionally the choice of initial values in order to find the global maximum in the fewer as possible iterations (Karlis and Xekalaki, 2003:578). "The choice of initial values is of great importance in the algorithm-based literature as it can heavily influence the speed of convergence of the algorithm and its ability to locate the global maximum" (Karlis and Xekalaki, 2003:579).

The EM algorithm starts from some initial value for model's parameters. Choosing starting values has an important role since "different starting strategies and stopping rules can lead to quite different estimates in the context of fitting mixtures of exponential components via EM algorithm" (McLachlan and Peel, 2000:54). Two problems emerge from the need of initial values. Firstly the number of components of the mixture must be known in advance, but in the majority of the cases this information is unavailable. Secondly even if the number of components is known, the suitable set of initial values must been determined in order to avoid being trapped in local optima (Panić et al., (2020:1-2).

In general, mixture models present an estimation difficulty in regards of the discrimination between a local optima and the global optimum. Mixture models may have several local optima and the normal mixture models may also come up with singularities where it means that there might be points that the likelihood function may go to infinity causing non-convergence to the model (Hipp and Bauer, 2006:36).

Hipp and Bauer (2006:36) state that the estimation of a mixture model should be done with several sets of starting values in order to avoid these kind of irregularities on the likelihood surface and in order to enable the determination of the global optimum. In short, the starting values for the EM algorithm are crucial since the convergence to the global maximum is strongly dependent on the starting values and additionally the speed of convergence of the EM algorithm depends to a high degree on the initial values (Biernachi, 2004:267).

4.2. Initialization techniques

The EM algorithm requires starting values for the model's parameters and several initialization strategies to determine starting values have been proposed in the literature. In the present section the most commonly used initialization techniques are introduced and briefly described.

Many techniques that have been proposed in order to choose initial values are Clustering based ideas. Cluster analysis includes a broad set of techniques in order to find subgroups of the observations within a dataset. Clustering is a very important task in data analysis. This task arranges a set of objects so the objects in the identical group are as related as possible to the objects included in the same group-cluster. In a few words, cluster analysis is done by separating the data into groups-clusters by detecting similarities among the data according to their characteristics and grouping similar data objects into clusters (Mann and Kaur, 2013:43-44).

Some clustering techniques have been suggested by Woodward et al. (1984) and McLachlan (1988). Woodward et al. used an ad hoc quasi-clustering technique for the case of a mixture of two Normals. More specifically, Woodward et al. allowed as possible initial values of the class probability p the values from 1-9. For each value the sample was divided in two subsamples and the initial value of the class probability was determined as “the value at which $p(1-p)(m_1-m_2)^2$ is maximized, where m_j is the sample median of the jth subsample” (Woodward et al., 1984:592). McLachlan (1988:418) proposed the use of the two-dimensional scatter plots combined with principal component analysis in order to search for the presence of clusters. The visual clustering of the data was used as initial values for the posterior probabilities.

Furthermore, Leroux (1992:1351-53) noticed that in order to obtain good estimates from maximum likelihood estimation, consideration should be given to the number of mixture components. In order to choose the number of components a number of criteria were used such as BIC and AIC comparing the constrained maximum-likelihood estimates for one, two and three components.

In particular settings another initialization strategy that can be applied only on simulated data is the use of the real cluster membership probability. This technique is proposed for cases where the researcher aims to investigate the behavior of the EM algorithm when the starting point is the optimal solution (Maruotti and Punzo, 2021:454). Another initialization technique is the “Random Short EM”. This procedure consists of a number of short runs of the EM algorithm with a certain number of iterations by starting from a different random position each time. The starting values for the EM algorithm are determined by the short EM with the biggest likelihood (Biernachi et al., 2013:567). As Maruotti and Punzo (2021:454) state, this procedure has S short runs of the EM algorithm and each short run has H iterations from different random positions. The random positions are obtained by selecting K centres randomly. The values that were considered for S and H were two, 1 or 10 and 1 or 5 respectively. In order to implement this initialization strategy they use the "rand.EM( )" function included in the "EMCluster" package which refers to finite Gaussian mixtures (Chen et al., 2021).

Other initialization techniques belong to partitional clustering algorithms. Partitional clustering algorithms obtain classification of the observations into a number of clusters, based on their similarity. The number of clusters is determined in advance by the researcher. Namely the most popular algorithms are K-means clustering, K-medoids clustering or PAM and CLARA algorithm (Kassambara, 2017:35). Additionally another clustering algorithm is the fuzzy C-means which is a variation of the K-means algorithm and SOM algorithm (Self-Organizing Map) (Brun et. al, 2007:813).

There are a number of R packages that provide initial values for the EM algorithm for the case of Gaussian mixtures. Namely there is the "EMcluster" package, which is already mentioned above (Chen et al., 2021), the "mclust" package (Fraley et al., 2022), the "mixtools" package (Young et al., 2020), the "mixture" package (Pocuca et al., 2021) and finally, the "Rough-Enhanced-Bayes mixture estimation (REBMIX) algorithm (Panic et al., 2020:1). The REBMIX algorithm provides an alternative to the EM algorithm for finding parameter estimates for mixture models where both estimate accuracy and estimation time are important criteria and therefore the resulting estimates, while close to the ones from the EM, might not be as good as the latter. Based on these observations, Panic et al. (2020:2) propose the use of the REBMIX algorithm as an initialization strategy for the EM.

For the initialization process two different techniques were adopted in order to initially partition the data into two classes and determine starting values for the remaining parameters. The first technique refers to a random classification applied in two different ways, namely only once and also multiple times. The second technique refers to another clustering technique known as the k-means clustering algorithm. In the following subsections, each one of the initialization techniques that has been applied in the proposed double-bounded mixture model is analyzed in detail.

4.2.1. Random initialization

To define starting values for the EM algorithm, a procedure was applied in order to assign randomly into two classes the data. In the present subsection the procedure that was applied in Chapters 2 and 3, namely random assignment, is analyzed. Random initialization might be the most employed technique in order to initialize EM algorithm (Biernachi et al., 2003:566).

Random draws were used from a uniform distribution to separate the data into two groups and an observation was classified in the first class whenever the draw was below 0.5. Random draws from a uniform distribution is a commonly used procedure for starting values (Hipp and Bauer, 2006:41, Shireman et al., 2017:284).

4.2.2. Random initialization multiple times

The present initialization technique is an extension of the previous one. More analytically, in this strategy the simple random initialization is applied several times, instead of only once and the selection of the “best” solution must take place. The best solution is defined as the one that returns the highest maximized likelihood. Briefly, this extended strategy refers to three steps, search/ run/ select in order to maximize the likelihood (Biernachi et al., 2003:563-66). Maximizing the likelihood contains the three steps. Firstly a search method is built in order to generate p initial positions, secondly, the EM algorithm must be run for a given number of iterations at each initial position and finally, the choice of the solution that provided the best likelihood value (Biernachi et al., 2003:563).

In general, usually for the selection of the best initial values a fit criterion is used, such as the BIC (Shireman et al, 2017:284). The random initialization technique can be applied together with the technique called the short runs of the EM algorithm. In short runs of EM the researches does not wait for convergence and the stopping rule of the EM algorithm is determined by a stopping rule with a specific number of iterations (Biernachi et al., 2003:567). Shireman, Steinley and Brusco (2016:477) recommended that the number of initializations must be large. More specifically, they suggest "that a "large" number of initializations for mixture modeling be over 1000". Although, in their research they found the optimal solution in a smaller number of iterations but they mentioned that this could be by chance.

In practice a safe number of random initializations exceeding 1000 is unfeasible, thus a smaller number of initializations is usually applied, for instance 100 random initializations (Shireman et al., 2017:284). To sum up, in the double-bounded DC format mixture model the random assignment to two classes applied 100 times was adopted and for each random assignment a short EM algorithm was adopted. By the end of the 100 assignments the starting values for the EM algorithm were selected by the BIC criterion. Moreover, since the number of parameters was the same across the different assignments, the BIC criterion is equivalent to using biggest log-likelihood.

For the short EM the number of iterations needed to be determined having in mind that the bigger the initial EM iteration number is the more time intensive the procedure will be. The literature does not offer a unique number for the choice of the number of iterations. For example, the Mplus software defaults to 10 iterations of the EM algorithm from 20 random starts (Shireman et al., 2017:285), while StataCorp (2021:4) in order to determine starting values for finite mixture models defaults to 20 iterations. On the other hand, Biernachi et al. (2003:574) concluded that after comparing eight different strategies with repeating algorithms, the random strategy 10EM (10 initializations with 100 iterations for each EM) returned the best results. Taking into consideration all the above it was decided to use 20 iterations for the short EM procedure.

4.2.3. Classification with k-means algorithm

"K-means clustering is the most commonly used unsupervised machine learning algorithm for partitioning a given data set into a set of k groups, where k represents the number of groups pre-specified by the analyst" (Kassambara, 2017:36). The basic idea is to define clusters in a way that the total intra-cluster variation is minimized. Each observation from the given dataset is assigned to the cluster in which the sum of squares distance of the specific observation to their assigned cluster center is minimized (Kassambara, 2017:36-7).

K-means has a number of advantages and disadvantages. The advantages are that it is a simple and fast algorithm but on the other hand a number of drawbacks occur. Firstly the number of clusters must be known in advance by the analyst, secondly if the data are rearranged it is possible that the analyst will receive a different solution and thirdly it is affected by outliers (Kassambara, 2017:46). Additionally, another drawback is that k-means tend to find spherical clusters, in cases where the clusters are highly heterogeneous and non spherical, k-means won't be able to find the exact representation of the data and consequently won't be able to provide appropriate starting values (Shireman et al., 2017:285).

Furthermore, because the final clustering result obtained by k-means is affected by the random starting assignments, Kassambara (2017:41) recommends that k-means clustering should be computed with a large number of different random starting assignments, since the algorithm will select the best result corresponding to the lowest within cluster variation. More specifically, the default in R is 1 thus the proposed number is 25 or 50. Additionally, since the number of times that k-means should perform is determined by the researcher and the accurate number is subjective, Shireman et al. (2017:289) propose to set the number of runs for the k-means initialization technique to be set to 100.

Taking into consideration all the above, in the present application, k-means will be performed 100 times by the using the function "kmeans( )" which is in the "stats package" of the R programming language (R Core Team, 2022).

4.3. Simulations for the initialization techniques for the double-bounded mixture model

The present section presents the simulation results for the three different initialization techniques described in section 4.2, namely, the 1 random initialization, the 100 random initializations and the k-means strategy. The simulations were conducted for 4 different cases considered in Chapter 3 and more specifically from Table 3.5 where in those cases it was assumed that the class membership probability p_1 was considered to be constant over respondents. The number of replications was set to 500 for a sample of 1000 observations.

4.3.1. Data generation

In the data generation process, for all cases that are going to be applied, the model that is going to be estimated is simple regression model

$$WTP_i^* = a + \beta x_i + \omega_i \qquad (4.14)$$

given by one explanatory-independent variable $x_i \sim N(4{,}1)$ where the coefficient of x_i, β is equal to 2 and the constant term α is equal to 5. Taking into consideration Eq. (4.6) and Eq. (4.14) the model becomes

$$WTP_i{}^* = \begin{cases} 5 + 2x_i + v_i & with\, probability\, p_1 \\ 5 + 2x_i + v_i + u_i & with\, probability\;\, p_2 \end{cases} \qquad (4.15)$$

where $v_i \sim iid\ N(0, \sigma_v^2)$ and $u_i \sim iid\ N^+(0, \sigma_u^2)$.

For the normal error term v, the values 1 and 1.25 were used for $\sigma_v$while for the half-normal error term u, the choice of values for σ_u followed Eq. (4.16) below.

$$\sigma_u = \begin{cases} 10\sigma_v \\ 5\sigma_v \end{cases} \qquad (4.16)$$

From Eq. (4.5) it follows that the parameter λ takes the values 5 and 10. The bids are determined by Eq. (4.17) and Table 4.1 illustrates the structure of bid_1 and bid_2.

$$bid_2 = \begin{cases} 25\% bid_1 + bid_1 & if\ yes\ to\ bid_1 \\ 25\% bid_1 - bid_1 & if\ \ no\ to\ bid_1 \end{cases} \qquad (4.17)$$

Table 4.1: bid_1 and bid_2

Bids given to respondents							
$\boldsymbol{bid_1}$ **=11**		$\boldsymbol{bid_1}$ **=12**		$\boldsymbol{bid_1}$ **=13**		$\boldsymbol{bid_1}$ **=15**	
Yes	No	Yes	No	Yes	No	Yes	No
$\boldsymbol{bid_2}$ **=14**	$\boldsymbol{bid_2}$ **=9**	$\boldsymbol{bid_2}$ **=15**	$\boldsymbol{bid_2}$ **=9**	$\boldsymbol{bid_2}$ **=17**	$\boldsymbol{bid_2}$ **=10**	$\boldsymbol{bid_2}$ **=19**	$\boldsymbol{bid_2}$ **=12**

The mean WTP at the mean value of x_i is given in Eq. (4.18) below

$$mean\,WTP = (2 + 5\bar{x}) = 12.994 \qquad (4.18)$$

And finally, regarding the class membership, the probability is a constant $p_1 = 0.75$ and $p_2 = 0.25$. In this case all respondents that belong to the same class have the same probability of belonging to the class.

4.3.2. Starting values and initialization strategy

The initialization techniques that have been applied in the present subsection aim to separate the data into two clusters in order to compute starting values for the EM algorithm. The procedure that it was followed in each case is described below.

A. Initial values obtained from 1 random initialization

This initialization strategy is a simple procedure in order to separate randomly the data into two classes. This procedure was followed in Chapter 2 and 3 in order to receive starting values for the EM algorithm. In the following description are analyzed in detail the steps that were followed in order to obtain starting values.

Step A1: Random assignment of observations to two classes

In the first step to determine starting values for the EM algorithm random draws from a Uniform(0,1) were generated and the observation was classified in the one class whenever the draw was below 0.5.

Step A2: ML estimations assuming that one class exists

For each class a probit model was estimated and the class with the biggest constant represents the one with overestimation. By ending this stage, starting values were obtained in order to apply ML estimation for a Normal error model and for a composed error model, assuming that all observations belong in one class, where initial value for σ_u was determined using the same formula as in Step2 in Chapter 2 (p.63-64). At this point the estimation results from the ML estimation for a Normal error model and from the composed error model were used as starting values for EM algorithm. More specifically, for the parameters α, β and σ_v^2 the starting values were determined from the ML of the Normal model, and the starting value of the parameter λ was determined by the ML estimate for the composed model. Additionally, from the clustering procedure the proportion was used as a starting value of the class membership probability.

Step A3: Applying EM algorithm

The EM algorithm was run for the mixture model until a tolerance criterion was reached, where the tolerance criterion was until the value of the log-likelihood obtained in each iteration, satisfies Eq. (4.19)

$$|loglikelihood(k) - loglikelihood(k-1)| < 0.001 \qquad (4.19)$$

where k is the number of the iteration.

Finally, the estimates produced from the EM algorithm were used as starting values to maximize the log-likelihood of the mixture model, since the EM algorithm was stopped before reaching convergence.

B. Initial values obtained from 100 random initializations

This initialization procedure is very similar to the previous initialization strategy. The main difference is that the random assignment is done multiple times and not only once. In the following description are presented the steps that were followed in order to obtain starting values.

Step B1:100 Random assignments

The first step of this procedure includes the Steps A1-A2 described in the previous initialization technique. The addition that is done by the second initialization strategy is the repetition of the first strategy by 100 times. In the current strategy though Step A3 is modified, instead of using a tolerance criterion the number of iterations for the EM algorithm was set to 20.

Step B2: Identifying the "best starting values"

In order to select the best starting values among the 100 initializations a criterion such as the BIC (Shireman et al., 2017:284) can be used. Nevertheless, since the sample size and the number of parameters is the same across the 100 cases, this amounts to basing comparisons on the log-likelihood. The estimates with the highest log-likelihood were therefore used as starting values for the EM algorithm.

Step B3: Applying EM algorithm

The EM algorithm was applied for the mixture model using the initial values determined in the previous step and using the tolerance criterion defined previously in Eq. (4.19). Finally, the estimates produced from the EM algorithm were used as starting values to maximize the log-likelihood of the mixture model.

C. Initial values obtained from k-means classification

For this initialization technique, the k-means algorithm was used to partition the data into two groups. The remaining Steps are the same as case A, thus the following steps were followed:

Step C1: Assigning to two classes with k-means

The data were separated into two classes by applying k-means algorithm.

Step C2: ML estimations assuming that one class exists

Same as in Step A2.

Step C3: Applying EM algorithm

Same as in Step A3.

Finally, the estimates produced from the EM algorithm were used as starting values to maximize the log-likelihood of the mixture model

4.3.3. Simulation results for the double-bounded method model for different initialization strategies

The present subsection presents the simulation results related to parameter estimates and mean Willingness To Pay estimates for the three different initialization techniques. The R programming language (R Core Team, 2022) was used to perform the simulations. More specifically version 4.1.1.

As was the case in Chapters 2 and 3, the class membership probability, p_1, was reparametrized as $p_1 = \frac{1}{(1+e^{kappa})}$ in order to ensure that the estimate of the class membership probability lies in the open unit interval. Additionally, a number of replications were dropped from the results as the standard errors of the maximum likelihood estimates could not be computed. Finally, in addition to the parameter estimates, the bias of WTP is illustrated for each case. More specifically the bias is given by the following equation:

$$Bias = expected\ value\ of\ estimated\ mean\ WTP - mean\ WTP \Rightarrow$$

$$Bias = \left(\hat{\alpha} + \hat{\beta}\bar{x}\right) - (5 + 2\bar{x}) \qquad (4.20)$$

where $\hat{\alpha} = \frac{\sum_{r=1}^{R} \widehat{\alpha}^r}{R}$, $\hat{\beta} = \frac{\sum_{r=1}^{R} \widehat{\beta}^r}{R}$ and R is the number of replications.

A. Simulation Results for the 1 random initialization strategy

The single or 1 random initialization has already been illustrated in Chapter 3 for 1000 simulations (Tables 3.5-3.7). For the sake of comparison the results for 500 replications are reported below. Table 4.2 reports the mean estimates of the parameters $\alpha, \beta, \sigma_v^2, \lambda$ and p_1. Additionally, the number of the replications that have been removed from each case is given in Table 4.3 and finally Table 4.4 presents the bias of the mean WTP for each case and the mean estimate of the WTP.

Table 4.2: Simulation Results from the 1 random initialization technique

Simulation Results for Double-Bounded DC and $p_1 = 0.75$ 1 random initialization					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^2$	$\hat{\lambda}$	$\hat{p}_1$
$\sigma_v = 1$ and $\lambda = 10$					
Mean	4.9631	2.0135	1.0379	10.4198	0.7507
Standard deviation	0.2601	0.0628	0.4752	2.4643	
$\sigma_v = 1$ and $\lambda = 5$					
Mean	4.9498	2.0143	1.0028	5.056	0.7486
Standard deviation	0.2429	0.0597	0.2207	0.7579	
$\sigma_v = 1.25$ and $\lambda = 10$					
Mean	4.9658	2.0135	1.6205	10.7376	0.7502
Standard deviation	0.294	0.0695	0.6545	3.6207	
$\sigma_v = 1.25$ and $\lambda = 5$					
Mean	4.9486	2.0164	1.5871	5.0952	0.7496
Standard deviation	0.2845	0.0689	0.3925	0.9435	

As Table 4.2 shows, in all cases the mean estimates are very close to the true values. Additionally, it can be noticed that the standard deviation of the parameters is small except for the parameter λ in the cases where $\lambda = 10$. In such cases due to the fact that σ_u has very large values compare to σ_v (which is 1 and 1.25 respectively) the parameter λ will probably get higher estimates and since the standard deviation is sensitive to extreme values, the standard deviation of λ might have higher values. Furthermore, Table 4.3 shows that the number of replications removed from all cases is between 15 and 25 replications, representing 3-5% of the replications, which is a very small percentage.

Table 4.3: Replications removed from the 1random initialization technique

Number of replications removed		
	$\boldsymbol{\lambda = 10}$	$\boldsymbol{\lambda = 5}$
$\boldsymbol{\sigma_v = 1}$	20	15
$\boldsymbol{\sigma_v = 1.25}$	20	25

Finally, Table 4.4 reports the bias of the mean WTP for each case, given by applying Eq. (4.20). It can be noticed that in all cases that have been examined the mean bias is very close to zero which means that the model succeeded to overcome hypothetical bias.

Table 4.4: Bias of the mean WTP from 1 random initialization technique

Bias of Mean WTP same probability for overestimation for all respondents 1 random initialization		
	$\boldsymbol{\sigma_v = 1 \; and \; \lambda = 10}$	$\boldsymbol{\sigma_v = 1 \; and \; \lambda = 5}$
$\mathbf{Mean\ W\widehat{T}P_{mixture\ model}}$	13.0111	13.0011
$\mathbf{Bias_{mixture\ model}}$	0.0171	0.0071

Table 4.4: (continued)

	$\sigma_v = 1.25\ and\ \lambda = 10$	$\sigma_v = 1.25\ and\ \lambda = 5$
Mean $\widehat{WTP}_{mixture\ model}$	13.0139	13.0081
$Bias_{mixture\ model}$	0.0199	0.014

B. Simulation Results for the 100 random initializations strategy

The results for the case where random initialization is conducted a 100 times and the "best" initialization is chosen are illustrated in Tables 4.5, 4.6 and 4.7. More specifically, Table 4.5 presents the mean estimates of the parameters α , β, σ_v^2, λ and p_1. Additionally, Table 4.6 shows the number of the replications that have been removed from each case and finally Table 4.7 reports the bias of the mean WTP for each case and the mean estimate of the WTP.

As can be observed from the results, in all cases the mean estimates are very close to the real parameter values. Additionally the standard deviation for all parameters is small except for parameter λ in cases with real λ equal to 10, as in the case with 1 random initialization. The main reason is that in these cases σ_u^2 has very large values compare to σ_v^2 thus the parameter λ will probably get higher estimates and since the standard deviation is sensitive to extreme values, the standard deviation of λ might have higher values.

Table 4.5: Simulation Results for the 100 random initializations strategy

Estimation Results for Double Bound DC and p_1=0.75 100 random initializations					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^2$	$\hat{\lambda}$	$\hat{p}_1$
	$\sigma_v = 1\ and\ \lambda = 10$				
Mean	4.9635	2.0135	1.0385	10.4274	0.7542
Standard deviation	0.2615	0.0624	0.4906	2.4547	

Table 4.5: (continued)

$\sigma_v = 1 \ and \ \lambda = 5$					
Mean	4.9477	2.0139	0.9929	5.0756	0.7562
Standard deviation	0.2375	0.0589	0.1437	0.6795	
$\sigma_v = 1.25 \ and \ \lambda = 10$					
Mean	4.9466	2.0148	1.5517	10.4962	0.7515
Standard deviation	0.2796	0.0702	0.1826	2.5383	
$\sigma_v = 1.25 \ and \ \lambda = 5$					
Mean	4.9419	2.0161	1.5603	5.1251	0.7514
Standard deviation	0.2791	0.0685	0.2703	0.8517	

Furthermore Table 4.6 shows that the number of replications removed from all cases is between 12 and 33. More generally, it can be stated that the replications removed represent the 2.5-6.5% of the replications which is a small percentage.

Table 4.6: Replications removed from the 100 random initializations technique

Number of replications removed		
	$\lambda = 10$	$\lambda = 5$
$\sigma_v = 1$	12	16
$\sigma_v = 1.25$	33	22

Finally, Table 4.7 reports the bias of the mean WTP, given by Eq. (4.20), together with the mean estimate of the WTP. The bias of the mean WTP estimate is very small for all cases, almost zero which means that the mean estimates are almost identical to the real values of the parameters.

Table 4.7: Bias of the mean WTP from the 100 random initializations technique

Bias of Mean WTP same probability for overestimation for all respondents 100 random initializations		
	$\sigma_v = 1$ *and* $\lambda = 10$	$\sigma_v = 1$ *and* $\lambda = 5$
Mean $\widehat{\mathbf{WTP}}_{\text{mixture model}}$	13.0115	12.9974
$\mathbf{Bias}_{\text{mixture model}}$	0.0175	0.0034
	$\sigma_v = 1.25$ *and* $\lambda = 10$	$\sigma_v = 1.25$ *and* $\lambda = 5$
Mean $\widehat{\mathbf{WTP}}_{\text{mixture model}}$	13.0003	13.0002
$\mathbf{Bias}_{\text{mixture model}}$	0.0063	0.0062

C. Initial values obtained from k-means algorithm

The results for parameter estimates when the k-means algorithm is used for initialization are illustrated in Table 4.8 and the number of replications that were removed for each case is illustrated in Table 4.9. Additionally, Table 4.10 reports the bias of the mean WTP and the mean estimate of the WTP.

As was the case for the previous two initialization techniques, the results show that for all cases the mean estimates of all parameters are very close to the real parameters. Additionally the standard deviation of the estimates is small except for parameter λ. The parameter λ in cases where its real value is equal to 10 has higher standard deviation due to extreme values that may occur during the estimation.

Table 4.8: Simulation Results for the k-means initializations strategy

Estimation Results for Double-Bounded DC and p_1=0.75 initialization with k-means					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_v^2$	$\hat{\lambda}$	$\hat{p}_1$
$\sigma_v = 1 \; and \; \lambda = 10$					
Mean	4.9598	2.0131	1.0143	10.0244	0.7526
Standard deviation	0.2516	0.0629	0.3385	2.3945	
$\sigma_v = 1 \; and \; \lambda = 5$					
Mean	4.9474	2.0144	0.9927	5.0772	0.7498
Standard deviation	0.2393	0.0596	0.1446	0.6958	
$\sigma_v = 1.25 \; and \; \lambda = 10$					
Mean	4.9544	2.0144	1.5801	10.7918	0.7518
Standard deviation	0.2843	0.0698	0.4114	3.4944	
$\sigma_v = 1.25 \; and \; \lambda = 5$					
Mean	4.9455	2.0161	1.57189	5.1266	0.7508
Standard deviation	0.2752	0.0681	0.3274	0.8746	

To continue with the analysis, Table 4.9 illustrates the number of replications that have been removed from each case due to standard error issues. The replications that have been removed are very few, the number of them ranges from 1 to 8 replications. In terms of percentage, the percentage of the replications that have been removed is smaller than 2%, which is a small proportion of the total replications.

Table 4.9: Replications removed from the k-means initialization technique

Number of replications removed		
	$\boldsymbol{\lambda = 10}$	$\boldsymbol{\lambda = 5}$
$\boldsymbol{\sigma_v = 1}$	1	1
$\boldsymbol{\sigma_v = 1.25}$	8	4

Table 4.10: Bias of the mean WTP from the k-means initialization technique

Bias of Mean WTP K-means initialization		
	$\boldsymbol{\sigma_v = 1\ and\ \lambda = 10}$	$\boldsymbol{\sigma_v = 1\ and\ \lambda = 5}$
$\mathbf{Mean\ W\widehat{T}P_{mixture\ model}}$	13.0062	12.999
$\mathbf{Bias_{mixture\ model}}$	0.0122	0.005
	$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 10}$	$\boldsymbol{\sigma_v = 1.25\ and\ \lambda = 5}$
$\mathbf{Mean\ W\widehat{T}P_{mixture\ model}}$	13.006	13.0039
$\mathbf{Bias_{mixture\ model}}$	0.012	0.0099

Finally Table 4.10 presents the mean estimated WTP and the bias of the mean WTP for each case, given by applying Eq. (4.20). Considering the results it can be noticed that the bias of the mean WTP for each case, estimated by the proposed mixture model, is very small, since the mean estimate of the WTP is very close to the real WTP.

Concluding the findings above it can be stated that the three initialization techniques produce very similar parameter estimates. However the k-means initialization technique has fewer replications with standard error issues. Additionally it can be seen that the standard deviation from the 100 random initialization technique, in almost all cases is smaller compare to the other initialization methods. Besides these small differences, all three methods can initialize the EM algorithm effectively and the results show that the proposed mixture model can overcome hypothetical bias in a satisfactory way in order to provide trustworthy estimates.

4.4. Comparison of initialization methods

In the previous section the three initialization techniques were compared in order to examine which technique returns better estimates to minimize the bias of the mean WTP. In the present section, additional performance criteria are going to be applied in order detect if there is an initialization strategy that performs better.

4.4.1. Methodology for comparing initialization methods

In order to compare the initialization techniques a number of performance criteria were applied. Examining in detail the literature that refers to comparison of initialization methods, the performance criteria that are commonly used in such researches are presented below.

One practical criterion for algorithm comparison is the running time (Meilă and Heckerman, 2001:16). Another criterion used is the number of iterations that are necessary until convergence. This criterion was performed by Karlis and Xekalaki (2003:580). Another criterion to examine the performance of an initialization strategy is its ability to reach the Global maximum. Karlis and Xekalaki (2003:581), assume that at least one of the methods has reached the global maximum.

In order to assume that each j^{th} set of initial values has succeeded in locating the global maximum two conditions given by Eq. (4.21) and Eq. (4.22), must be satisfied:

$$\text{Condition (1): } \left|\theta_{max} - \theta_j\right| < 10^{-5} \qquad (4.21)$$

and

$$\text{Condition (2): } \left|\frac{(L_{max} - L_j)}{L_{max}}\right| < 10^{-5} \qquad (4.22)$$

where L_{max} denotes the log-likelihood with the maximum value, considered as the global maximum and L_j is the value of the log-likelihood for the j-th set of initial values. Additionally, θ_{max} denotes the parameters corresponding to the global maximum.

In the present study since three techniques are going to be compared, for each technique the log-likelihood value is determined at each replication and L_{max} is determined as the log-likelihood with the maximum value among the three. Furthermore, as θ_{max} is considered the parameter estimates of the technique with the maximum log-likelihood.

Furthermore, another criterion is based on the ability to find the correct classification. In order to determine the ability to determine the right classification the measure that is used is the Adjusted Rand Index (ARI). The ARI has "a maximum value of 1 indicating identical solutions" and measures the accuracy of the classifications (Maruotti and Punzo, 2021:455-56). The ARI is included in the "fossil" package and was applied by using the adj.rand.index( ) function right after the classes were separated (Vavrek, 2020:33). The Rand Index was proposed by Rand in 1971 and in 1985 Huber and Arabie proposed the Adjusted Rand Index (Vavrek, 2020:33). The Adjusted Rand Index rescaled the Rand Index by taking into consideration that "random cause will cause some objects to occupy the same clusters" (Vavrek, 2020:33).

The Adjustment Rand Index is calculated by

$$ARI(P^*, P) = \frac{\sum_{ij}\binom{N_{ij}}{2} - \left[\sum_i\binom{N_i}{2}\sum_j\binom{N_j}{2}\right]/\binom{N}{2}}{\frac{1}{2}\left[\sum_i\binom{N_i}{2} + \sum_j\binom{N_j}{2}\right] - \left[\sum_i\binom{N_i}{2}\sum_j\binom{N_j}{2}\right]/\binom{N}{2}} \qquad (4.23)$$

Where

$P^* = \{C_1^*, \ldots, C_K^*\}$ is the partition of the data set based on the ground truth

$P = \{C_1, \ldots, C_K\}$ is the clustering results generated by the clustering algorithm

K the number of clusters

N is the number of data points in a given dataset

N_{ij} is the number of data points of the class label $C_j^* \in P^*$ assigned to cluster C_i in partition P.

N_i is the number of data points in cluster C_i of partition P and

N_j is the number of data points in class C_j^* (Yang, 2017:31).

Finally, two additional criteria are added. The proportion that condition 1 given by Eq. (4.21) is satisfied as well as the condition 2, given by Eq. (4.22).

4.4.2. Comparison results

In order to determine the technique that provides the best initial values for the EM algorithm, it should be in mind that the notion of "the best initialization method" involves a trade-off between computation cost and accuracy. Additionally the best initial values are an ill-defined notion since a formal delimitation among initial searches doesn't' exist (Meilă and Heckerman, 2001:14-5). The following Tables illustrate the results obtained by applying the performance criteria described in the previous subsection.

To begin with, Table 4.11 reports the proportion of the times that each initialization technique obtained the maximum log-likelihood value (column 1). It should be noted that in a number of replications there was more than one technique with the maximum log-likelihood value thus the percentages do not sum to 100%. The 100 Random Initializations method had a very high percentage of maximum log-likelihood values. More specifically, in the case where $\sigma_v = 1$ and $\lambda = 10,$ 67.81% of the replications the 100 Random Initializations Technique returned the highest log-likelihood value among the other techniques. Furthermore, in the remaining cases, 100 Random Initializations method has as well the higher log-likelihood value among the other methods for more than 50% of the replications.

Column 2 of Table 4.11 shows the percentage of times a method finds the global maximum or in other words satisfies both conditions given in Eq. (4.21) and Eq. (4.22). Note that condition (4.21) needs to hold for all parameters. As it can be seen the success rate for the 1 Random Initialization technique is between 24% and 33% of the total replications. The K-means algorithm has a very similar percentage of succeeding in reaching the global maximum since it ranges from 23% to 37%. On the other hand the 100 Random Initializations technique has a higher percentage of reaching the global maximum, since it varies between 50% and 70%, which is almost the double percentage of the other two methods.

Additionally, Table 4.11 presents the results for each condition separately. As it is shown, for each parameter separately that composes the global maximum conditions, it can be noticed that the parameters α, β and σ_v^2 have a high percentage of satisfying condition 1 given by Eq. (4.21). The parameter λ on the other hand although satisfies Eq. (4.21) it can be noticed that the percentage is smaller than the other parameters' percentages.

Although for all parameters and methods condition 1 is satisfied over 20% of the time, the 100 Random Initializations technique has a substantial advantage of achieving percentages more than 50%. Furthermore, in regards of succeeding in condition (2) separately, in all cases the initialization methods reached over 99% of satisfying Eq. (4.22).

Table 4.11: Comparing the initialization methods for Double-Bounded DC (1)

Comparing the initialization methods for Double-Bounded DC and p_1=0.75 (1)							
	% max log-likelihood	% global maximum	% $condition$ (1)				% $condition$ (2)
			$\hat{a}$	$\hat{\beta}$	$\widehat{\sigma_v^2}$	$\hat{\lambda}$	
$\sigma_v = 1$ and $\lambda = 10$							
1Random Init	18.24%	24.68%	41.63%	58.37%	45.49%	24.89%	99.14%
100Random Init	67.81%	68.45%	74.46%	86.48%	78.54%	69.74%	99.36%
K-means	15.45%	23.18%	40.56%	56.01%	42.27%	23.18%	99.57%
$\sigma_v = 1$ and $\lambda = 5$							
1Random Init	48.08%	30.13%	45.73%	60.9%	45.3%	39.53%	99.57%
100Random Init	70.3%	50.85%	66.45%	79.06%	65.17%	55.98%	100%
K-means	49.36%	32.05%	45.09%	63.25%	46.79%	41.67%	100%
$\sigma_v = 1.25$ and $\lambda = 10$							
1Random Init	19.1%	33.71%	48.99%	70.34%	48.54%	33.93%	98.88%
100Random Init	54.83%	55.28%	71.69%	86.52%	68.99%	55.96%	100%
K-means	26.07%	36.63%	53.26%	72.58%	49.21%	37.53%	99.78%
$\sigma_v = 1.25$ and $\lambda = 5$							
1Random Init	14.86%	26.61%	39.91%	60.31%	36.36%	28.6%	100%
100Random Init	69.84%	70.07%	76.05%	85.14%	74.94%	70.73%	100%
K-means	15.96%	25.72%	39.25%	60.31%	38.14%	27.72%	100%

Table 4.12 presents the remaining criteria that have been considered during the comparison of the methods. To begin with, one parameter that is examined in the comparison of the initialization strategies is the time that is needed in order to complete the estimation. In order to enable the time needed for the comparison, the models were run for the same number of replications. Additionally the time was calculated from the beginning of the procedure until the end of the total replications. The 100 Random Initializations technique is an extension of the 1Random Initialization and therefore the time needed will be approximately 100 times more. This technique is not taken into consideration in regards to the comparison of the time needed, thus the time needed to complete the 100 Random Initializations is not illustrated in Table 4.12. As it is shown in the first column of Table 4.12, the time that is needed in order to complete the process for the 1 Random Initialization and the k-means are very close. In the case where $\sigma_v = 1$ and $\lambda = 10$ the time needed is the same, in the rest of the cases the differences between the techniques is around 1 minute.

Another criterion that is taken under consideration is the number of replications that have been removed due to standard error issues. As it is shown in Table 4.12, in 1 Random Initialization and 100 Random Initializations techniques the replications that have been removed are about the same and vary from 13 to 25. On the other hand with k-means technique the replications removed were fewer and more specifically it can be noticed that the replications removed were 1 to 6 which are very few compared to the other two strategies.

Another performance criterion refers to the number of iterations that the estimation process runs until convergence. Table 4.12 reports the mean number of iterations needed for each technique. It can be observed that across cases the number of iterations needed doesn't vary. More specifically, the 1 Random Initialization and the k-means technique needed the same number of iterations for almost all cases. Furthermore the 100 Random Initializations technique doesn't differ from the other techniques much. More specifically the 100 Random Initializations technique has either the same number of iterations as the other techniques or it differs by 1 iteration. It should be noted that the number of iterations is calculated after the best of the 100.

The last performance criterion is the Adjusted Rand Index (ARI) which measures the accuracy of the classification. The criterion was applied in order to have a measurement of which criterion separates the two classes more accurately. As it is illustrated in Table 4.12, the three strategies have very similar results and there is no strategy that receives an ARI score over 60%. Although none of the initialization techniques received a high level for classification accuracy, it can be noticed that 100 Random Initializations technique achieved the highest score for three out of the four cases.

Table 4.12: Comparing the initialization methods for Double-Bounded DC (2)

Comparing the initialization methods for Double Bounded DC and p_1=0.75 (2)				
	Running Time in minutes	**Replications removed**	**Number of iterations**	**ARI**
$\sigma_v = 1\ and\ \lambda = 10$				
1Random Init	37.89	20	40	50.36%
100Random Init	-	13	38	52.57%
K-means	37.42	1	39	46.92%
$\sigma_v = 1\ and\ \lambda = 5$				
1Random Init	39.27	15	38	48.54%
100Random Init	-	16	38	44.24%
K-means	40.52	1	38	44.74%
$\sigma_v = 1.25\ and\ \lambda = 10$				
1Random Init	32.66	20	37	50.34%
100Random Init	-	22	36	57.19%
K-means	31.25	6	37	48.17%
$\sigma_v = 1.25\ and\ \lambda = 5$				
1Random Init	34.34	25	37	49.26%
100Random Init	-	22	38	55.79%
K-means	36.35	4	37	46.8%

Finally, in order to complete the comparison of the initialization techniques Table 4.13 presents the estimates of the mean WTP and the bias of the mean WTP for all cases and techniques. As it can be seen, the bias of the mean WTP for each technique is very small. Although the differences within the techniques are very small, the 100 Random Initializations has a slightly better performance in the majority of the examined cases, in overcoming hypothetical bias.

Table 4.13: Comparison of the Bias of the mean WTP

Comparison of the Bias of Mean WTP			
	1 Random Initialization	**100 Random Initializations**	**K-means**
$\sigma_v = 1$ *and* $\lambda = 10$			
Mean $\widehat{WTP}$	13.0111	13.0115	13.0062
Bias	0.0171	0.0175	0.0122
$\sigma_v = 1$ *and* $\lambda = 5$			
Mean $\widehat{WTP}$	13.0011	12.9974	12.999
Bias	0.0071	0.0034	0.005
$\sigma_v = 1.25$ *and* $\lambda = 10$			
Mean $\widehat{WTP}$	13.0139	13.0003	13.006
Bias	0.0199	0.0063	0.012
$\sigma_v = 1.25$ *and* $\lambda = 5$			
Mean $\widehat{WTP}$	13.0081	13.0002	13.0039
Bias	0.014	0.0062	0.0099

Conclusions

The initial values of the EM algorithm play a very crucial role since selecting starting values refers as a “well-documented” drawback of the method (Panić et al., 2020:1). In the present research, in order to overcome hypothetical bias a double-bounded mixture model has been proposed. Furthermore the present chapter examined the starting values issue thus different initialization methods were examined in order to determine the technique that provides the more proper starting values for the proposed mixture model.

After 500 replications and examination of different cases the three chosen initialization techniques have provided interesting findings. At first the main goal of overcoming hypothetical bias is achieved by all initialization techniques. By comparing the simulation results and the performance criteria it can be concluded that the three techniques achieve accurately outcomes. Table 4.14 summarizes analytically which technique performed better for each performance criterion. As is reported in Tables 4.11, 4.12 and 4.13 the differences are very small in the majority of the performance criteria that have been applied.

In order to determine which technique performed better at each criterion it was taken into account whether it obtained better results in most of the cases and more specifically a technique performed better compared to another one if it had better outcomes at least in 3 out of the 4 cases. As is illustrated in the following table, the initialization method that performed better in the majority of the criteria is the 100 Random Initializations technique.

Table 4.14: Technique that performed better at each criterion

Technique that performed better at each criterion		
Criterion applied	**Technique that performed better**	**Technique with the second best performance**
% maximum log-likelihood	100 Random Initializations	K-means
% global maximum	100 Random Initializations	1 Random Initialization and K-means (50-50)
Time	1 Random Initialization and K-means	-
Iterations	100 Random Initializations	1 Random Initialization
Replications removed	K-means	1 Random Initialization and 100 Random Initializations
ARI	100 Random Initializations	1 Random Initialization
Bias of the mean WTP	100 Random Initializations	K-means

Summarizing all the above, the 100 Random Initializations technique performed better for the majority of the criteria but the main drawback of this method is that it needs a lot of time compared to the other methods. Thus if time is not an issue the 100 Random Initializations technique is the preferred technique, on the other hand if the time is limited the other two techniques can determine as well proper starting values in order to overcome hypothetical bias.

References of Chapter 4

Biernachi, C., 2004, “Initializing EM using the properties of its trajectories in Gaussian mixtures”, *Statistics and Computing,* vol.14, p.267-279

Biernacki C., Celeux G. and Govaert G.,2003, “Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models”, *Computational Statistics and Data Analysis*, vol. 41, p.561–575.

Brun, M., Sima, C., Hua, J., Lowey, J., Carroll, B, Suh, E. and Dougherty, E R., 2007, “Model-based evaluation of clustering validation measures”, *The Journal of the Pattern Recognition Society,* vol. 40, p. 807-824

Chen, W-C., Maitra, R, Melnykov, V., Nettleton, D., Faden, D., Rostamian, R. and R Core team, 2021, EMCluster: EM Algorithm for Model-Based Clustering of Finite Mixture Gaussian Distribution, R package version 0.-13

Chien, Y-L., Huang, C. J. and Shaw, D., 2005, “A general model of starting point bias in double-bounded dichotomous contingent valuation surveys”, *Journal of Environmental Economics and Management,*vol.50, p.362-377

Fraley, C., Raftery, A. E., Scrucca, L., Murphy, T. B. and Fop, M., 2012, mclust: Gaussian Mixture Modelling for Model-Based Clustering, Classification, and Density Estimation, R package version 5.4.10

Hipp, J. R. and Bauer, D., J., 2006, “Local Solutions in the Estimation of Growth Mixture Models”, *Phychological Methods,* vol. 11, no. 1 p. 36-53

Hofler, R.A. and List, J.A., 2004, “Valuation on the frontier: Calibrating actual and hypothetical statements of value”, *American Journal of Agricultural Economics,*vol.86, no.1, p.213-221

Karlis, D. and Xekalaki, E., 2003, “Choosing initial values for the EM algorithm for finite mixtures”, *Computational Statistics and Data Analysis,* vol. 41, p.577-590

Kassambara, A., 2017. *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning*. Sthda.com

Kumbhakar, S.C. and Lovell, K.C.A., 2000, *Stochastic Frontier Analysis*, United States of America, Cambridge University Press

Kumbhakar, S.C., Parmeter C.F. and Tsionas E.G., 2013, “A zero inefficiency stochastic frontier model”,*Journal of Econometrics*, vol.172, p.66-76

Leroux, B.G., 1992, “Consistent Estimation of a Mixing Distribution”, *The Annals of Statistics,* vol.20, no.3, p.1350-1360

Mann, A.K. and Kaur, N., 2013, “Review Paper on Clustering Techniques”, *Global Journal of Computer Science and Technology,* vol.13,no.5-C,[pdf], available at: <https://computerresearch.org/index.php/computer/article/view/353>[access 31 March 2022]

Maruotti, A.and Punzo,A., 2021, “Initialization of Hidden Markov and Semi-Markov Models: A Critical Evaluation of Several Strategies”, *International Statistical Review,* vol.89, no.3, p. 447-480

McLachlan, G.T., 1988, “On the Choice of Starting Values for the EM Algorithm in Fitting Mixture Models”, *Journal of the Royal Statistical Society, Series D (The Statistician),* vol. 37, no.4/5, p.417-425

Meilă , M. and Heckerman,D., 2001, “An Experimental Coparison of Model-Based Clustering Methods”, *Machine Learning,* vol. 42, no.9, p.9-29

Panić, B., Klemenc, J. and Nagode, M., 2010, “Improved Initialization of the EM Algorithm for Mixture Model Parameter Estimation”, *Mathematics,* vol.8 no.3, p. 373, available at: <https://www.mdpi.com/2227-7390/8/3/373/htm>. [access 19 May 2022]

Pocuca, N., Browne, R. P. and McNicholas, P. D., 2021, mixture: Mixture Models for Clustering and Classification, R package version 2.0.4

R Core Team, (2022), R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rand, W.M., 1971, “Objective criteria for the evaluation of clustering methods”, *Journal of the American Statistical Association*, vol. 66, p.846–850.

Shireman, E., Steinley, D. and Brusco, M.J., 2017, “Examining the effect of initialization strategies on the performance of Gaussian mixture modeling”, *Behavior Research Methods,* vol.49, p.282–293

Tsay, W-J.,Huang C.J., Fu, T-T. and Ho, I-L., 2013, “a simple closed –form approximation for the cumulative distribution function of the composite error of stochastic frontier model” , *J Prod Anal*,vol.39, p.259-269

Vavrek, M. J., 2020, fossil: Palaeoecological and Palaeogeographical Analysis Tools, R package version 0.4.0

Woodward, W.A., Parr, W.C., Schucany, W.R. and Lindsey, H., 1984, “A comparison of Minimum Distance and Maximum Likelihood Estimation of a Mixture Proportion”, *Journal of the American Statistical Association,* vol.79, no.387, p.590-598

Young, D., Benaglia,T., Chauveau, D., Hunter, D., Elmore, R., Hettmansperger, T., Thomas, H. and Huan, F., 2020, mixtools: Tools for Analyzing Finite Mixture Models, R package version 1.2.0