ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΛΟΓΙΚΗΣ ΚΑΙ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Ανάλυση Εσόδων σε Επαναλαμβανόμενες Δημοπρασίες Πρώτης Τιμής με Αλγορίθμους Άμεσης Μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΒΑΣΙΛΙΚΗ ΜΠΑΤΣΑΡΗ

Επιβλέπων : Δημήτριος Φωτάκης
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2024

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών
Εργαστήριο Λογικής και Επιστήμης Υπολογιστών

Revenue Analysis in Repeated First-Price Auctions with Online Learning Bidders

DIPLOMA THESIS

VASILIKI BATSARI

Επιβλέπων : Δημήτριος Φωτάκης
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 1η Νοεμβρίου 2024.

....	....	....
Δημήτριος Φωτάκης	Ευστράτιος Σκουλάκης	Αριστείδης Παγουρτζής
Καθηγητής Ε.Μ.Π.	Επίκουρος Καθηγητής Aarhus University	Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2024

....
Βασιλική Μπάτσαρη
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Περίληψη

Οι δημοπρασίες αποτελούν θεμελιώδεις μηχανισμούς που χρησιμοποιούνται ευρέως σε πολλούς τομείς, από την τέχνη και τις αγορές ηλεκτρικής ενέργειας, μέχρι το ηλεκτρονικό εμπόριο και τη διαφήμιση. Ούσες κρίσιμα εργαλεία για την κατανομή πόρων και τη συσσώρευση εσόδων, έχουν μελετηθεί εκτενώς τόσο στη θεωρία όσο και στην πράξη. Ιστορικά, οι δημοπρασίες δεύτερης τιμής έχουν κυριαρχήσει, λόγω της ύπαρξης κυρίαρχης στρατηγικής για τους ενδιαφερόμενους αγοραστές. Ωστόσο, τα τελευταία χρόνια παρατηρείται μια αξιοσημείωτη στροφή στη βιομηχανία προς τις επαναλαμβανόμενες δημοπρασίες πρώτης τιμής. Η παρούσα διπλωματική διερευνά την ανάλυση εσόδων σε αυτό το πλαίσιο, εστιάζοντας σε ανεξάρτητους και πανομοιότυπα κατανεμημένους αγοραστές που προσομοιώνονται από αλγορίθμους άμεσης μάθησης.

Η διπλωματική εξετάζει τρία βασικά ερωτήματα: πρώτον, διερευνά την αποτελεσματικότητα ενός "τεχνητού" αγοραστή—ο οποίος συμμετέχει στη δημοπρασία με στόχο την αύξηση της τιμής πώλησης του αντικειμένου—ως υποκατάστατο μιας παραδοσιακής ελάχιστης τιμής. Δεύτερον, αναλύει τις πιθανές διαφορές στα έσοδα μεταξύ αγοραστών που χρησιμοποιούν no-regret και no-swap regret αλγορίθμους. Τέλος, εξετάζει τη σχέση μεταξύ των προκυπτόντων εσόδων από επαναλαμβανόμενες δημοπρασίες πρώτης τιμής και των αναμενόμενων εσόδων από δημοπρασίες δεύτερης τιμής υπό τις ίδιες Μπεϋζιανές συνθήκες. Τα πειραματικά μας ευρήματα καταδεικνύουν ότι ενώ η συμπερίληψη ενός "τεχνητού" αγοραστή αυξάνει τα έσοδα, η επίδρασή του στην αύξηση αυτή είναι μειωμένη σε σχέση με την επίδραση μιας βέλτιστης ελάχιστης τιμής. Επί προσθέτως, δεν παρατηρείται σημαντικό πλεονέκτημα, όσον αφορά στα έσοδα, μεταξύ της χρήσης no-regret έναντι no-swap regret αλγορίθμων ή το αντίστροφο. Τέλος, η σύγκριση των εσόδων μεταξύ δημοπρασιών πρώτης και δεύτερης τιμής δεν οδηγεί σε σαφές συμπέρασμα, καθώς τα αποτελέσματα διαφοροποιούνται ανά περίπτωση.

Λέξεις Κλειδιά

Κυρίαρχη Στρατηγική, Επαναλαμβανόμενες Δημοπρασίες, Δημοπρασίες Πρώτης Τιμής, Αλγόριθμοι 'Αμεσης Μάθησης, No-Regret Αλγόριθμοι, Ελάχιστη Τιμή

Abstract

Auctions are fundamental mechanisms widely used across various domains, from art and electricity markets to e-commerce and advertising. As critical tools for resource allocation and revenue generation, they have been extensively studied in both theory and practice. Historically, second-price auctions have been favored for their ease of comprehension and control, owing to the existence of a dominant strategy for bidders. However, in recent years, there has been a notable industry shift towards repeated first-price auctions. This thesis investigates revenue generation within this framework, focusing on independent and identically distributed (i.i.d.) bidders simulated by online learning algorithms.

The study addresses three main questions: first, it explores the efficacy of a "fake" bidder—an artificial agent inserted into the auction by the seller to raise prices—as a substitute for a traditional reserve price. Second, it analyzes potential differences in revenue between no-regret and no-swap regret bidders. Finally, it examines the relationship between the revenue generated from repeated first-price auctions with the expected revenue from second-price auctions under the same Bayesian conditions. Our experimental findings indicate that while including a "fake" bidder enhances revenue, its impact does not fully replicate that of an optimal reserve price. Furthermore, no significant revenue advantage is observed between no-regret and no-swap regret algorithms. Lastly, the comparison of revenue between first-price and second-price auctions does not lead to a definitive conclusion, as the results remain scenario-dependent.

Keywords

Dominant Strategy, Repeated Auctions, First-price auctions, Online Learning Algorithms, No-Regret Algorithms, Reserve Price

Ευχαριστίες

Η ολοκλήρωση της διπλωματικής ορίζει το τέλος ενός μεγάλου κεφαλαίου: της πολυετούς φοίτησης στην σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών (ΗΜΜΥ). Σηματοδοτεί ωστόσο και την αρχή ενός νέου μεγάλου κεφαλαίου, της ζωής μετά το προπτυχιακό-μεταπτυχιακό. Πολλές οι σκέψεις και τα συναισθήματα που κατακλύζουν το μυαλό μου σε αυτή τη μετάβαση... Δεν έχω αποφασίσει ποιο θα είναι το επόμενό μου βήμα, ωστόσο ξέρω με σιγουριά ότι στα φοιτητικά μου χρόνια υπήρξαν άτομα που αποτέλεσαν έμπνευση για μένα και έβαλαν το λιθαράκι τους στο να γοητευτώ από τον τομέα των διακριτών μαθηματικών, των αλγορίθμων, της θεωρητικής πληροφορικής· εκεί που τα μαθηματικά παύουν να είναι αφηρημένα και εναρμονίζονται με το φυσικό κόσμο και τη λογική. Συνεπώς, αξίζει τουλάχιστον μια μικρή αναφορά σε αυτά τα άτομα.

Θα ήθελα πρωτίστως να ευχαριστήσω τον καθηγητή Δημήτρη Φωτάκη, όχι μόνο για την επίβλεψη της διπλωματικής μου, τη στήριξη και την επίλυση όλων των αποριών μου—που δεν είναι καθόλου αμελητέα—αλλά και για κάτι ακόμη πιο σημαντικό για μένα: την αγάπη του για τον τομέα της θεωρητικής πληροφορικής, μια αγάπη κι ένα πάθος που το μοιράζεται, το μεταδίδει στους μαθητές του. Νομίζω ότι η επιστήμη, ανεξαρτήτως τομέα, κρύβει ένα είδος μαγείας μέσα της, και το όλο νόημα έγκειται στην ανακάλυψη αυτής της μαγείας. Ο κύριος Φωτάκης δίνει αυτή την ευκαιρία στους φοιτητές και γι' αυτό τον λόγο είναι και ο αγαπημένος μου καθηγητής. Είναι μεγάλη μου χαρά που έκανα την διπλωματική μου μαζί του.

Επι προσθέτως, θα ήθελα να ευχαριστήσω τον Στρατή Σκουλάκη, ο οποίος ξεκίνησε να επιβλέπει την διπλωματική μου ως μεταδιδακτορικός ερευνητής στο EPFL, στη Λωζάνη, και ολοκλήρωσε την επίβλεψη ως επίκουρος καθηγητής στο Aarhus University! Ο Στρατής πρότεινε το θέμα της διπλωματικής, έχοντας βαθειά γνώση του τομέα και όντας ικανός να απαντήσει οποιουδήποτε είδους απορία μου δημιουργούνταν. Η συμβολή του ήταν πολύτιμη και μέσω των τακτικών διαδικτυακών συναντήσεων που είχα μαζί του και με τον κύριο Φωτάκη έμαθα πολλά πράγματα και ξεδιάλυνα σημαντικά ερωτήματα. Το πηγαίο ενδιαφέρον του Στρατή για επιστημονικά ερωτήματα που προέκυπταν ανά καιρούς με έκανε να συνειδητοποιήσω ακόμη περισσότερο το πόσο όμορφη και δημιουργική είναι η ενασχόληση με την έρευνα.

Τέλος, θα ήθελα να ευχαριστήσω το οικογενειακό και φιλικό μου περιβάλλον, το οποίο πάντα με στήριζε και με στηρίζει σε κάθε μου βήμα, ό,τι κι αν επιλέξω. Αυτή η υποστήριξη είναι πολύτιμη· είμαι πολύ τυχερή που έχω αυτούς τους ανθρώπους στη ζωή μου.

Περιεχόμενα

Περίληψη **5**

Abstract **7**

Ευχαριστίες **9**

Περιεχόμενα **11**

Εκτεταμένη Ελληνική Περίληψη **13**
0.1 Θεωρία παιγνίων 15
0.2 Δημοπρασίες από την πλευρά του πωλητή 15
0.3 Αλγόριθμοι Άμεσης Μάθησης 16
0.4 Ενώνοντας τις τελείες 17
0.5 Πειραματικά Αποτελέσματα 18
0.5.1 Η Επιρροή ενός "Τεχνητού" Αγοραστή 18
0.5.2 No-Regret Έναντι No-Swap Regret 18
0.5.3 Σχέση Παραγόμενου Εισοδήματος Μεταξύ Επαναλαμβανόμενων Δημοπρασιών Πρώτης και Δεύτερης Τιμής 19

Κείμενο στα Αγγλικά **23**

1 Introduction **23**
1.1 Game Theory 24
1.2 Revenue in Auctions 25
1.3 Online Learning Software Agents 26
1.4 Putting Everything Together 26

2 Game Theory and Equilibrium Concepts **29**
2.1 Introduction to Game Theory 29
2.2 Equilibria in Games 30
2.2.1 Pure Nash Equilibrium (PNE) 30
2.2.2 Mixed Nash Equilibrium (MNE) 31
2.2.3 Correlated Equilibrium (CE) 31
2.2.4 Coarse Correlated Equilibrium (CCE) 33
2.2.5 Understanding equilibria with an example 33

3 Auctions: A Bidding Game **43**
3.1 Auctions as Games — The Setting . . . 43
3.2 Second-Price Auction . . . 45
3.3 The Vickrey Auction . . . 46
3.4 Allocation Rule: Implementable and Monotone . . . 46
3.5 Bayesian Setting . . . 47
3.6 Bayes-Nash Equilibrium . . . 48
3.7 Revenue Maximization . . . 49
3.7.1 Expected Revenue Equals Expected Virtual Welfare . . . 50
3.7.2 Repeated Auctions . . . 51

4 Online Learning Algorithms **53**
4.1 Notation and Setting . . . 53
4.2 The Notion of External Regret . . . 54
4.2.1 Weighted Majority (WM) . . . 56
4.2.2 Randomized Weighted Majority (RWM) . . . 58
4.2.3 Hedge . . . 59
4.3 The Notion of Swap Regret . . . 61
4.3.1 Reduction from Swap Regret to External Regret . . . 62
4.4 Linking Regret Minimization with Equilibrium Concepts . . . 64

5 Our Contribution **67**
5.1 Existing Literature . . . 67
5.2 Experimental Research . . . 69
5.2.1 Deterministic Valuations in First- and Second-Price Auctions - Reproduction of Results . . . 70
5.2.2 "Fake" Bidder as a Reserved Price . . . 71
5.2.3 Revenue Evolution within the Bayesian Setting . . . 79
5.2.4 No-Regret Versus No-Swap Regret . . . 79
5.2.5 Revenue of Bayesian First-Price Auctions vs Expected Second-Price 81
5.2.6 Summary of Our Experimental Results . . . 82

Bibliography **83**

Εκτεταμένη Ελληνική Περίληψη

Οι δημοπρασίες αποτελούν αναπόσπαστο κομμάτι του ανθρώπινου πολιτισμού από την αρχαιότητα έως σήμερα. Οι πρώτες καταγεγραμμένες δημοπρασίες ανάγονται περίπου στο 500 π.Χ. στη Βαβυλώνα [1]. Στη ρωμαϊκή εποχή, οι στρατιώτες διεξήγαγαν δημοπρασίες καρφώνοντας ένα δόρυ στο έδαφος για να σηματοδοτήσουν την πώληση πολεμικών λαφύρων—μια πρακτική γνωστή ως «sub hasta», που σημαίνει «κάτω από τη λόγχη» [1], [2]. Στην Κίνα, ιστορικά αρχεία καταδεικνύουν ότι ήδη από τον 7ο αιώνα μ.Χ. πραγματοποιούνταν δημοπρασίες για την πώληση περιουσιών αποθανόντων βουδιστών μοναχών [3]. Παρόλο που τα αντικείμενα που τίθενται σε δημοπρασία έχουν αλλάξει με την πάροδο του χρόνου, ο ρόλος και η σημασία των δημοπρασιών όχι μόνο παραμένουν ακέραιοι, αλλά έχουν επεκταθεί ουσιαστικά.

Οι σύγχρονες δημοπρασίες λειτουργούν σε ένα ευρύ φάσμα πλαισίων, κανόνων και διαδικασιών. Ανεξαρτήτως της μορφής τους, αποτελούν σημαντικό μέρος της σύγχρονης κοινωνίας. Παρά το γεγονός ότι η πλειονότητα των ανθρώπων δεν συμμετέχει σε δημοπρασίες σε καθημερινή βάση, οι τελευταίες ασκούν βαθιά επιρροή σε διάφορους τομείς υψίστης οικονομικής σημασίας. Οι δημοπρασίες επηρεάζουν ένα ευρύ φάσμα δραστηριοτήτων, από τις αγορές ηλεκτρικής ενέργειας [4], [5], [6] και την τέχνη [7], έως τις κυβερνητικές και πολιτικές διαδικασίες [8], [9]. Δημοπρασίες χρησιμοποιούνται επίσης σε λιγότερο αναμενόμενους τομείς, όπως οι διαδικασίες πτώχευσης [10], η διαχείριση της συμφόρησης στα αεροδρόμια [11], η πώληση συλλεκτικών γραμματοσήμων [12], αλλά και η εμπορία κρασιού [13]. Η άνοδος του διαδικτύου έχει ενισχύσει περαιτέρω την επιρροή των δημοπρασιών, επιτρέποντας τη συμμετοχή σε παγκόσμια κλίμακα για ένα ευρύ φάσμα αγαθών και υπηρεσιών. Το ηλεκτρονικό εμπόριο έχει φέρει επανάσταση στον τρόπο διεξαγωγής των δημοπρασιών [14], [15], [16]. Μάλιστα, η διαδικτυακή διαφήμιση, η οποία στηρίζεται σε μεγάλο βαθμό σε μηχανισμούς δημοπρασιών, αποτελεί κρίσιμο παράγοντα όσον αφορά στη μεγιστοποίηση εσόδων, τόσο για τις διαδικτυακές πλατφόρμες, όσο και για τους δημιουργούς περιεχομένου [17], [18], [19], [20], [21]. Όλα τα παραπάνω καταδεικνύουν τον σημαντικό οικονομικό αντίκτυπο που έχουν οι δημοπρασίες στον σύγχρονο κόσμο. Τέλος, αξίζει να σημειωθεί ότι ο William Vickrey τιμήθηκε το 1996 με το Βραβείο Νόμπελ Οικονομικών Επιστημών για την καθοριστική συμβολή του στη θεωρία των δημοπρασιών, γεγονός που αναδεικνύει τη σημασία του πεδίου αυτού.

Ο όρος «δημοπρασία» προέρχεται από τη λατινική λέξη «auctus», η οποία αποτελεί μετοχή του ρήματος «augeō», που σημαίνει «αυξάνω». Οι δημοπρασίες λειτουργούν ως μηχανισμοί όπου αγαθά ή υπηρεσίες πωλούνται στον πλειοδότη μέσω ανταγωνιστικών προσφορών. Με την πάροδο του χρόνου, έχουν αναδειχθεί διάφορες μορφές δημοπρασιών, κάθε μία με τη δική της ξεχωριστή δομή και τις δικές της στρατηγικές επιπτώσεις.

Σε μια δημοπρασία, κάθε συμμετέχων κατέχει μια ιδιωτική αποτίμηση (private valuation) του αντικειμένου που δημοπρατείται, η οποία αντιπροσωπεύει το μέγιστο ποσό που είναι διατεθειμένος να πληρώσει. Ο βασικός στόχος των ενδιαφερόμενων αγοραστών είναι να αποκτήσουν το αντικείμενο στην χαμηλότερη δυνατή τιμή. Πιο φορμαλιστικά, κάθε πιθανός αγοραστής επιδιώκει να μεγιστοποιήσει το όφελός του (utility), το οποίο ορίζεται ως η διαφορά μεταξύ της ιδιωτικής του αποτίμησης και της τιμής που πληρώνει αν κερδίσει· σε διαφορετική περίπτωση, το όφελός του ισούται με μηδέν.

Ένας από τους πιο διαδεδομένους τύπους δημοπρασιών είναι η αγγλική δημοπρασία, στην οποία οι ενδιαφερόμενοι αγοραστές ανταγωνίζονται αυξάνοντας σταδιακά τις προσφορές τους. Η διαδικασία συνεχίζεται έως ότου κανείς δεν είναι διατεθειμένος να υπερβεί την τρέχουσα προσφορά και ο ενδιαφερόμενος αγοραστής με την υψηλότερη προσφορά κερδίζει το αντικείμενο. Αυτή η μορφή χρησιμοποιείται συχνά, μεταξύ άλλων, σε δημοπρασίες έργων τέχνης, πωλήσεις ζώων φάρμας και φιλανθρωπικές εκδηλώσεις [22], [23], [24]. Αντίθετα, τομείς όπως η διαδικτυακή διαφήμιση και το ηλεκτρονικό εμπόριο χρησιμοποιούν συχνά δημοπρασίες με σφραγισμένες προσφορές (sealed-bid), σύμφωνα με τις οποίες οι ενδιαφερόμενοι υποβάλλουν τις προσφορές τους ιδιωτικά και ταυτόχρονα, χωρίς να γνωρίζουν τις προσφορές των ανταγωνιστών τους [25], [26].

Σε ορισμένα περιβάλλοντα δημοπρασιών, δημοπρατούνται ταυτόχρονα πολλά αντικείμενα· για παράδειγμα, στη διαδικτυακή διαφήμιση, μπορεί να διατίθενται διάφορες διαφημιστικές θέσεις σε μια σελίδα αναζήτησης. Σε αυτές τις περιπτώσεις, ο πλειοδότης εξασφαλίζει την πιο κερδοφόρα θέση στη σελίδα, ενώ ο ενδιαφερόμενος αγοραστής με την δεύτερη υψηλότερη προσφορά αποκτά την αμέσως επόμενη καλύτερη θέση, κ.ο.κ. [25], [26], [27]. Ωστόσο, το απλούστερο σενάριο αποτελεί η δημοπρασία ενός μοναδικού αντικειμένου (single-item), κατά την οποία δημοπρατείται ένα (αδιαίρετο) αντικείμενο. Με αυτή την περίπτωση ασχολείται η παρούσα διπλωματική. Μια άλλη σημαντική διάκριση μεταξύ των τύπων δημοπρασιών αφορά στην τιμή πώλησης του αντικειμένου, δηλαδή στο ποσό που πληρώνει ο νικητής. Στη δημοπρασία πρώτης τιμής (first-price), ο πλειοδότης πληρώνει το ακριβές ποσό της προσφοράς του. Αυτή η δομή προτρέπει τους ενδιαφερόμενους να ενεργούν στρατηγικά, προσφέροντας ποσά ελαφρώς μικρότερα από την ιδιωτική τους αποτίμηση, προκειμένου να αυξήσουν το πιθανό τους κέρδος [28], [29]. Αντίθετα, στη δημοπρασία δεύτερης τιμής, ο πλειοδότης πληρώνει τη δεύτερη υψηλότερη προσφορά. Αυτή η μέθοδος δίνει κίνητρο στους ενδιαφερόμενους να προσφέρουν ποσό ίσο με τις ιδιωτικές τους αποτιμήσεις, καθώς η πληρωμή τους δεν εξαρτάται από τη δική τους προσφορά, αλλά από την προσφορά του δεύτερου υψηλότερου ενδιαφερόμενου αγοραστή [29]. Πιο φορμαλιστικά, σε μια δημοπρασία δεύτερης τιμής, η υποβολή προσφοράς ίσης με την ιδιωτική αποτίμηση αποτελεί κυρίαρχη στρατηγική για κάθε πιθανό αγοραστή. Για την καλύτερη κατανόηση αυτής της ορολογίας, όπως και του ευρύτερου πλαισίου των δημοπρασιών από μια πιο φορμαλιστική, μαθηματική σκοπιά, είναι απαραίτητη η μελέτη ορισμένων εννοιών της θεωρίας παιγνίων, η οποία προσφέρει ένα αυστηρό και μαθηματικό πλαίσιο για την ανάλυση στρατηγικών αλληλεπιδράσεων, όπως αυτές που παρατηρούνται στις δημοπρασίες.

0.1 Θεωρία παιγνίων

Οι δημοπρασίες μπορούν να θεωρηθούν ένας συγκεκριμένος τύπος στρατηγικού παιγνίου, στο οποίο κάθε συμμετέχων επιδιώκει να μεγιστοποιήσει το όφελός του. Κατά συνέπεια, ο τομέας της θεωρίας παιγνίων, ο οποίος αναλύει μαθηματικά τα παίγνια και τις έννοιες της ισορροπίας, συνδέεται με τη θεωρία δημοπρασιών προσφέροντας πολύτιμα αποτελέσματα. Η ισορροπία σε ένα παίγνιο αντιπροσωπεύει μια κατάσταση όπου κάθε παίκτης έχει επιτύχει το μέγιστο δυνατό όφελος, δεδομένων των ενεργειών των άλλων παικτών. Παρέκκλιση από την επιλεγμένη ενέργεια δεν θα οδηγούσε σε περαιτέρω κέρδη για κανέναν παίκτη, γεγονός που υποδηλώνει ότι κανένας παίκτης δεν μπορεί να επωφεληθεί από την αλλαγή της στρατηγικής του, εφόσον οι στρατηγικές των άλλων παικτών παραμένουν σταθερές. Η στρατηγική κάθε παίκτη αναφέρεται στην επιλεγμένη ενέργεια/δράση του από ένα σύνολο πιθανών ενεργειών/δράσεων. Μια κυρίαρχη στρατηγική προκύπτει όταν ένας παίκτης διαθέτει μια συγκεκριμένη ενέργεια που εγγυάται το καλύτερο δυνατό αποτέλεσμα για αυτόν, ανεξαρτήτως των ενεργειών των άλλων παικτών [30]. Στην περίπτωση μιας δημοπρασίας δεύτερης τιμής, η προσφορά της ιδιωτικής αποτίμησης του παίκτη αναδεικνύεται ως κυρίαρχη στρατηγική για κάθε παίκτη, καθώς οποιαδήποτε παρέκκλιση από αυτή τη στρατηγική δεν αυξάνει το όφελός του. Μια πληρέστερη επεξήγηση ορισμένων εννοιών ισορροπίας παρουσιάζεται στο Κεφάλαιο 2, το οποίο περιλαμβάνει επίσης ένα ενδεικτικό παράδειγμα. Επιστρέφοντας στις δημοπρασίες, είναι ουσιώδους σημασίας να αναγνωριστεί η σημασία τους ως στρατηγικά παιχνίδια με σημαντικές οικονομικές επιπτώσεις.

0.2 Δημοπρασίες από την πλευρά του πωλητή

Όπως έχει αναφερθεί πρωθύστερα, υπάρχουν διάφοροι τύποι δημοπρασιών, καθένας από τους οποίους διαθέτει μοναδικά χαρακτηριστικά. Μία κρίσιμη πτυχή όλων των μορφών δημοπρασιών είναι η παραγωγή εσόδων. Ενώ οι ενδιαφερόμενοι αγοραστές επιδιώκουν να μεγιστοποιήσουν το όφελός τους, ο πωλητής επιδιώκει να πωλήσει το αντικείμενο στην μέγιστη δυνατή τιμή. Η επιλογή της μορφής δημοπρασίας δύναται να επηρεάσει το ύψος των εσόδων για τους πωλητές, καθώς τα πιθανά κέρδη διαμορφώνονται από τις στρατηγικές που επιλέγουν οι ενδιαφερόμενοι αγοραστές. Οι δημοπρασίες δεύτερης τιμής τείνουν να αποφέρουν πιο προβλέψιμα έσοδα, καθώς οι ενδιαφερόμενοι υποβάλλουν ειλικρινείς προσφορές, δηλαδή προσφέρουν ποσά ίσα με τις πραγματικές ιδιωτικές τους αποτιμήσεις. Ιστορικά, αυτές οι δημοπρασίες έχουν προτιμηθεί, καθώς οδηγούν σε αναμενόμενα έσοδα, και σημαντική έρευνα έχει αφιερωθεί σε αυτόν τον τύπο δημοπρασίας [31], [32], [33], [34]. Αντίθετα, οι δημοπρασίες πρώτης τιμής είναι λιγότερο προβλέψιμες ως προς τα πιθανά κέρδη [35].

Ένας επιπλέον παράγοντας που προσθέτει πολυπλοκότητα στην ανάλυση των δημοπρασιών είναι το γεγονός ότι αυτές δεν διεξάγονται πάντα σε ένα ντετερμινιστικό περιβάλλον, όπου κάθε ενδιαφερόμενος διαθέτει μια σταθερή ιδιωτική αποτίμηση. Σε ορισμένες περιπτώσεις, οι δημοπρασίες λειτουργούν σε ένα Μπεϋζιανό πλαίσιο, όπου οι ιδιωτικές αποτιμήσεις των ενδιαφερόμενων προέρχονται από μια κατανομή πιθανοτήτων. Μία κοινή παραδοχή είναι ότι αυτή η κατανομή είναι γνωστή τόσο στον πωλητή όσο και στους άλλους ενδιαφερόμενους αγοραστές. Διαισθητικά, αυτή η γνωστή κατανομή αντιπροσωπεύει τις πληροφορίες που κατέχει ο πωλητής σχετικά με τις πιθανές αποτιμήσεις του αντικειμένου που δημοπρατείται.

Δεδομένης της ύψιστης οικονομικής σημασίας των δημοπρασιών, η κατανόηση του τρόπου με τον οποίο οι πωλητές μπορούν να βελτιστοποιήσουν τις μορφές δημοπρασίας τους, ώστε να μεγιστοποιήσουν τα κέρδη τους, έχει μελετηθεί εκτενώς. Οι ερευνητές έχουν καταλήξει στην ύπαρξη ισοδυναμίας, όσον αφορά στα έσοδα, μεταξύ δημοπρασιών πρώτης και δεύτερης τιμής—σε Bayes-Nash ισορροπία—όταν οι προσωπικές αποτιμήσεις των ενδιαφερόμενων αγοραστών είναι ανεξάρτητες και πανομοιότυπα κατανεμημένες τυχαίες μεταβλητές [36], δηλαδή στην περίπτωση που κάθε ενδιαφερόμενος επιλέγει την αποτίμησή του ανεξάρτητα από τους άλλους αγοραστές, αλλά από την ίδια κοινή κατανομή.

Μια άλλη σημαντική έννοια που σχετίζεται με τη μεγιστοποίηση των εσόδων είναι η ελάχιστη τιμή. Σε πολλές μορφές δημοπρασιών, οι πωλητές καθορίζουν μια ελάχιστη τιμή—ένα ελάχιστο όριο κάτω από το οποίο το αντικείμενο δεν διατίθεται προς πώληση. Στο πλαίσιο των δημοπρασιών δεύτερης τιμής, ο καθορισμός μιας ελάχιστης τιμής συνεπάγεται ότι εάν μόνο ένας πιθανός αγοραστής υποβάλει προσφορά που υπερβαίνει αυτό το ελάχιστο όριο, θα πληρώσει για να αγοράσει το αντικείμενο ποσό ίσο με την ελάχιστη τιμή, αντί της δεύτερης υψηλότερης προσφοράς, όπως στην κλασική περίπτωση. Έρευνες έχουν αποδείξει ότι η χρήση βέλτισης ελάχιστης τιμής μπορεί να αυξήσει τα έσοδα του πωλητή [37], [38].

Με την πρόοδο της τεχνολογίας και την ανάπτυξη της διαδικτυακής διαφήμισης, ο τομέας των επαναλαμβανόμενων δημοπρασιών έχει αποκτήσει μεγάλη σημασία. Σε αυτό το πλαίσιο, δημοπρασίες για το ίδιο αντικείμενο (π.χ. μια συγκεκριμένη διαφημιστική θέση σε έναν ιστότοπο) διεξάγονται πολλές φορές, επιτρέποντας την επαναλαμβανόμενη συμμετοχή των ίδιων πιθανών αγοραστών. Αυτή η επαναλαμβανόμενη αλληλεπίδραση δίνει τη δυνατότητα στους αγοραστές να παρακολουθούν τη συμπεριφορά των ανταγωνιστών τους, τις προηγούμενες προσφορές τους και να προσαρμόζουν τις δικές τους προσφορές αναλόγως, μετατρέποντας τη δημοπρασία σε μια συνεχή στρατηγική αλληλεπίδραση. Ταυτόχρονα, παρατηρείται μια αξιοσημείωτη μετατόπιση από τις δημοπρασίες δεύτερης τιμής στις δημοπρασίες πρώτης τιμής στη διαδικτυακή διαφήμιση. Η τάση αυτή σηματοδοτήθηκε ιδιαίτερα από τη μετάβαση του Google Ad Exchange από τις δημοπρασίες δεύτερης τιμής στις δημοπρασίες πρώτης τιμής το 2019, υπογραμμίζοντας μια ευρύτερη κίνηση στον κλάδο [39], [35], [40], [41].

Υπάρχει εκτενής έρευνα στον τομέα των επαναλαμβανόμενων δημοπρασιών [42], [43], [44], [45]. Πώς όμως αναλύουν οι ερευνητές τις διαδικασίες μάθησης και τα αποτελέσματα των επαναλαμβανόμενων δημοπρασιών; Συχνά χρησιμοποιούν προσομοιώσεις για να μοντελοποιήσουν τον τρόπο με τον οποίο οι ενδιαφερόμενοι αγοραστές μαθαίνουν και προσαρμόζουν τις προσφορές τους με την πάροδο του χρόνου. Στις περισσότερες περιπτώσεις, οι πιθανοί αγοραστές αναπαρίστανται από no-regret αλγορίθμους. Για να κατανοηθεί πλήρως η σημασία και επίδραση αυτής της προσέγγισης, είναι απαραίτητο να εξερευνηθεί η έννοια των αλγορίθμων άμεσης μάθησης.

0.3 Αλγόριθμοι Άμεσης Μάθησης

Όπως αναφέρθηκε προηγουμένως, οι ενδιαφερόμενοι αγοραστές συχνά προσομοιώνονται από υπολογιστές οι οποίοι χρησιμοποιούν αλγορίθμους άμεσης μάθησης. Μια κεντρική έννοια στον κλάδο των αλγορίθμων αυτών είναι η «μέτανοια» (regret). Η έννοια της (εξωτερικής) μετάνοιας (external regret) αναφέρεται στη διαφορά μεταξύ του καλύτερου δυνατού οφέλους που θα μπορούσε να έχει επιτύχει ένας αγοραστής επιλέγοντας σε κάθε γύρο σταθερά μια συγκεκριμένη ενέργεια—την καλύτερη δυνατή ενέργεια, αν γνώριζε εκ των προτέρων τις στρατηγικές των ανταγωνιστών του για όλους τους γύρους—και του πραγματικού οφέλους που έλαβε. Η μετάνοια μετρά ουσιαστικά τον βαθμό στον οποίο ένας αγοραστής «μετανιώνει» που δεν επέλεξε τη βέλτιστη στρατηγική από την αρχή. Ένας αλγόριθμος χωρίς (εξωτερική) μετάνοια (no-regret) στοχεύει στην ελαχιστοποίηση αυτής της διαφοράς με την πάροδο του χρόνου, γεγονός που σημαίνει ότι, καθώς επαναλαμβάνεται η δημοπρασία, οι εκάστοτε ενέργειες του αγοραστή αποδίδουν σχεδόν τόσο καλά όσο η βέλτιστη σταθερή ενέργεια. Με άλλα λόγια, με τους αλγορίθμους χωρίς μετάνοια, η μέση τιμή της μετάνοιας τείνει στο μηδέν καθώς αυξάνεται ο αριθμός των γύρων της δημοπρασίας, υποδεικνύοντας ότι το όφελος του αγοραστή προσεγγίζει το όφελος της καλύτερης σταθερής στρατηγικής [46].

Αυτή η έννοια είναι ζωτικής σημασίας σε σενάρια άμεσης μάθησης, όπου οι συμμετέχοντες προσαρμόζουν τις στρατηγικές τους με βάση παρελθοντικά δεδομένα ώστε να λαμβάνουν όλο και πιο κερδοφόρες αποφάσεις με την πάροδο του χρόνου. Στην παρούσα διπλωματική αναλύονται αλγόριθμοι που ελαχιστοποιούν τη μετάνοια και διερευνάται η επιρροή τους όσον αφορά στα έσοδα του πωλητή. Πιο συγκεκριμένα, χρησιμοποιούνται τόσο αλγόριθμοι no-external μετάνοιας, όσο και αλγόριθμοι no-swap μετάνοιας (no-swap regret), και παρατηρείται η διαφορά μεταξύ τους. Η no-swap μετάνοια αποτελεί έναν ελαφρώς διαφορετικό και ισχυρότερο ορισμό της μέτανοιας, όπως αυτή δόθηκε πρωθύστερα· και οι δύο τύποι αλγορίθμων αναλύονται περαιτέρω στο Κεφάλαιο 4.

0.4 Ενώνοντας τις τελείες

Η θεωρία δημοπρασιών συνδέεται στενά με τη θεωρία παιγνίων, καθώς οι δημοπρασίες είναι ουσιαστικά ένα είδος στρατηγικού παγνίου και, κατά συνέπεια, οι έννοιες της ισορροπίας, που κυριαρχούν στον κόσμο των παιγνίων, τις διέπουν. Η εμφάνιση επαναλαμβανόμενων δημοπρασιών προσθέτει ένα επιπλέον επίπεδο πολυπλοκότητας, οδηγώντας στην χρήση αλγορίθμων άμεσης μάθησης για την προσομοίωση της συμπεριφοράς των ενδιαφερόμενων αγοραστών. Η παρούσα διπλωματική εξετάζει πτυχές της θεωρίας παιγνίων, των μηχανισμών δημοπρασιών και των αλγορίθμων άμεσης μάθησης, εστιάζοντας στον τρόπο με τον οποίο αυτά τα στοιχεία αλληλεπιδρούν και επηρεάζουν τα πιθανά έσοδα σε περιβάλλοντα επαναλαμβανόμενων δημοπρασιών πρώτης τιμής. Πιο συγκεκριμένα, η παρούσα διπλωματική αποσκοπεί στην επίτευξη των ακόλουθων στόχων: πρώτον, την διερεύνηση του κατά πόσο ένας "τεχνητός" αγοραστής—ο οποίος συμμετέχει στη δημοπρασία για λογαριασμό του πωλητή— μπορεί να λειτουργήσει αποτελεσματικά ως βέλτιστη ελάχιστη τιμή, δηλαδή κατά πόσο δύναται να αυξήσει τα έσοδα του πωλητή με τρόπο συγκρίσιμο με μια παραδοσιακή ελάχιστη τιμή· δεύτερον, την παρατήρηση πιθανών διαφορών μεταξύ των εσόδων που παράγονται κατά την εφαρμογή no-regret και no-swap regret αλγορίθμων· τέλος, την διερεύνηση της διαφοράς των εσόδων που παράγονται από δημοπρασίες πρώτης τιμής και των αναμενόμενων εσόδων από δημοπρασίες δεύτερης τιμής, δεδομένου ότι και οι δύο τύποι δημοπρασιών λειτουργούν υπό τις ίδιες Μπεϋζιανές συνθήκες, όσον αφορά στις ιδιωτικές αποτιμήσεις των ενδιαφερόμενων αγοραστών.

Με την καθιέρωση αυτών των θεμελιωδών εννοιών και τη σκιαγράφηση των πρωταρχικών στόχων της παρούσας διπλωματικής, μπορούμε πλέον να προβούμε σε μια σύνοψη των αποτελεσμάτων. Περαιτέρω εμβάθυνση σε κάθε συστατικό στοιχείο της παρούσας διπλωματικής—θεωρία παιγνίων, θεωρία δημοπρασιών, αλγορίθμους άμεσης μάθησης—καθώς και ανάλυση των πειραματικών ευρημάτων παρουσιάζονται εν συνεχεία, σε κείμενο γραμμένο στην αγγλική γλώσσα.

0.5 Πειραματικά Αποτελέσματα

Τα πειραματικά αποτελέσματα της παρούσας διπλωματικής παρουσιάζονται αναλυτικά στο Κεφάλαιο 5 στην αγγλική γλώσσα. Ωστόσο, στο παρόν κεφάλαιο γίνεται μια συνοπτική περίληψη αυτών στα ελληνικά.

0.5.1 Η Επιρροή ενός "Τεχνητού" Αγοραστή

Κύριος στόχος της παρούσας πειραματικής μελέτης είναι η διερεύνηση του κατά πόσο ένας "τεχνητός" αγοραστής θα μπορούσε να συμβάλλει στη μεγιστοποίηση των εσόδων, λειτουργώντας ως βέλτιστη ελάχιστη τιμή. Ο συγκεκριμένος αγοραστής συνεργάζεται με τον πωλητή και σκοπός του είναι να ωθήσει τους υπόλοιπους πιθανούς αγοραστές να αυξήσουν τις προσφορές τους· δεν έχει πρόθεση να αγοράσει το αντικείμενο. Για τον υπολογισμό του οφέλους (utility) του "τεχνητού" συμμετέχοντα στη δημοπρασία, εφαρμόστηκαν οι εξής δύο προσεγγίσεις:

Απλός αγοραστής: Σε αυτό το σενάριο, εάν ο "τεχνητός" αγοραστής κερδίσει τη δημοπρασία, το όφελός του είναι μηδενικό. Αντίθετα, εάν δεν κερδίσει το αντικείμενο—δηλαδή κάποιος άλλος ενδιαφερόμενος έχει καταθέσει ίση ή υψηλότερη προσφορά—το όφελός του ισούται με το ποσό της νικητήριας προσφοράς. Αυτή η προσέγγιση οδηγεί τον "τεχνητό" αγοραστή να επιλέγει τιμές που συνήθως "χάνουν", υποβάλλοντας με αυτό τον τρόπο πολύ χαμηλές προσφορές. Κατά συνέπεια, δεν καταφέρνει να παρακινήσει αρκετά αποτελεσματικά τους άλλους συμμετέχοντες ώστε να αυξήσουν τις δικές τους προσφορές.

Έξυπνος αγοραστής: Σε αυτή την προσέγγιση, ο "τεχνητός" αγοραστής και πάλι έχει μηδενικό όφελος εάν κερδίσει τη δημοπρασία, καθώς ο στόχος του δεν είναι αυτός. Αν όμως δεν κερδίσει το αντικείμενο, το όφελός του ισούται με το ποσό της προσφοράς του. Αυτό το σενάριο εξασφαλίζει ότι οι υψηλότερες—αλλά μη νικητήριες—προσφορές αποφέρουν μεγαλύτερο όφελος απ' ό,τι οι χαμηλότερες.

Συνοπτικά, τα ευρήματά μας καταδεικνύουν ότι η συμμετοχή ενός "τεχνητού" αγοραστή, αυξάνει τα έσοδα του πωλητή, αλλά όχι στον ίδιο βαθμό με την βέλτιστη ελάχιστη τιμή.

0.5.2 No-Regret Έναντι No-Swap Regret

Έπειτα από μια σειρά προσομοιώσεων επαναλαμβανόμενων δημοπρασιών πρώτης τιμής, στις οποίες οι αγοραστές μοντελοποιήθηκαν αρχικά με no-regret και στη συνέχεια με no-swap regret αλγορίθμους, δεν προέκυψε κάποιο πλεονέκτημα από τη χρήση των no-regret αλγορίθμων έναντι των no-swap regret ή το αντίστροφο, όσον αφορά στο παραγόμενο εισόδημα. Τα αποτελέσματα διαφοροποιούνται ανά περίπτωση.

0.5.3 Σχέση Παραγόμενου Εισοδήματος Μεταξύ Επαναλαμβανόμενων Δημοπρασιών Πρώτης και Δεύτερης Τιμής

Το τελευταίο κύριο ερώτημα της παρούσας διπλωματικής αφορά το κατά πόσο το παραγόμενο εισόδημα στις επαναλαμβανόμενες δημοπρασίες πρώτης τιμής, όπου οι ενδιαφερόμενοι αγοραστές μοντελοποιούνται με no-swap regret (ή no-regret) αλγορίθμους, είναι ισοδύναμο με το αναμενόμενο εισόδημα των αντίστοιχων δημοπρασιών δεύτερης τιμής. Τα πειραματικά αποτελέσματα καταδεικνύουν ότι η σχέση αυτή δεν είναι ξεκάθαρη, καθώς ανάλογα με τις πιθανές ιδιωτικές αποτιμήσεις των αγοραστών, το αναμενόμενο εισόδημα από τις δημοπρασίες δεύτερης τιμής μπορεί να υπερβαίνει το παραγόμενο εισόδημα των δημοπρασιών πρώτης τιμής, σε άλλες περιπτώσεις το αντίστροφο, ενώ σε ορισμένες περιπτώσεις τα δύο αποτελέσματα συγκλίνουν.

Κείμενο στα Αγγλικά

Chapter 1

Introduction

Auctions have been a part of human civilization from ancient times to the present day. The earliest recorded auctions date back to around 500 BC in Babylon [1]. Roman soldiers also conducted auctions, driving a spear into the ground to signify that spoils of war were being sold, a practice known as "sub hasta", meaning "under the spear" [1], [2]. In China, records suggest that as early as the 7th century AD, auctions were held to sell the possessions of deceased Buddhist monks [3]. Although the items auctioned have changed over time, the role and significance of auctions not only persist but have also expanded considerably.

Modern auctions are built on a wide range of frameworks, rules, and procedures. Despite the different formats, auctions are integral to contemporary society. Even though most people may not directly engage with auctions on a daily basis, they have a profound influence across diverse sectors with substantial economic importance. Auctions impact everything from electricity markets [4], [5], [6] and art [7], to governmental [8] and political processes [9]. Interestingly, auctions are also used in unexpected areas such as bankruptcy proceedings [10], airport congestion management [11], the sale of collectible postage stamps [12], and even wine trading [13]. The rise of the internet has amplified the influence of auctions, enabling global participation across a broad range of goods and services. E-commerce has revolutionized how auctions are conducted [14], [15], [16]. Online display advertising, which is predominantly driven by auction mechanisms, plays a pivotal role in generating revenue for online platforms and content creators [17], [18], [19], [20], [21]. This underscores the significant economic impact auctions have in today's world. To further emphasize their importance, it is worth mentioning that William Vickrey, a leading figure in auction theory, was awarded the 1996 Nobel Memorial Prize in Economic Sciences for his research in the field.

The term "auction" traces its roots to the Latin word "auctus", the past participle of the latin verb "augeō", which means "to increase". Auctions are mechanisms where goods or services are sold to the highest bidder through competitive offers. Over time, various auction formats have emerged, each with its own distinct structure and strategic implications. In any auction setting, each participant holds a **private valuation** of the item being auctioned, representing the maximum amount they are willing to pay. The primary objective for bidders is to acquire the item at the lowest possible price. Mathematically, a bidder seeks to maximize their **utility**, defined as the difference between their private valuation and the price they pay if they win; otherwise, their utility is zero.

One of the most well-known types of auctions is the **English** auction, in which bidders openly compete by gradually raising their bids. The process continues until no one is willing to outbid the current offer, and the highest bidder wins the item. This format is frequently used in art auctions, livestock sales, and charity events, among other areas [22], [23], [24]. In contrast, sectors such as online display advertising and e-commerce often utilize **sealed-bid** auctions, where bidders submit their offers privately and simultaneously, without knowing the bids of their competitors [25], [26].

In some auction contexts, multiple items are sold simultaneously; for instance, in online display advertising, various slots may be available. In such cases, the highest bidder might secure the most valuable slot, while the second-highest bidder may obtain the next best slot, and so on [25], [26], [27]. However, the simplest scenario is the **single-item** auction, where only one, undivided item is auctioned. This is the specific case that we address in this thesis. Another key distinction relates to the the price the winner pays. In a **first-price** auction, the highest bidder wins and pays the exact amount of their bid. This structure encourages bidders to act strategically, often lowering their bids slightly below their private valuation to enhance their potential profit [28], [29]. In contrast, a **second-price** auction allows the highest bidder to win, while only paying the second-highest bid. This approach incentivizes bidders to reveal their private true valuations, as their final payment is based on the next highest offer rather than their own bid [29]. More formally, in a second-price auction, bidding one's private valuation constitutes a dominant strategy for each bidder. To understand this terminology, as well as the overall auction landscape from a mathematical perspective, it is essential to explore certain aspects of the world of game theory, which provides a formal and mathematical framework for analyzing strategic interactions, including auction scenarios.

1.1 Game Theory

Auctions can be conceptualized as a specific type of strategic game in which each player strives to maximize their utility. Consequently, the field of game theory, which mathematically examines games and equilibrium concepts, intersects with auction theory to yield insightful results. An **equilibrium** in a game represents a state where each player has achieved the maximum utility possible, given the actions of other players. Deviating from their chosen strategy would not result in further utility gains for any player, meaning that no player can benefit from changing their strategy, while keeping the strategies of others constant. Each player's strategy refers to their selected action from a set of possible actions. A **dominant strategy** occurs when a player possesses a particular action that guarantees the best possible outcome, regardless of the actions of other players [30]. Thus, in a second-price auction, bidding one's private valuation emerges as a dominant strategy, as any deviation from this strategy will not increase a player's utility. A deeper understanding of key equilibrium concepts is presented in Chapter 2, which also provides an illustrative example. Returning to auctions, it is essential to recognize their significance as strategic games with substantial economic implications.

1.2 Revenue in Auctions

As previously discussed, there are multiple auction types, each with unique characteristics. A vital aspect of all auction formats is revenue generation. While bidders seek to maximize their utility, the seller—who owns the item for sale—must also be considered. The choice of auction format significantly influences revenue outcomes for sellers, as each format carries distinct implications for potential earnings, shaped by the strategies employed by bidders. Second-price auctions tend to yield more predictable revenue, as bidders bid truthfully by submitting their actual private valuations. Historically, these auctions were favored for their ability to provide controlled revenue outcomes, and substantial research has been dedicated to this auction type [31], [32], [33], [34]. In contrast, first-price auctions are less predictable [35].

Another factor that adds complexity to auction analysis is that auctions are not always conducted in a deterministic setting, where each bidder has a fixed private valuation. In some instances, auctions operate under a **Bayesian setting**, where bidders' private valuations are drawn from a probability distribution. A common assumption is that this distribution is known to both the seller and other bidders. Intuitively, this known distribution represents the information the seller has about the potential valuations for the item being auctioned.

Given the paramount importance of revenue, significant research has focused on understanding how sellers can optimize their auction formats. Researchers have concluded that in a Bayes-Nash equilibrium, when bidders' valuations are independently and identically distributed (i.i.d.) random variables—meaning each bidder draws their valuation from the same distribution—the revenue generated from a first-price auction with the optimal reserve price equals that of a second-price auction with the optimal reserve price [36].

Another important concept related to revenue maximization is the **reserve price**. In numerous auction formats, sellers establish a reserve price—a minimum threshold below which the item will not be sold. In the context of second-price auctions, setting a reserve price implies that if only one bidder submits a bid exceeding this minimum, they will pay the reserve price instead of the second-highest bid. Research has demonstrated that using a reserve price can enhance the seller's revenue [37], [38].

With advancements in technology and the growth of online advertising, the field of **repeated auctions** has gained prominence. In this context, auctions for the same item (e.g., a specific advertising slot on a website) occur multiple times, allowing repeated participation from the same bidders. This repeated interaction enables bidders to observe the behavior of their competitors through time and adjust their bids accordingly, transforming the auction into an ongoing strategic interaction rather than a one-time event. Concurrently, there has been a notable shift from second-price to first-price auctions in online advertising. This trend was particularly marked by Google Ad Exchange's transition from second-price to first-price auctions in 2019, signaling a broader move within the industry [39], [35], [40], [41].

A wealth of research exists in the area of repeated auctions [42], [43], [44], [45]. How do researchers analyze the learning processes and outcomes of repeated auctions? They often employ simulations to model how bidders learn and adapt over time. In most of the cases, bidders are represented by regret-minimizing agents. To fully grasp the implications of this approach, it is essential to explore the concept of online learning algorithms.

1.3 Online Learning Software Agents

As previously mentioned, bidders are often simulated by computers using online learning algorithms. In instances where the same auction occurs multiple times per day, these computer simulations reflect real-world dynamics, as actual bidders may not always be directly involved. A central concept in this context is "**regret**". In online learning, the notion of (external) regret refers to the difference between the best possible utility a bidder could have achieved from the best fixed action in hindsight—had they known the strategies of their competitors in advance—and the actual utility obtained. Regret measures how much a bidder "regrets" not choosing the optimal strategy from the beginning. A **no-(external) regret** algorithm aims to minimize this regret over time, meaning that as the auction is repeated, the chosen actions by the bidder perform nearly as well as the optimal fixed action. In other words, with no-regret algorithms, the average regret tends to zero as the number of auction rounds increases, indicating that the bidder's performance approaches that of the best fixed strategy in hindsight [46].

This concept is crucial in online learning scenarios, where agents aim to adapt their strategies based on past experiences to improve their decision-making over time. In this thesis, we analyze regret-minimizing algorithms and explore their implications for revenue outcomes. More precisely, both no-external regret and **no-swap regret** algorithms are used and the difference between them is measured. No-swap regret is a slightly different and stronger definition of regret than external regret, which is analyzed in Chapter 4.

1.4 Putting Everything Together

Auction theory is inherently linked to game theory, as auctions represent strategic games, allowing for the application of equilibrium concepts common to game theory. The emergence of repeated auctions adds another layer of complexity, wherein online learning algorithms are employed to simulate bidders' behavior over time. This thesis examines the intersection of game theory, auction mechanisms, and online learning algorithms, focusing on how these elements interact to influence outcomes in repeated first-price auction environments.

More specifically, this thesis aims to achieve the following objectives: first, to investigate whether a "fake" bidder—essentially an agent acting on behalf of the seller—can function effectively as a reserve price, thereby maximizing the seller's revenue in a manner comparable to a traditional reserve price. Second, to assess potential differences in revenue generated when applying no-regret algorithms in contrast to no-swap regret algorithms. Finally, to explore whether the revenue produced by first-price auctions aligns with the expected revenue of second-price auctions, given that both types operate under the same Bayesian conditions regarding bidders' valuations. In light of the previously mentioned revenue equivalence, this inquiry can be reframed as examining whether repeated first-price auctions converge to a Bayes-Nash Equilibrium.

With these foundational concepts established and the primary objectives of this thesis outlined, we can now delve deeper into each component—game theory, auction theory, online learning algorithms—as well as analyze our findings.

Chapter 2

Game Theory and Equilibrium Concepts

2.1 Introduction to Game Theory

We have all played games during our lives; from a very young age, we engage in such interactions. The word "game" is familiar to us and often evokes feelings of joy. Games are meant to be entertaining, but they are also competitive, offering rewards to winners. Some games require cooperation between players, while others do not. However, a common denominator exists in every game: each player wants to win, i.e., to maximize their profit.

While luck may influence some games, most involve strategic thought. This means that if a player chooses specific combinations of actions from a set of possible actions, their probability of winning increases. This thought process can be simple or complex, but it is always worth exploring. The fascinating field which studies these strategic interactions from a formal and mathematical perspective is called **Game Theory**. We focus on simultaneous move games, in which all players make their decisions at the same time. A more formal description of this type of game is presented below.

A **simultaneous move game** is a formal model of an interactive situation involving a finite number K of decision-makers, referred to as players. Each player $i \in [K]$ has their own finite set of possible actions, also called strategies, denoted by S_i. At each round, every player i selects and plays a strategy $s_i \in S_i$, doing so simultaneously with the other players. The strategies chosen by all players at a given round are represented by the vector $\mathbf{s} = (s_1, \ldots, s_k)$, where $\mathbf{s}$ is an element of the set S, with S being the cartesian product $S_1 \times \cdots \times S_k$. The vector $\mathbf{s}$ is referred to as the **strategy profile** or **outcome** of this round. Based on the strategy profile, each player receives a value, which can either be a cost (in cost-minimization games) or a utility (in utility-maximization games). For each player i, the cost function c_i and the utility function u_i are defined as $c_i : S \to \mathbb{R}$ and $u_i : S \to \mathbb{R}$ respectively. It follows that in a cost-minimization game, each player aims to minimize their cost, while in a utility-maximization game, each player seeks to maximize their utility.

Every cost-minimization game has a corresponding utility-maximization counterpart, as the cost and utility functions can be interchanged by considering $u_i(\mathbf{s}) = -c_i(\mathbf{s})$. In this formulation, the utility function becomes the negative of the cost function, effectively transforming the minimization problem into a maximization one. The cost or utility of each player is determined by the strategies chosen by all players in the game, not just by the individual player's own strategy.

This raises a fundamental question: is there a strategy profile where all players reach an equilibrium, a state in which each player is satisfied with their utility (or cost) given the strategies of the others, and thus has no incentive to deviate from their current strategy? In other words, is there a strategy profile where no player can improve their utility (or reduce their cost) by unilaterally changing their own strategy? This constitutes the conceptually simplest and strictest form of equilibrium—the Nash Equilibrium. Nevertheless, there are various types of equilibria, ranging from more stringent ones like Nash Equilibria to more flexible ones such as Correlated Equilibria. In the following sections, we will define and examine four different types of equilibria [47] to gain a deeper understanding of these concepts.

2.2 Equilibria in Games

We begin by introducing the concept of a player's dominant strategy. A player's strategy is considered **dominant** if it is at least as good as all other strategies, regardless of the strategies chosen by other players.

Definition 2.1. *In a utility-maximization game, a player i has a dominant strategy if following this strategy maximizes their payoff, irrespective of the strategies chosen by other players.*

If all players in every game possessed dominant strategies, the analysis of equilibria would be straightforward—perhaps to the point where further exploration would hardly be necessary. However, whether we desire it or not, it is relatively rare to find games where players have dominant strategies.

Having introduced the concept of dominant strategies, we will now delve into certain equilibrium concepts. The equilibria we examine are within the context of utility-maximization games, where each player seeks to maximize their individual payoff. For cost-minimization games, these equilibrium concepts can be adapted by simply reversing the inequalities employed in the utility-maximization framework.

2.2.1 Pure Nash Equilibrium (PNE)

The strictest equilibrium concept in Game Theory, as mentioned in the previous section, is the **Pure Nash Equilibrium** (PNE). In a Pure Nash Equilibrium, no player can increase their profit by unilaterally deviating from the chosen strategy profile. A formal definition of a PNE is provided below:

Definition 2.2. *A strategy profile $\mathbf{s}$ of a utility-maximization game is a* ***pure Nash equilibrium (PNE)*** *if for every player $i \in \{1, 2, \ldots, k\}$ and every unilateral deviation $s'_i \in S_i$,*

$$U_i(\mathbf{s}) \geq U_i(s'_i, \mathbf{s}_{-i}).$$

PNE are easy to understand and interpret but, unfortunately, do not exist in all games [48] [49].

2.2.2 Mixed Nash Equilibrium (MNE)

As mentioned earlier, the Pure Nash Equilibrium is straightforward and strict, but many games, such as the well-known Rock-Paper-Scissors, lack a pure strategy profile. However, according to Nash's Theorem, given a finite number of players and strategies, an equilibrium can always be achieved if players adopt a **mixed strategy**, meaning they randomize independently over their possible strategies [50]. This type of equilibrium is known as a **Mixed Nash Equilibrium** (MNE).

Before proceeding with the formal definition of a MNE, it is important to explain the symbol $\mathbf{s}_{-i}$, as it will be used frequently in the following discussions. Given a finite number of K players and a strategy profile $\mathbf{s} = (s_1, s_2, \ldots, s_i, \ldots, s_k)$, we denote by $\mathbf{s}_{-i}$ the strategy profile of all players except for player i. Specifically, $\mathbf{s}_{-i} = (s_1, s_2, \ldots, s_{i-1}, s_{i+1}, \ldots, s_k)$. After having clarified that, we can proceed with a formal definition of a MNE:

Definition 2.3. *Distributions $\sigma_1, \ldots, \sigma_k$ over strategy sets S_1, ..., S_k of a utility-maximization game constitute a **mixed Nash equilibrium (MNE)** if for every player $i \in \{1, 2, \ldots, k\}$ and every unilateral deviation $s'_i \in S_i$,*

$$\mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(\mathbf{s})] \geq \mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(s'_i, \mathbf{s}_{-i})], \tag{2.1}$$

where σ denotes the product distribution $\sigma_1 \times \cdots \times \sigma_k$.

In a MNE, each player's strategy is represented by a probability distribution over their possible actions, and these distributions are independent of one another. As a result, the overall strategy profile—the combination of all players' strategies—forms a product distribution. A product distribution means that the joint distribution of players' strategies can be expressed as the product of their individual distributions. It is important to note that this definition of a MNE considers pure-strategy unilateral deviations. However, allowing mixed-strategy unilateral deviations does not alter the definition.

As inferred from the previous definitions, every PNE is a special case of a MNE, where each player plays deterministically—assigning a probability of zero to all but one action, which has a probability of one. Finding a MNE, even in a two-player game, is often computationally intractable. Consequently, two more permissive equilibrium concepts have been proposed: the **Correlated Equilibrium** and the **Coarse Correlated Equilibrium**, both of which are computationally tractable.

2.2.3 Correlated Equilibrium (CE)

In the two Nash equilibria previously discussed, the players act independently, with no cooperation between them; their actions are not *correlated*. We will now introduce two equilibrium concepts that incorporate the notion of cooperation. First, we will provide a formal definition of a **Correlated Equilibrium** (CE), before explaining it further.

Definition 2.4. *A distribution σ on the set $S_1 \times \cdots \times S_k$ of outcomes of a utility-maximization game constitutes a **Correlated equilibrium (CE)** if for every player $i \in \{1, 2, \ldots, k\}$, strategy $s_i \in S_i$ and every deviation $s_i' \in S_i$,*

$$\mathbb{E}_{\mathbf{s}\sim\sigma}\left[U_i(\mathbf{s}) \mid s_i\right] \geq \mathbb{E}_{\mathbf{s}\sim\sigma}\left[U_i(s_i', \mathbf{s}_{-i}) \mid s_i\right] \tag{2.2}$$

A correlated equilibrium can be conceptualized as involving a trusted third party who participates in the game. This third party has access to the distribution σ over the possible outcomes, which is known to all players. The role of this third party is to sample an outcome $\mathbf{s}$ according to σ and privately suggest the strategy s_i to each player i, where $i \in \{1, 2, \ldots, k\}$. Each player can then decide whether to follow the suggested action or not. At the time of decision-making, a player i is aware of both the distribution σ and the suggested strategy s_i, and thus, forms a posterior distribution on the suggested strategies of other players. This posterior distribution refers to the player's updated belief about the strategies of the other players, after receiving the recommendation. In a CE, each player maximizes their expected payoff by following their suggested strategy. The expectation is conditioned on the information available to player i — σ and s_i — and assumes that the other players will follow their own recommended strategies $\mathbf{s}_{-i}$.

It is important to note that the distribution σ in a CE does not have to be a product distribution. In other words, the strategies selected by the players are not independent; they are, in fact, correlated. The joint distribution in a CE reflects the correlation between players' choices, which results from the recommendations made by the trusted third party. These recommendations are based on a joint probability distribution that cannot necessarily be decomposed into the product of individual distributions.

It follows that every MNE is a special case of a CE, where the distribution σ is a product distribution. Since a MNE exists in every game within our framework, it follows that every such game has also a CE. Fortunately, finding a CE in a game is computationally tractable.

An alternative but equivalent definition of a correlated equilibrium in a utility-maximization game is as follows:

Definition 2.5. *A distribution σ on the set $S_1 \times \cdots \times S_k$ of outcomes of a utility-maximization game constitutes a **Correlated equilibrium (CE)** if for every player $i \in \{1, 2, \ldots, k\}$, strategy $s_i \in S_i$ and every switching function $\delta : S_i \to S_i$,*

$$\mathbb{E}_{\mathbf{s}\sim\sigma}\left[U_i(\mathbf{s})\right] \geq \mathbb{E}_{\mathbf{s}\sim\sigma}\left[U_i(\delta(s_i), \mathbf{s}_{-i})\right] \tag{2.3}$$

This formulation will be particularly useful in the upcoming chapters, where we will merge the concept of correlated equilibria with a specific class of online learning algorithms known as no-swap regret algorithms.

We acknowledge that this definition may seem somewhat stringent, so an example is provided to clarify the meaning of the δ function, or at least offer some intuition behind it. Consider the Rock-Paper-Scissors game, where the set of actions is represented as $\{R, P, S\}$. According to the definition, a CE is reached if, for every player i, their utility cannot be increased if, every time they chose a specific action—whether R, P, or S—they had instead consistently chosen a different specific action, given that the other players stuck to their strategies. For instance, if every time player i selected R, they had chosen P instead, and/or if every time player i selected S, they had switched to P instead, their utility would not improve.

2.2.4 Coarse Correlated Equilibrium (CCE)

While CE already offers computational tractability, an 'even more tractable' concept has been proposed, known as the **Coarse Correlated Equilibrium** (CCE). Like CE, CCE relies on the cooperation between players' strategies. A formal definition of a CCE is given below:

Definition 2.6. *A distribution σ on the set $S_1 \times \cdots \times S_k$ of outcomes of a utility-maximization game constitutes a **Coarse Correlated equilibrium (CCE)** if for every player $i \in \{1, 2, \ldots, k\}$ and every unilateral deviation $s'_i \in S_i$,*

$$\mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(\mathbf{s})] \geq \mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(s'_i, \mathbf{s}_{-i})] \tag{2.4}$$

In a CCE, each player i is aware only of the overall distribution σ and not of the specific component s_i from the realization. In this setting, no player can improve their payoff by unilaterally deviating. This means that if a player i were to deviate, by consistently choosing a single action instead of following the distribution, their utility would not increase. Every CE is a special case of CCE; therefore, every game within our framework has a CCE as well.

2.2.5 Understanding equilibria with an example

We have now provided formal definitions of these four types of equilibria. However, these definitions may be difficult to grasp at first. To facilitate understanding, we will present an example that illustrates these concepts.

Consider the following game involving a network with a common source vertex s, a common sink vertex t and 6 parallel edges between s and t denoted by $E = \{0, 1, 2, 3, 4, 5\}$. Each edge represents a route from s to t. There are 4 players $A = \{1, 2, 3, 4\}$, each starting from s and aiming to reach t, while maximizing their points. Points are awarded based on the travel time each player takes to reach t. Specifically, each player earns $10 - m$ points, where m is the time in minutes it took them to reach t. The time m for a route is determined by the number of players choosing that route. For example, if players 1 and 2 choose route 1, player 3 chooses route 4 and player 4 chooses route 5, then players 1 and 2 will each take 2 minutes to reach t and therefore earn $10 - 2 = 8$ points each. Players 3 and 4, who take 1 minute, will earn $10 - 1 = 9$ points each. The goal of each player is to maximize their points, which corresponds to minimizing their travel time.

This scenario illustrates a utility-maximization game in which each player aims to maximize their points by minimizing their travel time, highlighting the principle that every utility-maximization game has an equivalent cost-minimization counterpart.

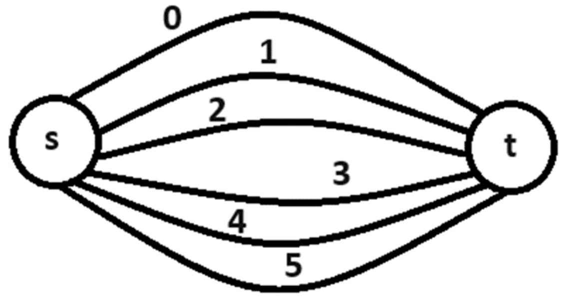


Figure 2.1: Game's Graph

Let's examine the four types of equilibria previously discussed:

PNE: In this game, the pure Nash equilibrium occurs when each route is chosen by at most one player. To determine the number of such equilibria, we consider the following: there are $\binom{6}{4} \times 4! = 15 \times 24 = 360$ such possible combinations, since there are $\binom{6}{4}$ ways to choose 4 out of the 6 available routes and for each combination of 4 routes, there are 4! ways to assign them to the 4 players. In each of these equilibria, every player chooses a different route, thus takes exactly 1 minute to reach the destination t and earns $10 - 1 = 9$ points. This is the maximum possible number of points a player can achieve, given the time constraints.

In each of these 360 combinations no player can increase their points by deviating to a different route, since all players are already earning the maximum possible points (9); therefore, each of those combinations consists of a pure Nash equilibrium.

Furthermore, there are no other PNE in this setup, because any other combination of strategies implies that at least two players choose the same route, experiencing a longer travel time (2 minutes or more) and as a consequence, earning less than 9 points. Therefore, each player would prefer to deviate to a route that is currently unchosen, which always exists as there are more routes than players.

MNE: A mixed Nash equilibrium in this game occurs when each player independently selects a route uniformly at random. In other words, in this MNE, every player chooses each of the 6 available routes with a probability of $\frac{1}{6}$. Under this strategy, the expected payoff for each player is $10 - \frac{3}{2} = \frac{17}{2}$, since the expected time to reach the destination is $\frac{3}{2}$. This can be calculated as follows:

Since there are 6 possible routes and 4 players, the total number of ways to assign the players to the routes is $6^4 = 1296$. These 1296 configurations can be categorized in the following cases:

- Case (1-1-1-1): Each player chooses a different route.
- Case (2-1-1): Two players choose the same route and the other two choose different routes.

- Case (2-2): Two players choose one route and two others choose a different route.
- Case (3-1): Three players choose the same route, while the remaining one selects a different one.
- Case (4): All players choose the same route.

Now, we calculate the expected time for player i to reach the destination by analyzing each case:

- (1-1-1-1): As explained in the PNE analysis, there are 360 ways in which each player chooses a different route. In all these cases, player i takes 1 minute to reach the destination, as no player shares their route.
- (2-1-1): There are $\binom{4}{2} \times 6 \times \binom{5}{2} \times 2! = 720$ possible configurations for this setup. This is because there are $\binom{4}{2}$ ways to select 2 players out of 4, and for each selected pair, there are 6 available routes to choose from. From the remaining 5 routes, there are $\binom{5}{2}$ ways to select 2, and for those 2 selected routes, there are 2! ways to assign the remaining 2 players. In half of these configurations (360 out of 720), player i shares a route with another player, taking 2 minutes. This is because there are $\binom{4}{2} = 6$ ways to select 2 players from 4 to take the same route, and in 3 out of those 6 cases, player i is one of the selected players. In the other half, player i takes a unique route and reaches the destination in 1 minute.
- (2-2): In this case, there are $\binom{6}{2} \times \binom{4}{2} = 15 \times 6 = 90$ ways, as there are $\binom{6}{2} = 15$ ways to choose 2 routes of out 6, and for each of these configurations, there are $\binom{4}{2} = 6$ ways to assign 2 players to each route. In all 90 configurations, player i takes 2 minutes to reach the destination.
- (3-1): There are $\binom{6}{2} \times 4 \times 2 = 15 \times 8 = 120$ configurations in this case; there are $\binom{6}{2} = 15$ ways to select 2 out of 6 routes, for each of these ways, there are $\binom{4}{3} = 4$ ways to assign 3 players to one route (with the remaining player on the other route) and for each of those configurations, there are 2 available routes that the 3 players can select from. In $\frac{3}{4} \times 120 = 90$ of these configurations, player i shares a route with 2 others, taking 3 minutes. In the remaining 30 configurations, player i is alone on their route and takes 1 minute. This is because, as mentioned earlier, there are 4 possible ways to select 3 player from 4, and in only 1 of those 4 cases, player i is not selected to be part of the group of 3.
- (4): There are 6 ways for this setup, since there are 6 possible routes. In these 6 configurations, player i takes 4 minutes to reach the destination.

We have analyzed all possible cases. Next, we will calculate the probability associated with each possible travel time for player i:

- **1 minute**: There are 360 configurations from [1-1-1-1], 360 from [2-1-1], and 30 from [3-1], giving a total of $360 + 360 + 30 = 750$ configurations. Thus, the probability of a 1-minute travel time is $\frac{750}{1296}$.

- **2 minutes**: There are 360 configurations from [2-1-1] and 90 from [2-2], for a total of $360 + 90 = 450$ configurations. So, the probability of a 2-minute travel time is $\frac{450}{1296}$.
- **3 minutes**: There are 90 configurations from [3-1], so the probability of a 3-minute travel time is $\frac{90}{1296}$.
- **4 minutes**: There are 6 configurations from [4], therefore, the probability of a 4-minute travel time is $\frac{6}{1296}$.

Bringing everything together, the expected travel time for player i can be calculated as:

$$\mathbb{E}[T] = \frac{750}{1296} \times 1 + \frac{450}{1296} \times 2 + \frac{90}{1295} \times 3 + \frac{6}{1296} \times 4 = \frac{1944}{1296} = \frac{3}{2}$$

To verify that this scenario describes a MNE, we need to assess whether a player who unilaterally deviates ends up with a lower expected travel time (and thus a higher payoff). Suppose player i decides to deviate and chooses route j. We will analyze the travel time for each of the previously discussed cases.

- (1-1-1-1): There are $\binom{6}{4} = 15$ ways to select 4 out of 6 routes. Out of these, in 5 of the ways, route j is not chosen. Thus, there are $5 \times 4! = 120$ configurations where route j is empty, meaning player i can deviate to route j and incur a travel time of 1 minute. In the $10 \times 4! = 240$ ways where route j is one of the 4 routes, there are $10 \times 6 = 60$ configurations where player i is already assigned to route j. This is because for each combination of 4 routes (including route j), there are $3! = 6$ ways to assign player i to route j and allocate the remaining 3 players to the other 3 routes. In this scenario, the travel time remains 1 minute. In the remaining $10 \times (4! - 3!) = 10 \times 18 = 180$ configurations, route j is occupied by another player, and player i deviating to route j, would result in a travel time of 2 minutes.
- (2-1-1): Out of the $\binom{6}{3} = 20$ ways to select 3 out of 6 routes, in half of them route j is not included. So, in $10 \times 36 = 360$ configurations, player i, by deviating to route j, is alone on the route and incurs a travel time of 1 minute. In the other 360 out of 720 configurations, route j is selected; in half of those cases, as previously explained, player i shares the route with another player. If player i shares route j, the travel time remains 2 minutes. If player i shares another route, then deviating to route j results again in a travel time of 2 minutes, since route j is already occupied by another player. So, in 180 configurations, the travel time for player i is 2 minutes. For the remaining 180 configurations, route j is included and player i does not share the route that they are assigned to with any other player. In that scenario, in 60 configurations, player i is already assigned to route j, so the travel time remains 1 minute; in another 60 configurations, player i is assigned to another route and a single player is assigned to route j, so, deviating to route j would result in a travel time of 2 minutes; in the remaining 60 configurations, player i is assigned to another route and 2 other players are assigned to route j. In that case, if player i deviates, the travel time increases to 3 minutes.

- (2-2): Out of $\binom{6}{2} = 15$ ways to select 2 routes, in $\frac{2}{3} \times 15 = 10$ of them route j is not chosen. So, in these $10 \times 6 = 60$ configurations of the (2-2) case, player i can deviate to route j and travel alone, resulting in a 1-minute travel time. In the remaining $5 \times 6 = 30$ configurations, route j is selected. Half of the time, player i is already assigned to route j with another player, thus the travel time remains 2 minutes. In the other half (so, in other 15 configurations), player i is assigned to the other route, meaning that deviating to route j, which is occupied by 2 other players, increases the time travel for player i to 3 minutes.

- (3-1): Similarly with the previous case, in $\frac{2}{3} \times 120 = 80$ configurations, route j is not part of the selected pair of routes. Thus, player i incurs a travel time of 1 minute by deviating. In the remaining 40 configurations, route j is included. Of these 40 ways, in $\frac{1}{8} \times 40 = 5$ configurations, player i is alone on route j, thus the travel time remains 1 minute. In $\frac{1}{8} \times 40 = 5$ configurations, i is alone on the other route, and deviating to route j increases the travel time to 4 minutes. In $\frac{3}{8} \times 40 = 15$ configurations, player i is assigned to route j with 2 other players, so the travel time remains 3 minutes. In the last $\frac{3}{8} \times 40 = 15$ configurations, player i is assigned to the other route with 2 other players, and deviating to route j (with only 1 player already on it) reduces the travel time to 2 minutes.

- (4): In 1 configuration, player i is already assigned to route j along with every other player, so the travel time remains 4 minutes. In the other 5 configurations, all players are assigned to a different route, (not route j). Therefore, if player i deviates to route j, their travel time decreases to 1 minute.

Now, we can calculate the probability of each possible travel time for player i, if they deviate to route j:

- **1 minute**: There are 120+60 configurations from [1-1-1-1], 360+60 from [2-1-1], 60 from [2-2], 80+5 from [3-1] and 5 from [4], giving a total of $180+420+85+5 = 750$ configurations. So, the probability of a 1-minute travel time is $\frac{750}{1296}$.

- **2 minutes**: There are 180 configurations from [1-1-1-1], $180 + 60$ from [2-1-1], 15 from [2-2] and 15 from [3-1], resulting in a total of $180 + 240 + 15 + 15 = 450$ configurations. Thus, the probability of a 2-minute travel time is $\frac{450}{1296}$.

- **3 minutes**: There are 60 configurations from [2-1-1-1], 15 from [2-2] and 15 from [3-1], for a total of $60 + 15 + 15 = 90$ configurations. Therefore, the probability of a 3-minute travel time is $\frac{90}{1296}$.

- **4 minutes**: There are 5 configurations from [2-2] and 1 from [4], resulting in a total of $5 + 1 = 6$ configurations. So, the probability of a 4-minute travel time is $\frac{6}{1296}$.

Thus, the expected travel time for player i, after deviating, is still:

$$\mathbb{E}[T] = \frac{750}{1296} \times 1 + \frac{450}{1296} \times 2 + \frac{90}{1295} \times 3 + \frac{6}{1296} \times 4 = \frac{1944}{1296} = \frac{3}{2}$$

Since the expected travel time remains unchanged, unilateral deviation does not reduce the expected travel time for player i, and as a consequence, it does not increase their payoff. In addition, since both player i and route j were chosen arbitrarily, a unilateral deviation by any player to any route would yield the same outcome. This confirms that the strategy in which each player selects each route with a probability of $\frac{1}{6}$ is indeed a MNE.

It is clear that this MNE is not a PNE, as the players are randomizing over the set of possible actions rather than choosing a specific strategy deterministically.

CE: The uniform distribution over all outcomes where 1 route has 2 players and 2 other routes have 1 player each constitutes a (non-product) correlated equilibrium. There are 720 possible configurations for this setup, as explained in the MNE analysis—case (2-1-1). Hence, in this CE, each of these configurations is selected with a probability of $\frac{1}{720}$. Now, consider the expected profit for a player i given that they are suggested to take route j. There are $\binom{4}{2} = 20$ ways to select 3 routes out of 6, with 10 of these selections including route j, since the probability that route j (where $j \in \{1, 2, \ldots, 6\}$) is among the 3 chosen routes is $\frac{1}{2}$. For each combination of 3 routes, there are $\binom{4}{2} \times 3 \times 2! = 36$ ways to assign the players, as there are $\binom{4}{2}$ ways to choose 2 players out of 4, and for each choice, there are 3 possible routes to assign to those 2 players.The remaining 2 routes are then assigned to the remaining 2 players in 2! ways. Therefore, when player i is advised to take route j, they understand that the other players have been given suggestions based on one of the $10 \times 36 = 360$ possible combinations that include route j.

In the scenario with 4 players and 3 routes, the likelihood that a player i is assigned to an empty route is the same as the likelihood of being assigned to a route with another player; both of which have a probability of $\frac{1}{2}$. This is analogous to the probability for a given item to be chosen if we have to choose 2 items out of 4 (let's assume that the chosen ones will be assigned to the same route). Given that the 360 possible combinations are equally likely and that, for each of these 360 triplets, player i is assigned to an empty route half of the time and to an occupied route the other half, the expected time for player i to reach their destination by following the recommended route j is:

$$\mathbb{E}[T] = \frac{1}{2} \times 1 + \frac{1}{2} \times 2 = \frac{3}{2}$$

As a result, the expected payoff is $10 - \frac{3}{2} = \frac{17}{2}$. To verify that this scenario describes a CE, we need to examine whether a player who deviates from the suggested route ends up with a lower expected time (and thus higher payoff) compared to following the suggestion. Assume player i is suggested to take route j. This means that the trusted third party has selected one of the 360 possible combinations where route j is included. Now, suppose player i ignores the suggestion and chooses another route j'; we need to calculate the expected time and consequently, the expected payoff in this case. There are 10 triplets that include route j, each with 36 possible ways of assigning the 4 players to the corresponding 3 routes. The probability that route j' is also included in the selected triplet is $\frac{2}{5}$, since there are 5 remaining routes and we need to choose 2 more routes for the triplet. Thus, with a probability of $1 - \frac{2}{5} = \frac{3}{5}$, player i will choose an empty route, resulting in a travel time of 1 minute. With a probability of $\frac{2}{5}$, route j' is included in the triplet chosen by the trusted third party, meaning that route j' is already occupied by 1 or 2 other players. To determine the probability of route j' being occupied by 2 other players, consider the following: suppose the triplet consists of routes j, j' and l. The remaining 3 players are assigned to these routes. There are $3! = 6$ possible ways where each of those 3 players is assigned to a different route. There are 3 different ways where 2 of the 3 remaining 3 players take route j' and the last player takes route l. Similarly, there are 3 different ways where 2 of the remaining 3 players take route l and the last player takes route j'. Thus, there are $6 + 3 + 3 = 12$ possible ways to assign the 3 players, with only 3 of these ways resulting in route j' being occupied by 2 other players. Therefore, the probability that route j' is occupied by 2 other players is $\frac{3}{12} = \frac{1}{4}$. Putting them all together:

- If route j' is empty, which occurs with probability $\frac{3}{5}$, the time is 1 minute.
- If route j' is occupied by 1 or 2 other players, which occurs with probability $\frac{2}{5}$ then:
  - With probability $\frac{1}{4}$, route j' has already 2 players, resulting in a time of 3 minutes.
  - With probability $\frac{3}{4}$, route j' has already 1 player, resulting in a time of 2 minutes.

Therefore, the expected time for player i to reach the destination, if they deviate to route j', is:

$$\mathbb{E}[T] = \frac{3}{5} \times 1 + \frac{2}{5} \times \left(\frac{1}{4} \times 3 + \frac{3}{4} \times 2\right) = \frac{3}{2}$$

Since the expected time when deviating is the same as the time when following the suggesting route, the player's payoff does not improve by deviating. Since player i, route j and route j' were chosen arbitrarily, a unilateral deviation by any player to any route would yield the same outcome. Hence, no player can increase their profit by unilaterally deviating from the suggested route. Therefore, this scenario indeed constitutes a correlated equilibrium (CE). It is clear that this is not a MNE, as there is a correlation between players' actions rather than each player choosing their strategy independently.

CCE: The uniform distribution over the subset of these outcomes in which the set of chosen routes is either $\{0, 2, 4\}$ or $\{1, 3, 5\}$ constitutes a coarse correlated equilibrium. As previously stated, there are 36 distinct ways to assign 4 players to a triplet of routes, resulting in a total of $36 + 36 = 72$ possible outcomes under this distribution. Each of the two triplets is chosen with equal probability of $\frac{1}{2}$. Given that 3 routes are selected, it means that one route is assigned to 2 players and the other 2 routes are assigned to 1 player each. As previously explained, for each triplet, the probability that a player i will be assigned to a non-empty route is $\frac{1}{2}$. Therefore, the expected time for a player is:

$$\mathbb{E}[T] = \frac{1}{2} \times \left(\frac{1}{2} \times 1 + \frac{1}{2} \times 2\right) + \frac{1}{2} \times \left(\frac{1}{2} \times 1 + \frac{1}{2} \times 2\right) = \frac{3}{2}$$

Thus, each player's expected profit is $10 - \frac{3}{2} = \frac{17}{2}$. To verify if this scenario describes a CCE, we need to ensure that a player cannot increase their profit by unilaterally deviating.

Assume player i decides to deviate chooses route j. Given the distribution, we know that each route is included in exactly one of the two possible triplets. Thus, there is a $\frac{1}{2}$ probability that the chosen triplet does not include route j, which means route j will be empty. In this case, the travel time for player i would be 1 minute. However, with a probability of $\frac{1}{2}$, route j is included in the chosen triplet. Out of the 36 possible combinations for assigning the 4 players to the triplet of routes, the following scenarios occur:

- In 6 combinations, player i is alone on route j, resulting in a 1-minute travel time.
- In 6 combinations, player i shares route j with one other player, resulting in a travel time of 2 minutes.
- In 6 combinations, route j is occupied by 2 other players and thus, the travel time for player i increases to 3 minutes.
- In 18 combinations, route j has one other player, resulting in a travel time of 2 minutes.

Thus, the overall expected time when deviating is:

$$\mathbb{E}[T] = \frac{1}{2} \times 1 + \frac{1}{2} \times \left(\frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \frac{3}{6} \times 2 \right) = \frac{3}{2}$$

Hence, the expected payoff for player i after deviating remains $10 - \frac{3}{2} = \frac{17}{20}$. Since the expected payoff does not increase with deviation, and both player i and route j were chosen arbitrarily, a unilateral deviation by any player to any route would yield the same outcome. This confirms that the given distribution is indeed a CCE.

Another, more conceptual way to assess whether this scenario represents a CCE is to consider the probabilities associated with being assigned to different routes. The probability of being assigned to a route with 2 other players is the same for any route, and similarly, the probability of being alone on a route is consistent across routes. Since the chances of being in a route with 2 other players or being alone are equal for all routes, players have no incentive to deviate and choose one specific route. Deviating would not improve their expected outcome, because the distribution ensures that the expected payoff is the same regardless of the route chosen.

It is important to note that this distribution does not qualify as a CE. The reason is that if player i is suggested to follow route j, they can deduce the triplet of routes that has been chosen. By selecting a route j' that is not part of that triplet, player i can ensure their route remains unoccupied, reducing their expected travel time to the minimum of 1 minute, and consequently, increasing their payoff to 9. In contrast, if player i follows the suggested route, there is a $\frac{1}{2}$ probability that they will share the route with another player, resulting in an expected travel time greater than 1 minute and a payoff of less than 9. Therefore, by deviating, player i can improve their payoff, whihc implies that the case does not satisfy the conditions for a CE.

In this example, we observe that the mixed Nash equilibrium (MNE) is not a pure Nash equilibrium (PNE), the correlated equilibrium (CE) is not a mixed Nash equilibrium (MNE) and the coarse correlated equilibrium (CCE) is not a correlated equilibrium (CE). This demonstrates that the inclusions among these equilibrium concepts are strict, with each equilibrium type being a broader or more general concept than the previous one.

Chapter 3

Auctions: A Bidding Game

3.1 Auctions as Games — The Setting

So far, we have developed a basic understanding of games from a mathematical perspective, along with certain equilibrium concepts. It is now time to introduce a specific type of game—the auction game [29]—which is a primary focus of this thesis. When we hear the word "auction", most of us—or at least I—imagine a large room filled with wealthy individuals raising their paddles to bid higher, driving up the price of the item being auctioned. This type of auction is known as the **English auction**. However this is not the only type; there are numerous kinds and categories of auctions, enough to fill more than a book. Some of these will be discussed in this thesis. But before delving into specific types, let's first establish some broader definitions and a common setting.

An **auction** can be defined as a market mechanism in which goods or services are sold to the highest bidder. A **bidder** is a potential buyer who places **bids**: offers of a price they are willing to pay to acquire the good or service being auctioned. The **seller** or **auctioneer** is the owner of the good or service—referred to hereafter as the item—and aims to sell it at the highest possible price. Each bidder is interested in acquiring the item, but seeks to do so at the lowest possible price. Each bidder i has a **private valuation** v_i, representing the maximum amount they are willing to pay for the item. If the price of the item exceeds a bidder's valuation, that bidder is no longer interested in purchasing it. Each bidder's valuation is private, meaning it is unknown to the other bidders and the seller. To determine the profitability of the auction outcome for bidder i—given their bid and valuation—we use the concept of utility. In this thesis, we apply the quasilinear utility model, where a bidder i's **utility** is defined as follows: if they lose the auction, their utility is 0; if they win the auction at a price p, their utility is $v_i - p$. More formally:

$$u_i = \begin{cases} v_i - p, & \text{if bidder } i \text{ is the winner;} \\ 0, & \text{otherwise} \end{cases}$$

As we can see, an auction is essentially a utility-maximization game. From the seller's perspective, the goal is to sell the item at the highest possible price. From the bidders' perspective, the goal is to obtain the item at the lowest possible price—provided this price is below their private valuation—thereby maximizing their utility.

As mentioned before, apart from the well-known English auction format—where bidders openly raise their bids until no one is willing to go higher, and the highest bidder wins the item—there are other formats that, while less familiar to the general public, are quite significant and widely used. One such format is the **sealed-bid auction**, which is the main focus of this thesis. More precisely, we focus exclusively on *sealed-bid single-item auctions*. **Single-item auctions** are those in which only one, undivided item is available for sale. The term "sealed-bid" refers to the manner in which the bidding process is conducted. In this format, all bidders privately submit their bids to the auctioneer—without knowing the bids of others. The bids are then opened simultaneously and the result of the auction is revealed. In every auction, the auctioneer determines the **allocation rule**, which dictates who wins the item (if anyone), and the **payment rule**, which specifies the price at which the item is sold. Bidders are informed in advance about the seller's strategy (i.e., the allocation and payment rules) and adjust their bids accordingly.

Even though formalization can make the text somewhat rigid and impersonal, it is essential for establishing a common understanding and ensuring clear communication. Therefore, more formal definitions of some previously mentioned terms are provided below.

Consider there are n bidders, each with a corresponding bid: $\mathbf{b} = (b_1, \ldots, b_n)$ and a *feasible set* X, where each element of X is an n-vector $(x_1, x_2, \ldots, x_n)$, with x_i representing the amount of items given to bidder i. In the single-item auction we are analyzing, X is the set of $0-1$ vectors that have at most one 1, i.e., $\sum_{i=1}^{n} x_i \leq 1$. When the auctioneer selects a feasible allocation, they actually select a $\mathbf{x}(\mathbf{b}) \in X \subseteq \mathbb{R}^n$ as a function of the bids **(allocation rule)**. Regarding the payments, we have $\mathbf{p}(\mathbf{b}) \in \mathbb{R}^n$, as a function of the bids **(payment rule)**. Thus, in an auction with allocation and payment rules $\mathbf{x}$ and $\mathbf{p}$, respectively, bidder i, given the bid profile (i.e., bid vector) $\mathbf{b}$, has utility:

$$u_i(\mathbf{b}) = v_i \cdot x_i(\mathbf{b}) - p_i(\mathbf{b})$$

So far, we have outlined the basic components of an auction. However, one crucial question remains: how much does that winner pay to acquire the item? This question is of outmost importance to all participants in the auction, both bidders and the seller. Depending on the price the winner pays, there are different types of auctions; the most common are the first-price auction and the second-price auction.

Even though the natural assumption might be that the winner of an auction should pay their own bid, auctions where this happens—called **first-price auctions**—are actually quite unpredictable and thus complex to analyze. In game theory terms, in a first-price auction, there is no straightforward strategy that leads all players to an equilibrium. The unpredictability of first-price auctions makes them challenging, yet also intriguing to study, which is why they are a focus in this thesis. However, before diving into the complexities of these auctions, it seems wiser to first explain an easier-to-digest type of auction to set the foundation for our discussion. This auction, which is widely used, is known as the **second-price auction**.

3.2 Second-Price Auction

In a second-price auction, the highest bidder wins the item but pays the amount of the second-highest bid (plus a small increment). At first glance, this might seem counter-intuitive, doesn't it? However, as we will demonstrate shortly, this type of auction has certain qualities that make it easier to manage for both the seller and the bidders.

Proposition 3.1. *In a second-price auction, bidding truthfully—i.e., setting one's bid equal to their private valuation—is a dominant strategy for all bidders.*

Proof. Let v_i denote the private valuation of bidder i and let $B = max_{j \neq i} b_j$ represent the highest bid of the competition. We will demonstrate that bidding v_i is the optimal strategy by comparing the utility of bidding b (where $b \neq v_i$) with the utility of bidding the private valuation v_i.
Consider two possible cases:

- Bidding more than the private valuation ($b > v_i$):
  The only situation where the bidder i's utility changes (compared to bidding v_i) is when $v_i < B < b$. In this case, by bidding b, bidder i wins the auction, but their utility becomes negative, since $v_i - B < 0$.

- Bidding less than the private valuation ($b < v_i$):
  The only situation where the bidder i's utility changes (compared to bidding v_i) is when $v_i > B > b$. By bidding b, bidder i loses the auction and misses out on a positive utility, that would have been obtained by bidding their private valuation.

Hence, bidder i has no incentive to deviate from truthful bidding. □

Another, straightforward proposition is as follows:

Proposition 3.2. *In a second-price auction, every truth-telling bidder is guaranteed non-negative utility.*

Proof. If a bidder i does not win the auction, their utility is 0. If bidder i does win the auction, their utility is given by $v_i - p$, where p is the second-highest bid. Since all bidders bid their true valuations and bidder i is the highest bidder—since this bidder is the winner—it follows that $p \leq v_i$ and hence $v_i - p \geq 0$. Therefore, in every case, the utility of each bidder remains non-negative. □

With those two propositions, we can now form the following theorem:

Theorem 3.1. *Every second-price auction is* ***DSIC (Dominant Strategy Incentive Compatible)****, meaning that the two previous propositions hold.*

A DSIC auction is considered highly desirable for two main reasons. From the bidders' perspective, it simplifies decision-making: they have a straightforward strategy—bidding their private valuation—that leads to optimal results (it maximizes their utility). This simplicity ensures that bidders can achieve the best possible outcome without needing to anticipate the strategies of others. From the seller's perspective, the DSIC property enhances predictability. Since bidders reveal their private valuations, the seller can more easily forecast the auction's outcome. This predictability helps advertisers and market analysts to better understand market dynamics and make more informed decisions.

3.3 The Vickrey Auction

Although this thesis mainly addresses first-price auctions, Vickrey auctions have played a significant role in the history of auction theory. First introduced academically by William Vickrey, a professor at Columbia University, in 1961, they are considered a foundation milestone, paving the way for further research in this field [51], [52]. Therefore, dedicating a brief section to them is the least we could do. A **Vickrey auction** refers to any sealed-bid, second-price auction. In other words, Vickrey auctions are a specific subset within the broader category of second-price auctions. This type of auction is particularly noteworthy because it combines three different and desirable properties:

- It is **DSIC**, encouraging truthful bidding, which helps to maintain control over the outcome.
- Assuming that bidders report truthfully, this auction maximizes **social surplus** by ensuring that the item is awarded to the bidder who values it the most. The social surplus can be expressed as:

  $$\sum_{i=1}^{n} v_i x_i,$$

  where x_i is 1 if bidder i wins and 0 if bidder i loses. This is, of course, subject to the feasibility constraint $\sum_{i=1}^{n} x_i \leq 1$,, which ensures that at most one item is awarded—meaning only one bidder can win in the case of single-item auctions.
- The auction can be implemented in **polynomial time**, specifically linear time.

A remark on surplus maximization that should not go unnoticed: although bidders' valuations are private, a priori unknown to the seller, this auction mechanism effectively identifies the bidder with the highest valuation, ensuring that the item goes to the one who desires it the most.

Given these three undeniably desirable properties, the Vickrey auction is a remarkable and elegant auction.

3.4 Allocation Rule: Implementable and Monotone

Having defined what constitutes a DSIC auction, we can now introduce two important definitions. In this thesis, these definitions are limited to the context of single-item auctions; however, they can be extended to broader settings with slight modifications.

Definition 3.1. *An allocation rule $\boldsymbol{x}$ for a single-item auction is* ***implementable*** *if there is a payment rule $\boldsymbol{p}$ such that the sealed-bid auction $(\boldsymbol{x}, \boldsymbol{p})$ is DSIC.*

Definition 3.2. *An allocation rule $\boldsymbol{x}$ for a single-item auction is* ***monotone*** *if, for every bidder i and bids $\boldsymbol{b}_{-i}$ from the other bidders, the allocation $x_i(z, \boldsymbol{b}_{-i})$ to bidder i is non decreasing in their bid z.*

After presenting the preceding two definitions, we can now articulate Myerson's Lemma in the context of single-item auctions.

Lemma 3.1. *For every single-item auction:*

- *(a) An allocation rule $\boldsymbol{x}$ is implementable if and only if it is monotone.*
- *(b) If $\boldsymbol{x}$ is monotone, then there is a unique payment rule such that the sealed-bid mechanism $(\boldsymbol{x}, \boldsymbol{p})$ is DSIC (assuming that $b_i = 0$ implies $p_i(b) = 0$).*
- *(c) The payment rule in (b) is given by the formula: $p_i(b_i, \boldsymbol{b}_{-i}) = \int_0^{b_i} z \cdot \frac{d}{dz} x_i(z, \boldsymbol{b}_{-i}) dz$*

This lemma is crucial and the payment formula will be addressed again in the upcoming sections.

3.5 Bayesian Setting

Up to this point, we have discussed auctions where each bidder i has a private and deterministic valuation v_i. We now broaden this setting by introducing non-deterministic valuations for the bidders. In this context, each bidder i does not have a fixed valuation, but instead draws their valuation from a **value distribution**. A common assumption, that we also adopt in this thesis, is that these value distributions are known to all bidders, as well as the seller. We represent the value distribution of each bidder i by F_i, which refers to the cumulative distribution function (CDF) of the bidder's valuation. The corresponding probability density function (PDF), if it exists, is denoted by f_i. For simplicity, we often use F_i to refer to both the distribution and its CDF. An important assumption throughout our discussion is that the values x_i are drawn independently for each bidder i, meaning that these values are statistically independent random variables. The joint distribution of these values, represented as $F = F_1 \times \cdots \times F_n$, is thus a product distribution of the individual distributions.

While auctions under the deterministic setting are treated as games described in the first chapter—with the equilibrium concepts applied accordingly—auctions in the Bayesian setting introduce an additional element of uncertainty, as the players' valuations are no longer deterministic. Consequently, these auctions are associated with a different equilibrium concept which incorporates this uncertainty, known as the Bayes-Nash Equilibrium. In Game Theory terminology, this unknown information is referred to as a player's type. The type profile of the n players is denoted by $\mathbf{t} = (t_1, ..., t_n)$. A strategy in this case is defined as a function mapping a player's type to one of the player's possible actions in the game (or to a distribution over actions in the case of mixed strategies). The strategy of player i is denoted by $b_i(\cdot)$ and a strategy profile by $\mathbf{b} = (b_1, \ldots, b_n)$.

Thus, in the Bayesian auction setting, t_i corresponds to v_i. For instance, in a second-price auction, the strategy for each player i could be $b_i(v_i) = \{\text{bid truthfully: } b_i = v_i\}$.

In this setting, with the added uncertainty, the definition of a dominant strategy is slightly modified, though it retains a clear correspondence with the dominant strategy defined in the first chapter, where players did not have multiple types.

Definition 3.3. *A player i's strategy b_i is a **dominant strategy** if, for all t_i and $\boldsymbol{b}_{-i}$ (where $\boldsymbol{b}_{-i}$ refers to the actions of all players except i), player i's utility is maximized by following strategy $b_i(t_i)$.*

3.6 Bayes-Nash Equilibrium

In our setting, as mentioned before, we assume that the distribution from which each player draws their type is common knowledge and that the distribution of types is independent. The equilibrium concept related to this type of game is the Bayes-Nash equilibrium, which is defined as follows:

Definition 3.4. *In a game with a known joint probability distribution over type profiles denoted as F, a **Bayes-Nash equilibrium (BNE)** is a strategy profile $\boldsymbol{b}$ such that, for all i and t_i, $b_i(t_i)$ is a best response when the other players play $\boldsymbol{b}_{-i}(\boldsymbol{t}_{-i})$ where $\boldsymbol{t}_{-i} \sim F_{-i}$.*

To better understand this concept, consider the case of a single-item first-price auction with two bidders, where each bidder's valuation is drawn uniformly from [0,1]. We will prove that a BNE of this game is given by the strategies $b_1(x) = \frac{x}{2}, b_2(x) = \frac{x}{2}$. Consider player 1 with valuation v_1 and bid b_1. Player 1's probability of winning the item, by biding b_1 is:

$$\Pr[Player1 \text{ wins with bid } b_1] = \Pr[b_2 \leq b_1] = \Pr[\frac{v_2}{2} \leq b_1] = \Pr[v_2 \leq 2b_1] = F(2b_1) = 2b_1$$

Thus, player 1's expected utility can be calculated as follows:

$$\mathbb{E}[u_1] = (v_1 - b_1) \cdot \Pr[1 \text{ wins with bid } b_1] = (v_1 - b_1) \cdot 2b_1 = 2v_1b_1 - 2b_1^2$$

The function $f(x) = 2ax - 2x^2$ achieves its maximum at $x = \frac{a}{2}$. Therefore, the expected utility for bidder 1 is maximized at $b_1 = \frac{v_1}{2}$. A similar analysis applies to player 2, yielding the same result, as both players have identical strategies. Since the expected utility for both players is maximized by bidding half of their valuation, given that the other player does the same, this strategy profile constitutes a BNE.

In this thesis, we mostly consider non-deterministic valuations for each player and we also examine bidders within a symmetric setting, defined as follows:

Definition 3.5. *A **symmetric setting** refers to an auction environment where all bidders have the same value distribution.*

3.7 Revenue Maximization

So far, we have explored various types of auctions that operate within both deterministic and Bayesian frameworks. Some auctions, such as the Vickrey auction, offer significant advantages, as their ability to accurately identify the bidder with the highest private valuation and thereby maximize the social surplus. However, other types of auctions may not have these desirable features. Regardless of an auction's properties, an important question remains to be addressed: what about the seller's perspective? How can auction design be optimized to ensure that the seller maximizes their profit? This consideration falls under the category of **revenue maximization**, which focuses on how the seller can achieve the highest possible profit. A significant portion of auction theory literature focuses on this objective, which is not surprising, given that auctioneers have control over the allocation and payment rules, as previously discussed. Consequently, they can design auction mechanisms with the goal of maximizing their revenue. To determine the optimal revenue-maximizing auction, we assume that the seller has prior knowledge of the value distribution F_i for each bidder's valuation. These value distributions quantify the information that the seller has on each bidder. One notable strategy for increasing or optimizing the seller's revenue in traditional auctions is the use of a **reserve price**.

Definition 3.6. *The **reserve price** is the minimum amount a bidder must offer to purchase an item.*

The reserve price is considered **anonymous** when it is set at the same level for all bidders, and **personalized** if it varies among bidders. In this thesis, we focus exclusively on cases where the reserve price is anonymous. In the following discussion, we deal with single-item, DSIC auctions. In the case of a single bidder, the expected revenue from a posted price r (a take-it-or-leave-it offer) is $r(1 - F(r))$, where $1 - F(r)$ represents the probability of a sale, as $F(r)$ denotes the probability that the bidder's valuation is less than or equal to the price of the item. Solving for the optimal r—meaning the r which maximizes the expected revenue—known as the **monopoly price**, is usually a simple process given a distribution F. For example, if F is uniform on $[0, 1]$, the monopoly price is $\frac{1}{2}$, yielding an expected revenue of $\frac{1}{4}$. However, with two bidders, the situation becomes more complex as the range of DSIC (Dominant Strategy Incentive Compatible) auctions expands. Consider an auction with two bidders whose valuations are independently drawn from a uniform distribution on [0, 1]. The Vickrey auction in this case generates revenue equal to the expected value of the lower bid, $\frac{1}{3}$. Adding a reserve price, similar to an eBay auction's opening bid, modifies this. In a Vickrey auction with a reserve price r, the item is awarded to the highest bidder unless all bids fall below r. The winner's payment is the higher of the second-highest bid or r. Introducing a reserve price can both increase and decrease revenue—higher prices lead to more revenue when only one bid exceeds r, but no sale occurs if both bids fall short. In this setting, a reserve price of $\frac{1}{2}$ increases expected revenue from $\frac{1}{3}$ to $\frac{5}{12}$.

This raises the question: can an even higher revenue be achieved with a different reserve price or auction format? Despite the extensive variety of DSIC auctions, Myerson provided a comprehensive solution to this problem.

3.7.1 Expected Revenue Equals Expected Virtual Welfare

Our objective is to characterize the optimal (regarding revenue) DSIC auction for any single-item environment and any distributions $F_1, \ldots, F_n$. Having assumed truthful bidding ($b = v$), the expected revenue of such an auction $(\mathbf{x}, \mathbf{p})$ is:

$$\mathbb{E}_{\mathbf{v}}\left[\sum_{i=1}^{n} p_i(\mathbf{v})\right],$$

where the expectation is taken over the distribution $F_1 \times \cdots \times F_n$ of bidders' valuations. However, maximizing this expression directly over all DSIC mechanisms seems complex. Instead, we derive an alternative revenue formula that depends only on the allocation rule $\mathbf{x}$, which simplifies the maximization process. Myerson's formula for the expected payment of bidder i, as mentioned in 3.1 is:

$$p_i(b_i, \mathbf{b}_{-i}) = \int_0^{b_i} z \cdot x_i'(z, \mathbf{b}_{-i}) dz,$$

which holds for any monotone allocation function $x_i(z, \mathbf{b}_{-i})$. This shows that the payments are fully determined by the allocation rule, allowing to express expected revenue in terms of its allocation rule alone. By fixing i and $\mathbf{v}_{-i}$, the expected payment for bidder i can be written as:

$$\mathbb{E}_{v_i \sim F_i}\left[p_i(\mathbf{v})\right] = \int_0^{v_{\max}} \int_0^{v_i} z \cdot x_i'(z, \mathbf{v}_{-i}) dz\, f_i(v_i) dv_i$$

Reversing the order of integration simplifies this to:

$$\begin{aligned}\mathbb{E}_{v_i \sim F_i}\left[p_i(\mathbf{v})\right] &= \int_0^{v_{\max}} \left(\int_z^{v_{\max}} f_i(v_i) dv_i\right) \cdot z \cdot x_i'(z, \mathbf{v}_{-i}) dz = \\ &= \int_0^{v_{\max}} (1 - F_i(z)) \cdot z \cdot x_i'(z, \mathbf{v}_{-i}) dz\end{aligned}$$

Applying integration by parts to further simplify the expression:

$$\begin{aligned}\mathbb{E}_{v_i \sim F_i}\left[p_i(\mathbf{v})\right] &= (1 - F_i(z)) \cdot z \cdot x_i(z, \mathbf{v}_{-i})\Big|_0^{v_{\max}} - \int_0^{v_{\max}} x_i(z, \mathbf{v}_{-i}) \cdot (1 - F_i(z) - z f_i(z)) dz = \\ &= \int_0^{v_{\max}} \left(z - \frac{1 - F_i(z)}{f_i(z)}\right) \cdot x_i(z, \mathbf{v}_{-i}) f_i(z) dz\end{aligned}$$

We define the **virtual valuation** $\varphi_i(v_i)$ for bidder i with valuation v_i drawn from F_i as:

$$\varphi_i(v_i) = v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}$$

This transforms the expected payment of bidder i into:

$$\mathbb{E}_{v_i \sim F_i}\left[p_i(\mathbf{v})\right] = \mathbb{E}_{v_i \sim F_i}\left[\varphi_i(v_i) \cdot x_i(\mathbf{v})\right]$$

Taking the expectation over all bidders' valuations, we obtain:

$$\mathbb{E}_{\mathbf{v}}\left[p_i(\mathbf{v})\right] = \mathbb{E}_v\left[\varphi_i(v_i) \cdot x_i(\mathbf{v})\right]$$

Using linearity of expectation, we conclude that:

$$\mathbb{E}_{\mathbf{v}}\left[\sum_{i=1}^{n} p_i(\mathbf{v})\right] = \sum_{i=1}^{n} \mathbb{E}_{\mathbf{v}}\left[p_i(\mathbf{v})\right] =$$

$$= \sum_{i=1}^{n} \mathbb{E}_{\mathbf{v}}\left[\varphi_i(v_i) \cdot x_i(\mathbf{v})\right] = \mathbb{E}_{\mathbf{v}}\left[\sum_{i=1}^{n} \varphi_i(v_i) \cdot x_i(\mathbf{v})\right]$$

Thus, the expected revenue of an auction is equal to its expected **virtual welfare**—the virtual welfare of an auction on the valuation profile $\mathbf{v}$ is defined as $\sum_{i=1}^{n} \varphi_i(v_i) \cdot x_i(\mathbf{v})$. Therefore, maximizing expected revenue in DSIC auctions is equivalent to maximizing expected virtual welfare. Before presenting an important result, we must introduce one more definition.

Definition 3.7. *A distribution F is called **regular** if the corresponding virtual valuation function $\varphi(v) = v - \frac{1-F(v)}{f(v)}$ is strictly increasing.*

Taking all these points into account, we can now state the following result: in a single-item auction with i.i.d. bidders, where the valuation distribution is regular, the maximum expected revenue is equivalent to the revenue obtained from a Vickrey auction with a reserve price of $\varphi^{-1}(0)$. In fact, this simple format—a Vickrey auction with a reserve price—is sufficient to provide the seller with the maximum possible revenue. There is no need for more complex mechanisms or designs, as no better results, regarding revenue, can be achieved.

The question of which auction format generates higher revenue has been a major focus of research. One significant result related to revenue equivalence is the following: in a Bayes-Nash equilibrium, when bidders' valuations are independent and identically distributed random variables (i.i.d.), the revenue generated from a first-price auction with the optimal reserve price equals that of a second-price auction with the optimal reserve price [36].

With these key points on revenue established, we now turn to the final significant aspect of auctions for this thesis: repeated auctions.

3.7.2 Repeated Auctions

Up to this point, the auction setting has been treated as a one-time game. However, there is increasing interest in repeated auctions, where the same item is auctioned multiple times. At first glance, the concept of repeated auctions may seem counterintuitive, especially when thinking of items like artwork, which are typically sold only once, or at most, infrequently. Nevertheless, repeated auctions are prevalent in various contexts. A prominent example is the auctioning of advertising slots on platforms like Google [39]. Every day, companies (bidders) compete for prime advertising positions, which are crucial to their revenue generation. In the digital era, with the rapid growth of technology, online advertising—such as the aforementioned Google ads—has become a significant factor in a company's visibility and profitability, making it strategically important. Thus, repeated auctions are quite common.

What makes repeated auctions particularly interesting is the fact that bidders can learn from past experiences and adapt their bidding strategies over time. By observing how other bidders behave, they can adjust their own bids accordingly. In this dynamic environment, modeling bidder behavior requires a structured approach. To address this, we borrow terminology and concepts from the field of online learning algorithms. Understanding these algorithms and their role in repeated auctions is essential for simulating bidder behavior in a meaningful way and for presenting the results of this thesis.

Chapter 4

Online Learning Algorithms

4.1 Notation and Setting

As outlined in the previous chapter, auctions are a form of strategic game, where bidders aim to maximize their utility and the seller seeks to maximize revenue. We have already examined equilibrium concepts, focusing on one-time auctions. However, repeated auctions introduce new dynamics. Do strategic players converge to an equilibrium, and if so, how quickly does this occur? To investigate these questions, we rely on online learning algorithms to simulate how players evolve their strategies and adapt their behavior over time.

In computer science, an online algorithm is an algorithm that processes its input incrementally, handling data as it is received, without having access to the entire input upfront. In our framework, we deal with **online learning algorithms**, which adapt and learn from new data as it becomes available, offering continual updates to the predictive model [46]. Our setting can be viewed as a structured repeated game. In this framework, a decision-maker repeatedly makes choices without knowing the outcomes in advance. Each choice incurs a certain loss. These losses, unknown beforehand, may even be adversarially chosen, depending on the player's actions. However, it is crucial to assume that losses are bounded; otherwise, an adversary could gradually decrease the scale of the loss at each step, preventing the algorithm from recovering after incurring an initial large loss, and, in that case, the framework would not be really meaningful. The total number of iterations (or game rounds) is represented by T, the decision set (i.e. the possible actions a player can take) by K, and the losses (or costs) by bounded convex functions over this set. The setup is as follows.

At each iteration t, the player selects a mixed strategy p_t, i.e., a probability distribution over the actions in K. After making this choice, the adversary reveals a convex cost function $f_t \in \mathcal{F} : K \to \mathbb{R}$, chosen from $\mathcal{F}$, a bounded set of possible functions. An action $k_t \in K$ is then chosen according to the distribution p_t and the corresponding cost $f_t(k_t)$ is incurred, representing the value of the function at the chosen action k_t. After this iteration, the player learns the entire cost vector f_t, not just the realized cost $f_t(k_t)$.

To illustrate this setup, consider the following example: the decision set K represents the possible routes a driver can take to travel from point A to point B, and the cost function represents the travel time. This function is unknown to the driver in advance, as it depends on the traffic conditions on a given day. Each day, the driver must choose a route. After making their choice, the driver learns the travel time for that route, as well as what the travel times would have been for the other routes. Based on this information, the driver may adjust their strategy by assigning different probabilities to each route the next day. Although there is no adversary actively trying to increase the driver's travel time, this scenario—hopefully—gives a clear intuition for how certain real-world problems can be modeled as online decision-making problems. In the auction framework, the decision set represents a player's strategy set, and the cost function is determined by the strategies chosen by all other players.

After understanding the setting, a natural question arises: how can we determine if an algorithm performs well in online decision-making problems? There should be some performance metric to evaluate it. Initially, this setup might seem unfair—the adversary has the flexibility to select the cost function *after* the decision-maker has committed to a strategy. For that reason, it is unrealistic to expect an online decision-making algorithm to match the performance of the best possible sequence of actions in hindsight. This hindsight benchmark, denoted by $\sum_{t=1}^{T} \min_{\mathbf{k}\in K} f_t(k)$, is too stringent.

Consider a simple example to illustrate this: suppose a player is using an algorithm to select between two actions: heads (H) and tails (T)—so the decision set is $K = \{H, T\}$. The adversary can manipulate the cost function as follows: if the player selects heads with a probability of at least $\frac{1}{2}$, then $f_t(H) = 1$ and $f_t(T) = 0$. Otherwise, $f_t(H) = 0$ and $f_t(T) = 1$. This strategy guarantees that the algorithm's expected cost is at least $\frac{T}{2}$, while the optimal action sequence in hindsight would incur a total cost of 0.

As seen in this example, comparing the expected cost of an online decision-making algorithm with that of the best action sequence in hindsight does not yield meaningful results because the latter is too powerful. To address this, experts in the field introduced an alternative performance metric known as "regret". In this thesis, we focus on two types of regret: external regret and swap regret.

4.2 The Notion of External Regret

One method to evaluate the performance of an online algorithm is by comparing it to the best fixed action in hindsight. This comparison is measured using a metric known as external regret, which is formally defined below:

Definition 4.1. *The time-averaged* ***external regret*** *of an action sequence $k_1, k_2, \ldots, k_T$ relative to a fixed action k is defined as:*

$$R = \frac{1}{T}\left(\sum_{t=1}^{T} f_t(k_t) - \sum_{t=1}^{T} f_t(k)\right)$$

From this point onward, the term regret will refer to the time-averaged external regret as defined above. An algorithm is considered no-regret, and therefore effective, when its performance is almost as good as the best fixed strategy in hindsight. This can be formally expressed as follows:

Definition 4.2. *An online decision-making algorithm is a* ***no-(external) regret*** *algorithm if, for any $\epsilon > 0$ there exists a sufficiently large time horizon $T = T(\epsilon)$ such that, for any possible adversary, the expected regret (4.1), with respect to any action $k \in K$, is at most ϵ.*

Typically, when we talk about regret, we consider scenarios where minimizing cost is the goal. In the context of utility-maximization, however, regret can be defined by switching the terms: rather than comparing our cost to the cost of the best fixed action in hindsight, we instead compare our utility to the utility of the best fixed action. Instead of subtracting from our cost, we subtract our utility from that of the best.

We previously mentioned that, in our framework, players choose a probability distribution over their decision set. The following example highlights why this is important. If a player were to choose actions deterministically, the adversary, by selecting the cost function after observing the player's choice, would always be able to exploit this and prevent no-regret performance. Consider a case with $K \geq 2$ possible actions, where the player uses a deterministic algorithm. At each time step t, the algorithm commits to a specific action k_t. The adversary's optimal response—assuming costs lie in the range $[0, 1]$—is to assign a cost of 1 to the chosen action k_t and a cost of 0 to all other actions. Under this setup, the algorithm incurs a total cost of T, while the best fixed action in hindsight incurs a cost of at most T/n. Therefore, the regret of the deterministic algorithm remains constant as $T \to \infty$, with respect to some action k. This demonstrates that no algorithm can be both deterministic and no-regret.

Even for non-deterministic algorithms, there is an upper bound on their performance in terms of regret that, unfortunately, cannot be surpassed:

Theorem 4.1. *For every (randomized) algorithm, the expected regret cannot decrease faster than $\Theta(\sqrt{(\ln K)/T})$, where K denotes the number of possible actions and T the number of iterations.*

The following example, where the decision set consists of only two actions ($K = 2$), demonstrates that no randomized algorithm can achieve an expected regret that decreases faster than $\Theta(1/\sqrt{T})$. A similar reasoning shows that for K actions, the theorem (4.1) holds. Consider an adversary who, at each round T, randomly selects between two cost vectors, $(1, 0)$ and $(0, 1)$, with equal probability. This means that at each round, the adversary assigns a cost of 1 to one action and 0 to the other, or vice versa. No matter how sophisticated or simple an online decision-making algorithm is, its cumulative expected cost after T rounds will be $T/2$. However, when looking in hindsight, one of the two fixed actions will have a cumulative cost of $T/2 - \Theta(\sqrt{T})$. This is analogous to flipping a fair coin T times: the expected number of heads is $T/2$, but the standard deviation is $\Theta(\sqrt{T})$. Thus, there exists a distribution of 2^T adversarial strategies such that any algorithm will have an expected regret of at least $\Omega(1/\sqrt{T})$, where the expectation is over both the algorithm's random choices and the adversary's strategy. Consequently, for every algorithm, there is an adversary that ensures the algorithm's expected regret cannot be better than $\Omega(1/\sqrt{T})$.

Hence, deterministic no-regret algorithms do not exist and for randomized ones, there are established upper bounds on their performance. However, an important question remains: how can we be confident that a no-regret randomized algorithm actually exists? Fortunately, this has been proven.

Theorem 4.2. *There exist simple no-regret algorithms with expected regret* $O\left(\sqrt{\frac{\ln K}{T}}\right)$ *with respect to every fixed action, where* K *represents the number of possible actions and* T *the number of iterations.*

A direct consequence of this result is that a relatively small number of iterations, growing logarithmically with K, is enough to reduce the expected regret to a low constant value.

Corollary 4.1. *There exists an online decision-making algorithm that, for every* $\epsilon > 0$, *has expected regret at most* ϵ *with respect to every fixed action after* $O\left(\frac{\ln K}{\epsilon^2}\right)$ *iterations.*

The task of designing a no-regret algorithm is often referred to as "combining expert advice." This analogy arises because each action can be viewed as an "expert" offering recommendations, and a no-regret algorithm performs asymptotically as well as the best expert. While this equivalence may not be immediately intuitive, it becomes clearer by considering what no-regret entails. In essence, no-regret means that, given a set of recommendations/actions/experts, the algorithm can quickly identify which option leads to the best outcome if followed consistently over time. Thus, the question "can the algorithm perform as well as the best fixed action in hindsight?" can be reframed as "can the algorithm perform as well as the best expert in hindsight?"

There are some simple no-regret algorithms. They are based on the following key idea: past performance of action guides which action is selected at present. The probability of choosing an action should be decreasing in its cumulative cost. 'Bad' actions should be punished, meaning that their probability of being played in the next round should be decreased. Two of the most famous no-regret algorithms are the **Randomized Weighted Majority** (RWM) algorithm and the **Hedge** algorithm. They both reserve the main key idea but have slight differences. We will analyze both of them in the following subsections and prove the theorem (4.2). However, first, we will discuss the deterministic algorithm **Weighted Majority** (WM), since it serves as the foundation for the Randomized Weighted Majority algorithm.

4.2.1 Weighted Majority (WM)

The **Weighted Majority (WM)** algorithm can be described as follows: each expert i is assigned a weight $W_t(i)$ at each time step t. Initially, every expert is given equal weight, $W_1(i) = 1$ for all $i \in [N]$. For each time step $t \in [T]$, let $S_t(A)$ and $S_t(B) \subseteq [N]$ represent the groups of experts that select action A and B, respectively, at time t. The weights for the actions are then computed as follows:

$$W_t(A) = \sum_{i \in S_t(A)} W_t(i), \quad W_t(B) = \sum_{i \in S_t(B)} W_t(i)$$

The algorithm predicts the action at time t by selecting:

$$a_t = \begin{cases} A, & \text{if } W_t(A) \geq W_t(B) \\ B, & \text{otherwise} \end{cases}$$

After making the prediction, the weights are updated based on the correctness of the expert's prediction. The update rule for the weights is as follows:

$$W_{t+1}(i) = \begin{cases} W_t(i), & \text{if expert } i \text{ was correct} \\ W_t(i)(1-\varepsilon), & \text{if expert } i \text{ was wrong} \end{cases}$$

The parameter ε influences the performance of the algorithm. Intuitively, when ε is very small, the probability distribution p_t moves closer to a uniform distribution. This means that smaller values of ε promote more exploratory behavior. On the other hand, as ε increases toward 1, the distribution p_t shifts towards concentrating all of its weight on the action with the lowest accumulated cost up to that point. So, larger values of ε drive more exploitative behavior. Hence, ε serves as a tuning parameter that balances exploration and exploitation. Typically, ε is chosen to be between 0 and $\frac{1}{2}$.

Now we move on to bound the number of mistakes the algorithm makes. Let N represent the number of experts, M_T the total number of mistakes made by the algorithm up to time T, and $M_T(i)$ the number of mistakes made by expert i up to time T. We will show that for any expert $i \in [N]$, the following holds:

$$M_T \leq 2(1+\varepsilon)M_T(i) + \frac{2\ln N}{\varepsilon}$$

Let $\Gamma_t = \sum_{i=1}^{N} W_t(i)$ for all $t \in [T]$, with $\Gamma_1 = N$. Since the WM algorithm decreases weights over time, we have:

$$\Gamma_{t+1} \leq \Gamma_t$$

In the rounds that the algorithm makes a mistake, at least half of the total weight is associated with experts who were wrong. Therefore,

$$\Gamma_{t+1} \leq \frac{1}{2}\Gamma_t(1-\varepsilon) + \frac{1}{2}\Gamma_t = \Gamma_t\left(1 - \frac{\varepsilon}{2}\right)$$

So, we conclude that:

$$\Gamma_t \leq \Gamma_1\left(1-\frac{\varepsilon}{2}\right)^{M_T} = N\left(1-\frac{\varepsilon}{2}\right)^{M_T}$$

By definition, for any expert i, we have:

$$W_T(i) = (1-\varepsilon)^{M_T(i)}$$

Since $W_T(i) \leq \Gamma_T \leq N\left(1-\frac{\varepsilon}{2}\right)^{M_T}$, we can take the logarithm of both sides to get:

$$M_T(i)\ln(1-\varepsilon) \leq \ln N + M_T \ln\left(1-\frac{\varepsilon}{2}\right)$$

Recall the Taylor series expansion for $\ln(1-x)$ is given by:

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots$$

To upper bound $\ln(1-\varepsilon)$, we discard all the higher-order (negative) terms except the first one, which gives us $-\varepsilon$. Similarly, we can lower bound $\ln(1-\varepsilon)$ by retaining only the first two terms and doubling the second term for a tighter bound (assuming $\varepsilon \leq \frac{1}{2}$), resulting in $-\varepsilon - \varepsilon^2$. Therefore, for $\varepsilon \in (0, \frac{1}{2}]$, we can simplify the expression:

$$-M_T(i)(\varepsilon + \varepsilon^2) \leq \ln N - M_T \frac{\varepsilon}{2}$$

Thus, diving by ε, we conclude that:

$$M_T \leq 2M_T(i)(1 + \varepsilon) + \frac{2 \ln N}{\varepsilon}$$

This completes the proof.

4.2.2 Randomized Weighted Majority (RWM)

The randomized version of the Weighted Majority (WM) algorithm, called **Randomized Weighted Majority (RWM)**, is similar to the WM, with the difference that at each iteration t, we choose expert i with probability $p_t(i) = \frac{W_t(i)}{\sum_{j=1}^{N} W_t(j)}$.

Let M_t represent the number of errors made by the RWM algorithm up to iteration t. As before, let $\Gamma_t = \sum_{i=1}^{N} W_t(i)$ for all $t \in [T]$, noting that $\Gamma_1 = N$. Define the indicator variable $\tilde{m}_t = M_t - M_{t-1}$, which equals one if the RWM algorithm makes an error at iteration t and zero otherwise. Respectively, let $m_t(i)$ be an indicator equal to one if expert i makes an error at time t and zero otherwise.

By summing over the weights, we obtain:

$$\Gamma_{t+1} = \sum_i W_t(i)(1 - \varepsilon m_t(i)) = \Gamma_t \left(1 - \varepsilon \sum_i p_t(i) m_t(i)\right),$$

$$\text{where } p_t(i) = \frac{W_t(i)}{\sum_j W_t(j)}.$$

This simplifies to:

$$\Gamma_{t+1} = \Gamma_t(1 - \varepsilon \mathbb{E}[\tilde{m}_t]) \leq \Gamma_t e^{-\varepsilon \mathbb{E}[\tilde{m}_t]}$$

The last inequality holds, since $1 + x \leq ex$. For any expert i, we have by definition:

$$W_T(i) = (1 - \varepsilon)^{M_T(i)}$$

Since $W_T(i)$ is always less than the total weight Γ_T, we conclude:

$$(1 - \varepsilon)^{M_T(i)} = W_T(i) \leq \Gamma_T \leq N e^{-\varepsilon \mathbb{E}[M_T]}$$

Taking the logarithm of both sides gives:

$$M_T(i) \ln(1 - \varepsilon) \leq \ln N - \varepsilon \mathbb{E}[M_T]$$

Using the same approximation as before, for $\varepsilon \in (0, \frac{1}{2}]$, we obtain:

$$-M_T(i)(\varepsilon + \varepsilon^2) \leq \ln N - \varepsilon \mathbb{E}[M_T]$$

By dividing by ε, we get:

$$\mathbb{E}[M_T] \leq \frac{\ln N}{\varepsilon} + M_T(i)(1+\varepsilon)$$

Therefore, for any expert $i \in [N]$, the expected number of mistakes is bounded by:

$$\mathbb{E}[M_T] \leq (1+\varepsilon)M_T(i) + \frac{\ln N}{\varepsilon} \leq M_T(i) + \varepsilon T + \frac{\ln N}{\varepsilon}$$

We can optimize ε to minimize the above bound. The right-hand side of the bound takes the form $f(x) = ax + \frac{b}{x}$, which reaches its minimum at $x = \sqrt{\frac{b}{a}}$. Therefore, the bound is minimized at:

$$\varepsilon^\star = \sqrt{\frac{\ln N}{T}}$$

By using this optimal value of ε, we obtain for the best expert $i^\star$:

$$M_T \leq M_T(i^\star) + 2\left(\sqrt{T \ln N}\right)$$

So, the cumulative expected cost is at most $2\sqrt{T \ln n}$ more than the cumulative cost of the best expert (best fixed action in hindsight); dividing both sides by T shows that the time-averaged regret, as defined in 4.1 is at most $2\sqrt{\frac{\ln N}{T}}$. This completes the proof of theorem (4.2).

4.2.3 Hedge

The **Hedge** algorithm closely resembles the Randomized Weighted Majority (RWM) algorithm, with one key distinction: instead of counting discrete mistakes, it assesses the performance of each expert using a non-negative real value $\ell_t(i)$, which represents the loss of expert i at time step t. Like RWM, Hedge guarantees that a decision maker following its strategy will experience an average expected loss that approaches the loss of the best expert in hindsight.

Initially, all weights are set to 1. During each round t, an action (expert) i_t is selected based on the current weight distribution, where the probability of selecting expert $i_t = i$ is $x_t(i) = \frac{W_t(i)}{\sum_j W_t(j)}$. The loss for the chosen action is $\ell_t(i_t)$. In the Hedge algorithm, the weights are updated after each round according to the rule:

$$W_{t+1}(i) = W_t(i) e^{-\epsilon \ell_t(i)}$$

The expected loss of the algorithm is denoted in vector notation as:

$$\mathbb{E}[\ell_t(i_t)] = \sum_{i=1}^{N} x_t(i)\ell_t(i) = x_t^\top \ell_t,$$

where $x_t(i)$ is the probability of selecting expert i at time step t, and $\ell_t(i)$ is the loss of expert i at time t.

Let ℓ_t^2 denote the N-dimensional vector of squared losses, i.e., $\ell_t^2(i) = \ell_t(i)^2$, and let $\varepsilon > 0$. Assuming all losses are non-negative, the Hedge algorithm satisfies the following bound for any expert $i^\star \in [N]$:

$$\sum_{t=1}^{T} x_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t(i^\star) + \varepsilon \sum_{t=1}^{T} x_t^\top \ell_t^2 + \frac{\ln N}{\varepsilon}$$

As in previous analyses, let $\Gamma_t = \sum_{i=1}^{N} W_t(i)$ for all $t \in [T]$, noting that $\Gamma_1 = N$. Analyzing the sum of weights at each iteration, the total weight at time $t+1$ is given by:

$$\Gamma_{t+1} = \sum_{i=1}^{N} W_t(i) e^{-\varepsilon \ell_t(i)}$$

This can be rewritten as:

$$\Gamma_{t+1} = \Gamma_t \sum_{i=1}^{N} x_t(i) e^{-\varepsilon \ell_t(i)}$$

$$\text{where } x_t(i) = \frac{W_t(i)}{\sum_{j=1}^{N} W_t(j)}.$$

Using the approximation $e^{-x} \leq 1 - x + x^2$ for $x \geq 0$, we obtain:

$$\Gamma_{t+1} \leq \Gamma_t \sum_{i=1}^{N} x_t(i) \left(1 - \varepsilon \ell_t(i) + \varepsilon^2 \ell_t(i)^2\right)$$

This simplifies to:

$$\Gamma_{t+1} \leq \Gamma_t \left(1 - \varepsilon x_t^\top \ell_t + \varepsilon^2 x_t^\top \ell_t^2\right),$$

where $x_t^\top \ell_t$ represents the expected loss of the algorithm at time t, and $x_t^\top \ell_t^2$ is the expected squared loss. Using $1 + x \leq e^x$, we further bound the sum of weights by:

$$\Gamma_{t+1} \leq \Gamma_t e^{-\varepsilon x_t^\top \ell_t + \varepsilon^2 x_t^\top \ell_t^2}$$

For fixed expert $i^\star$, we have:

$$W_{T+1}(i^\star) = e^{-\varepsilon \sum_{t=1}^{T} \ell_t(i^\star)}$$

Since $W_T(i^\star)$ is always less than the sum of all weights Γ_T, we have:

$$W_{T+1}(i^\star) \leq \Gamma_{T+1} \leq N e^{-\varepsilon \sum_{t=1}^{T} x_t^\top \ell_t + \varepsilon^2 \sum_{t=1}^{T} x_t^\top \ell_t^2}$$

Taking the logarithm of both sides gives:

$$-\varepsilon \sum_{t=1}^{T} \ell_t(i^\star) \leq \ln N - \varepsilon \sum_{t=1}^{T} x_t^\top \ell_t + \varepsilon^2 \sum_{t=1}^{T} x_t^\top \ell_t^2$$

Thus, by rearranging, we conclude that:

$$\sum_{t=1}^{T} x_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t(i^\star) + \frac{\ln N}{\varepsilon} + \varepsilon \sum_{t=1}^{T} x_t^\top \ell_t^2$$

If the losses are bounded between $[0, 1]$, we get:

$$\sum_{t=1}^{T} x_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t(i^\star) + \frac{\ln N}{\varepsilon} + \varepsilon \sum_{t=1}^{T} x_t^\top \ell_t^2 \leq \sum_{t=1}^{T} \ell_t(i^\star) + \frac{\ln N}{\varepsilon} + \varepsilon T$$

For $\varepsilon = \sqrt{\frac{\ln N}{T}}$, we obtain:

$$\sum_{t=1}^{T} x_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t(i^\star) + 2\sqrt{T \ln N}$$

If the losses are bounded between $[0, G]$, where G is a real, positive number, we have:

$$\sum_{t=1}^{T} x_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t(i^\star) + \frac{\ln N}{\varepsilon} + \varepsilon T G^2$$

For $\varepsilon = \sqrt{\frac{\ln N}{TG^2}}$, we obtain:

$$\sum_{t=1}^{T} x_t^\top \ell_t \leq \sum_{t=1}^{T} \ell_t(i^\star) + 2G\sqrt{T \ln N}$$

Thus, in all cases, the time-averaged regret is $O\left(\sqrt{\frac{\ln N}{T}}\right)$. Therefore, the Hedge algorithm is also classified as a no-regret algorithm, as it satisfies the conditions of Theorem (4.2).

4.3 The Notion of Swap Regret

In this section, we introduce a concept that is slightly different from, and more powerful than external regret, known as **swap regret**. Similar to the notion of external regret, the time-averaged swap regret of an online decision-making algorithm (referred to as swap regret) is defined as follows:

Definition 4.3. *The time-averaged **swap regret** of an action sequence $k_1, k_2, k_3, \ldots, k_T$, with respect to every switching function $\delta : K \to K$, where K denotes the number of possible actions, is defined as:*

$$SwapR = \frac{1}{T}\left(\sum_{t=1}^{T} f_t(k_t) - \sum_{t=1}^{T} f_t(\delta(k_t))\right)$$

An online learning algorithm is said to have no-swap regret if, for every adversary, the expected swap regret is $o(1)$ as $T \to \infty$. Since fixed actions are a special case of constant switching functions, any algorithm with no-swap regret also exhibits no-external regret. Although formulated differently, this definition parallels that of no-external regret, with the key difference being that here we are dealing with a switching function δ rather than a constant action k.

4.3.1 Reduction from Swap Regret to External Regret

Instead of describing from scratch a no-swap regret algorithm, as was done for the no-regret algorithm, we will convert a no-regret algorithm into a no-swap regret one. This reduction preserves computational efficiency. For instance, plugging the Hedge algorithm into this reduction yields a polynomial time no-swap regret algorithm.

Theorem 4.3. *If there is a no-external regret algorithm, then there is also a no-swap regret one.*

Proof. Let n denote the total number of actions ($K = [n]$) and let $M_1, M_2, \ldots, M_n$ represent no-regret algorithms, such as RWM, Hedge or other similar algorithms. These algorithms may be different instances of the same algorithm. Each of these no-regret algorithms corresponds to one action. The proposed no-swap regret algorithm integrates these individual no-regret algorithms as follows.

At each time step $t = 1, 2, \ldots, T$, the algorithm:

- Receives the distributions $q_1^t, q_2^t, \ldots, q_n^t$ over actions from the algorithms $M_1, M_2, \ldots, M_n$
- Computes the "consensus" distribution p^t
- Receives the cost vector c^t from the adversary and incur the appropriate costs
- Provides M_j with the cost vector $p^t(j) \cdot c^t$, meaning it "deceives" M_j by scaling the true cost vector by $p^t(j)$

The time-averaged expected cost of this algorithm is:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} p^t(i)c^t(i) \tag{I}$$

Consider a switching function $\delta : K \to K$. If the algorithm follows the switching function δ, the time-averaged expected cost can be expressed as:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} p^t(i)c^t(\delta(i)) \tag{II}$$

To demonstrate that this algorithm achieves no-swap regret, we need to establish that the difference between (I) and (II) is asymptotically small, specifically $o(1)$, for every switching function δ.

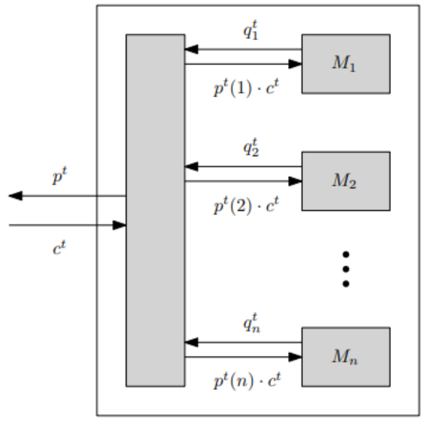


Figure 4.1: Reduction from swap regret to external regret

Now, adopt the perspective of the algorithm M_j. Since M_j is a no-regret algorithm, it exhibits no-regret with respect to $\delta(j)$. However, this regret applies to the perceived cost vectors, not the actual ones. More formally, this can be expressed as:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} q_j^t(i)(p^t(j)c^t(i)) - \frac{1}{T}\sum_{t=1}^{T} p^t(j)c^t(\delta(j)) \leq R_j,$$

where the first term represents M_j's perceived cost, the second term reflects the perceived cost in hindsight of always playing action $\delta(j)$ and R_j denotes the regret which approaches zero as $T \to \infty$.

Summing this inequality over all $j \in [n]$, we get:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n} q_j^t(i)p^t(j)c^t(i) - \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{n} p^t(j)c^t(\delta(j)) \leq \sum_{j=1}^{n} R_j$$

The term $\sum_{j=1}^{n} R_j$ goes to 0 as $T \to \infty$ (we think of n as fixed as $T \to \infty$). Notice that the second term on the left-hand side corresponds to expression (II). Therefore, the remaining objective is to show that the first term matches expression (I). In other words, we need to establish that:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n} q_j^t(i)p^t(j)c^t(i) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n} p^t(i)c^t(i),$$

which holds by defining p^t, for all t and i, as: $p^t(i) = \sum_{j=1}^{n} q_j^t(i)p^t(j).$

Although it might seem odd to define p^t in terms of itself, this can be achieved by defining a Markov chain with states $K = [n]$ and transition probabilities $q_j^t(i)$ from state j to state i. Since q_j^t is a valid probability distribution (i.e., $\sum_{i=1}^{n} q_j^t(i) = 1$), a stationary distribution for this Markov chain always exists. This stationary distribution will satisfy the last equation, as it is essentially the definition of such a distribution. This completes the proof of the theorem (4.3). □

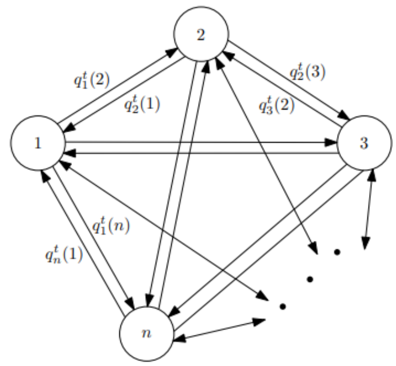


Figure 4.2: Markov Chain

4.4 Linking Regret Minimization with Equilibrium Concepts

So far, we have focused on no-external regret and no-swap regret algorithms from the perspective of a single player. Now, we extend this analysis to multi-player scenarios within the framework of utility-maximization games. During each time step $t = 1, 2, \ldots, T$, in no-regret (or no-swap regret) dynamics, each player i independently selects a mixed strategy p_i^t using a no-regret (or no-swap regret) algorithm, concurrently with the other players. Afterwards, each player i receives a utility vector u_i^t, where $u_i^t(s_i)$ represents the expected utility of strategy s_i, assuming the other players are following their respective mixed strategies. Specifically:

$$u_i^t(s_i) = \mathbb{E}_{\mathbf{s}_{-i} \sim \sigma_{-i}}[U_i(s_i, \mathbf{s}_{-i})],$$
$$\text{where } \sigma_{-i} = \prod_{j \neq i} \sigma_j.$$

It is important to highlight that each player can utilize any no-regret (or no-swap regret) algorithm, provided that all players use the same type of algorithm—either all employ no-regret algorithms or all use no-swap regret algorithms.

No-regret dynamics can be implemented efficiently. For instance, if every player applies the Randomized Weighted Majority (RWM) or the Hedge algorithm, then in each round, each player performs a straightforward update of one weight per strategy. After $O\left(\frac{\ln K}{\epsilon^2}\right)$ iterations, every player achieves an expected regret of at most ϵ for all strategies, where N is the maximum size of a player's strategy set.

The time-averaged play history of joint play under no-regret dynamics converges to a coarse correlated equilibrium. This reveals a key connection between static equilibrium concepts and outcomes generated by natural learning processes.

Proposition 4.1. *Suppose that after T iterations of no-regret dynamics, each player in a utility-maximization game experiences a regret of at most ϵ for each of their strategies. Let $\sigma^t = \prod_{i=1}^{k} p_i^t$ represent the outcome distribution at time t, and let $\sigma = \frac{1}{T}\sum_{t=1}^{T}\sigma^t$ represent the time-averaged history of these distributions. Then, σ is an ϵ-approximate coarse correlated equilibrium, meaning that:*

$$\mathbb{E}_{\boldsymbol{s}\sim\sigma}[U_i(\boldsymbol{s})] \geq \mathbb{E}_{\boldsymbol{s}\sim\sigma}[U_i(s_i', \boldsymbol{s}_{-i})] - \epsilon,$$

for every player i and any unilateral deviation s_i'.

Proof. By definition, for each player i:

$$\mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(\mathbf{s})] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\mathbf{s}\sim\sigma^t}[U_i(\mathbf{s})],$$

where the right-hand side of this equation represents the time-averaged expected cost for player i when following the no-regret algorithm. In addition,

$$\mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(s_i', \mathbf{s}_{-i})] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\mathbf{s}\sim\sigma^t}[U_i(s_i', \mathbf{s}_{-i})],$$

where the right-hand side of this equation represents the time-averaged expected cost for player i when playing the fixed action s_i' in each iteration. Since the regret for each player is at most ϵ, and we are dealing with a utility-maximization game, the second expression is at most ϵ larger than the first:

$$\mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(s_i', \mathbf{s}_{-i})] \leq \mathbb{E}_{\mathbf{s}\sim\sigma}[U_i(\mathbf{s})] + \epsilon$$

This confirms the condition for an approximate coarse correlated equilibrium within the framework of utility-maximization games. □

Respectively, the time-averaged play history of joint play under no-swap regret dynamics converges to a correlated equilibrium.

Proposition 4.2. *Suppose that after T iterations of no-swap-regret dynamics, each player in a utility-maximization game experiences a swap regret of at most ϵ for each of its switching functions. Let $\sigma^t = \prod_{i=1}^{k} p_i^t$ denote the outcome distribution at time t and $\sigma = \frac{1}{T}\sum_{t=1}^{T}\sigma^t$ the time-averaged history of these distributions. Then σ is an ϵ-approximate correlated equilibrium, in the sense that:*

$$\mathbb{E}_{\boldsymbol{s}\sim\sigma}[U_i(\boldsymbol{s})] \geq \mathbb{E}_{\boldsymbol{s}\sim\sigma}[U_i(\delta(s_i), \boldsymbol{s}_{-i})] - \epsilon,$$

for every player i and any switching function $\delta : S_i \to S_i$.

Chapter 5

Our Contribution

In this thesis, we investigate revenue in repeated sealed-bid first-price auctions, mostly within a Bayesian framework involving two or three bidders, simulated by software agents using no-regret and no-swap regret algorithms. In this setting, bidders acquire information about the actions of others over time and adjust their strategies accordingly. More specifically, we examine cases where bidders' valuations are independent and identically distributed (i.i.d.) random variables. This implies that each bidder draws their valuation from the same probability distribution, with no correlation between individual bidders' choices; instead, their valuations are mutually independent. Our objective is to deepen the understanding of how this framework functions, with a focus on addressing three open questions, to the best of our knowledge.

First, we aim to explore whether a "fake" bidder—essentially an agent working on behalf of the seller—could effectively function as a reserve price. Specifically, we wish to examine whether the inclusion of such a bidder would maximize the seller's revenue to the same extent as a conventional reserve price. Second, we seek to determine if there is any significant difference in revenue when utilizing no-regret algorithms as compared to no-swap regret algorithms. Finally, we aim to investigate whether the revenue generated by a first-price auction is equivalent to the expected revenue in a second-price auction, assuming both auctions operate under the same Bayesian conditions regarding bidders' valuations. Before delving into the details of our work, it is important to first explore the existing literature and how it relates to and informs our research.

5.1 Existing Literature

The study of auctions has been a longstanding topic in academic literature. As discussed in the earlier chapter on auctions, second-price auctions are generally simpler to understand and provide more predictable outcomes. However, there has been a noticeable shift from second-price to first-price auctions, particularly in the context of repeated auctions for online advertising [40]. Recent research has increasingly focused on the dynamics of first-price repeated auctions, especially in the context of online learning agents [44], [45]. Among the various works, we will primarily examine the findings of [43], not only because of its relevance to our research but also due to the foundational and extensive insights it provides in the field.

In [43], the authors examine the interaction of software agents, designed as regret-minimizing algorithms, in repeated first-price and second-price auctions on behalf of their users. The study demonstrates that, in second-price auctions with two bidders having different valuations and with Hedge agents that engage in a repeated second-price auction on their behalf with discrete bid levels, the higher-valued bidder's bids converge to a uniform distribution between the lower bidder's valuation and their own, while the lower bidder's to one that spans the entire range from zero to their private valuation. The higher bidder always wins, as expected, but pays strictly less than the second-highest price, which contradicts the traditional dominant-strategy outcome. The paper contrasts this with results from [42], which shows that when starting with a long pure-exploration phase, dominant strategies do emerge in second-price auctions. According to the authors, the failure to converge to dominant strategies can be explained by the existence of alternative Nash equilibria. In these, the first player arbitrarily bids well above the second price while the second player underbids arbitrarily. Since regret dynamics do not always favor dominant strategies, they often lead to outcomes associated with 'low-revenue' equilibria. This lower average price compared to the second price allows users with values between these points to benefit from inflating their reported value. For example, a player with 0.4 bidding truthfully against a player with 0.5 will always lose, but by exaggerating their value to 1, they can consistently win and secure strictly positive utility on average.

In repeated first-price auctions between two regret-minimizing agents, with valuations $v \geq w$, the expectation is that the dynamics converge to an equilibrium. The only pure Nash equilibrium occurs when the lower bidder bids close to their valuation and the higher bidder bids slightly more, leading to the higher bidder consistently winning at near the second price. Both mixed and correlated Nash equilibria in such auctions typically yield this outcome [53]. However, regret-minimizing dynamics do not necessarily converge to a Nash or correlated equilibrium. Instead, they may reach coarse correlated equilibria (CCE) [53], where the revenue can fall below the second price, and occasionally the lower bidder may even win. This phenomenon persists even when bid levels are discretized. For every CCE, there exist regret-minimizing algorithms that converge to it [54]. Agents can be designed to follow a pre-set schedule that converges to the desired CCE, and in case of any agent's deviation, they could revert to a standard regret-minimization strategy. Thus, regret-minimizing algorithms may not always lead to the second-price outcome. Nevertheless, algorithms in the "mean-based" class (as defined in prior work [55], consisting of algorithms which assign minimal probability to actions that previously performed poorly), may converge only to the second-price result. In this setting, it has been proven that in repeated first-price auctions with discrete ϵ-grid bidding levels, if the dynamics converge to any single distribution, the higher-valued player will almost always win and pay close to the second price. Unlike second-price auctions where initial exploration phases influence convergence to dominant strategies, first-price auctions reach similar outcomes without requiring this additional exploration phase. Consequently, for human users interacting with regret-minimizing agents, the observed auction dynamics resemble a second-price auction, which is incentive-compatible. This removes any long-term incentive for users to misreport their values to their agents.

5.2 Experimental Research

As seen in the previously discussed research, significant findings have been made regarding both first- and second-price auctions in deterministic settings, through the application of no-regret algorithms. However, further exploration is required in the Bayesian setting and also in the context of no-swap regret algorithms. Additionally, the use of software agents acting on behalf of bidders in repeated auctions is an established concept. However, the idea of misreporting valuations, which falls under the broader theme of deception, introduces new avenues for exploration. In our work, we adopt this core idea of deception but apply it in a different context. Here, there is no deception between bidders and their software agents; instead, the deception comes from a "fake" bidder, who manipulates the other bidders. With the foundation of existing research now established, we turn to our contributions.

To address our primary research questions, we conducted experimental investigations through computer simulations. Specifically, we used Python to model the repeated auction interactions, modeling each software agent—representing a bidder and their behavior—using object-oriented programming. The Hedge algorithm was employed for the simulations, with the tuning parameter set to $\varepsilon = \sqrt{\frac{\ln K}{T}}$, where K represents the total number of possible bids (each bid corresponding to an action), and T the number of iterations in the repeated auction. The utility was constrained between zero and one. When we refer to revenue, we are discussing the average revenue per auction. Importantly, bidders did not overbid (i.e., they never bid more than their actual valuation, which was capped at 1), and all bids were non-negative. As a result, revenue was bounded between 0 and 1.

We employed both no-regret and no-swap regret algorithms (using the reduction discussed in the chapter on online learning algorithms) to simulate how each player learns and adapts their strategies over time. In terms of the possible bids for each player, we experimented with various levels of discretization. The simplest discretization we used involved bids drawn from the set $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$. We also examined cases where bids were drawn from the set $\{0, 0.1, 0.2, ..., 0.9, 1\}$, as well as finer discretizations, such as $\{0, 0.01, 0.02, ..., 0.98, 0.99, 1\}$ and $\{0, 0.001, 0.002, ..., 0.999, 1\}$. Additionally, we considered scenarios where bidders had deterministic valuations and cases where their valuations were i.i.d. random variables. Throughout these simulations, our primary objective was to analyze the bid dynamics and evaluate the seller's revenue. The Python code for all the simulations can be accessed at the following link: `https://github.com/ntua-el18046/IntegratedMasterThesis`.

The following principles apply to all experiments presented in this thesis:

- **Utility** = actual value − bid
- **Loss** = − utility

The step-by-step process of the Hedge Algorithm under the utility framework is described below:

Algorithm 1 Hedge Algorithm

Initialize: $\forall i \in [K], W_1(i) = 1$
for $t = 1$ to T **do**
    Pick $i_t \sim W_t$, i.e., $i_t = i$ with probability $x_t(i) = \frac{W_t(i)}{\sum_j W_t(j)}$
    Receive utility $u_t(i_t)$
    Update weights $W_{t+1}(i) = W_t(i)e^{\epsilon u_t(i)}$
end for

In first-price auctions, taking into consideration both the maximum bid of the opponents and the player's true valuation, the utility for each possible bid $i \in K$ is computed as follows:

$$\text{utility}(i) = \begin{cases} \text{actualValue} - \text{bid[i]} & \text{if bid}[i] > \text{opponentBid} \\ \frac{1}{2}(\text{actualValue} - \text{bid[i]}) & \text{if bid}[i] = \text{opponentBid} \\ 0 & \text{otherwise} \end{cases}$$

In second-price auctions, considering both the maximum bid of the opponents and the player's true valuation, the utility for each possible bid $i \in K$ is computed as follows:

$$\text{utility}(i) = \begin{cases} \text{actualValue} - \text{opponentBid} & \text{if bid}[i] > \text{opponentBid} \\ \frac{1}{2}(\text{actualValue} - \text{opponentBid}) & \text{if bid}[i] = \text{opponentBid} \\ 0 & \text{otherwise} \end{cases}$$

5.2.1 Deterministic Valuations in First- and Second-Price Auctions - Reproduction of Results

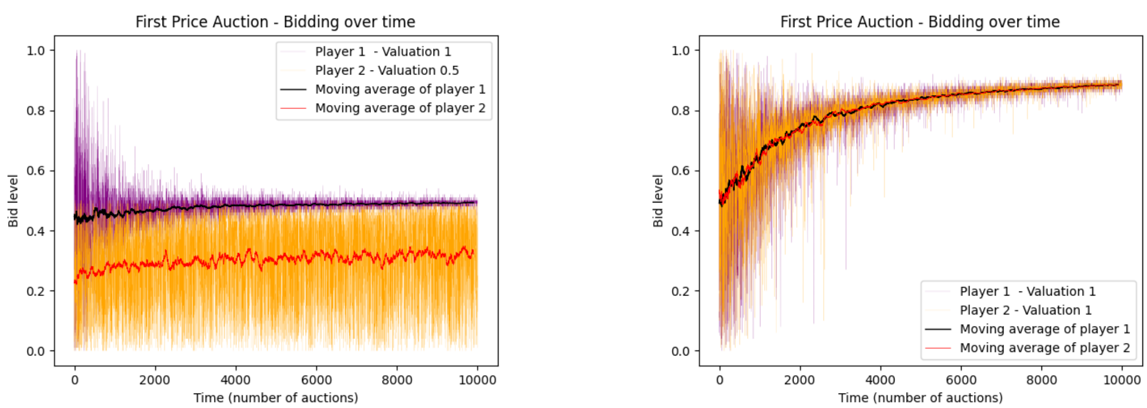
At the initial phase of our experimental research, we replicated the graphs depicting the bid dynamics in both first- and second-price repeated auctions with two bidders holding deterministic valuations, which were presented in the previously discussed paper [43]. In our configuration, we used the Hedge algorithm, $K = 100$ potential bids (ranging from 0 to 1 in increments of 0.01) and repeated the auction $T = 10,000$ times, employing a sliding window of size 100.

We implemented Python code to reproduce the results of [43] for two key purposes: first, to verify the findings of the aforementioned paper, and second, to ensure that our code functions as expected. Our results were entirely consistent with theirs:

- In the second-price repeated auction with two bidders where valuations are $w < v$, the high player bids uniformly in $(w, v]$, so the average bidding is around $w + \frac{(v-w)}{2}$, while the low player bids with full support on $[0, w]$, so their average bid level is $\frac{w}{2}$. The price paid by the higher bidder is strictly less than the valuation of the lower bidder.

- In the first-price repeated auction, the higher bidder almost always bids and pays approximately as much as the valuation of the low bidder.

Figure 5.1: Second-Price Repeated Auction
$v1 = 1$ and $v2 = 0.5$
Revenue $= 0.271$



(a) $v1 = 1$ and $v2 = 0.5$
Revenue $= 0.482$

(b) $v1 = 1$ and $v2 = 1$
Revenue $= 0.825$

Figure 5.2: First-Price Repeated Auctions

5.2.2 "Fake" Bidder as a Reserved Price

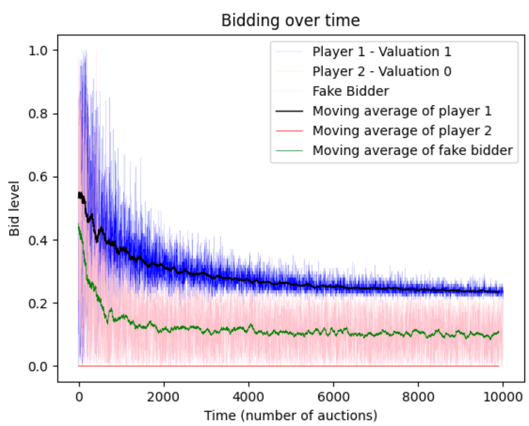
Following our examination of the two-bidder scenario in a deterministic setting, we proceeded to analyze the inclusion of a "fake" bidder, considering both deterministic and Bayesian contexts. We aimed to determine whether a "fake" bidder could contribute to maximizing revenue by serving as a reserve price in first-price auctions. This bidder is affiliated with the seller's team and aims to make other bidders to raise their bids.

Notably, this bidder does not intend to purchase the item, as it is already in their possession. To determine how the utility of the "fake" bidder is calculated, we employed two different approaches: one simpler and one more complex.

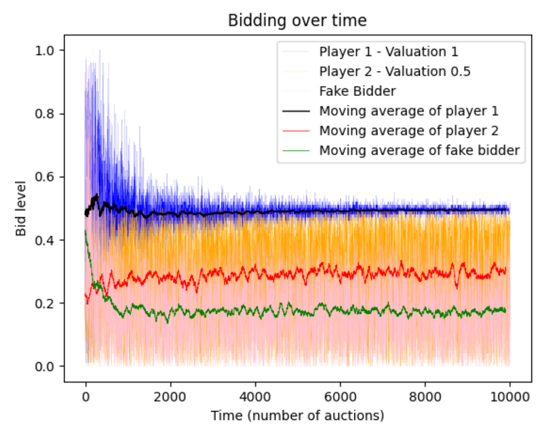
Simple bidder: In this scenario, if the "fake" bidder wins the auction, their utility is zero. Conversely, if they do not win the item—meaning another bidder has placed a higher or equal bid—their utility corresponds to the amount of the winning bid. This setup leads to a tendency for the "fake" bidder to select values that consistently lose, typically resulting in very low bids. Consequently, the "fake" bidder does not effectively encourage higher bids from the other players.

Smart Bidder: In this approach, the "fake" bidder again has a utility of zero if they win the auction, as their objective is not to win. Conversely, if they do not win the item—indicating that another bidder has submitted a higher or equal bid—their utility is equal to the amount of their bid. This structure ensures that higher non-winning bids yield greater utility than lower ones.

In all of our experiments involving a "fake" bidder, we measured the revenue generated in repeated first-price auctions, where each bidder is modeled by an instance of the Hedge algorithm. We initially applied the concept of a "fake bidder" in a deterministic setting concerning the valuations. The results are presented below.



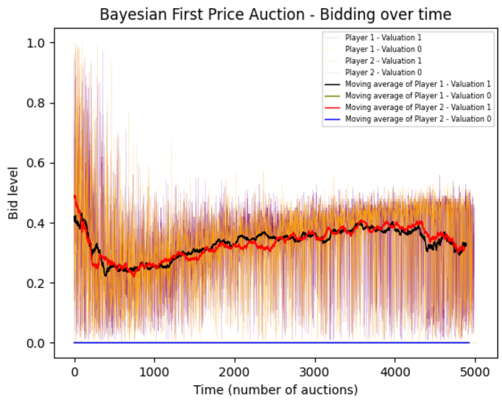
(a) $v1 = 1$ and $v2 = 0$
Revenue $= 0.271$



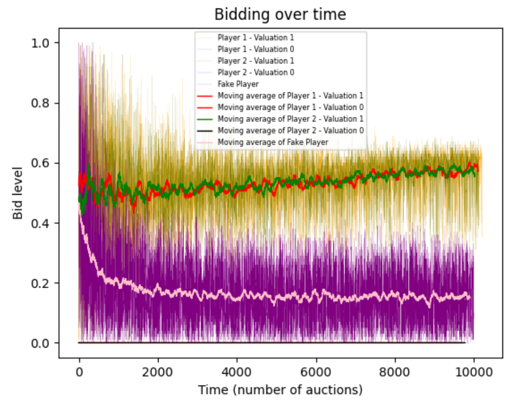
(b) $v1 = 1$ and $v2 = 0.5$
Revenue $= 0.489$

Figure 5.3: Simple Fake Bidder in a Deterministic Setting

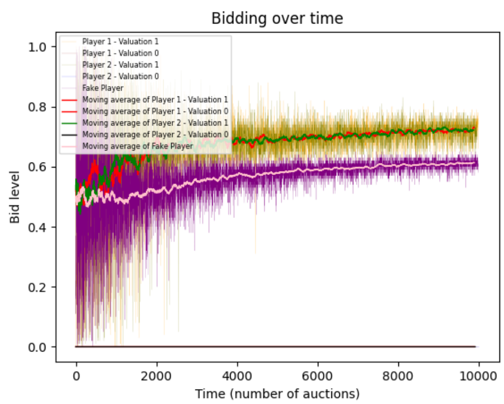
As observed, there are no significant differences in the deterministic setting; the "fake" bidder contributes negligibly to revenue enhancement. In contrast, the situation in the Bayesian setting is somewhat different. We will present our findings mostly through graphical representations. Our analysis included scenarios with two bidders or three bidders, each having two possible valuations. Firstly we present the case of two bidders and a "fake" bidder.



(a) No fake bidder
Revenue = 0.258



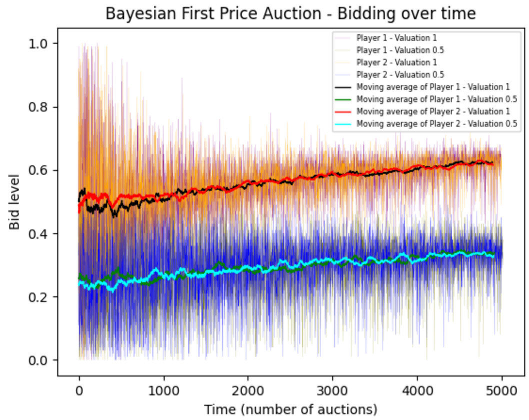
(b) A simple fake bidder
Revenue = 0.403



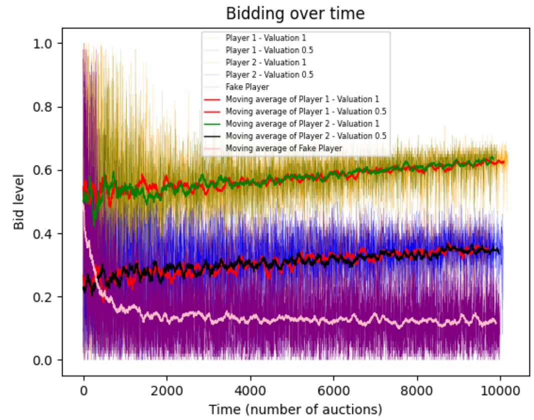
(c) A smart fake bidder
Revenue = 0.501

Figure 5.4: Bayesian First-Price Auctions
$v1 = 1$, $v2 = 0$ with probability $\frac{1}{2}$ each

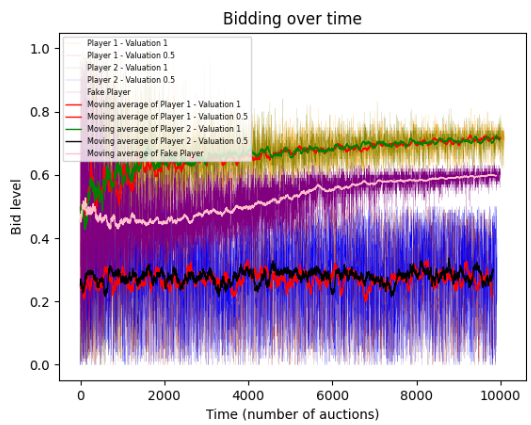
As observed in the graphs, when each bidder has a valuation of 1 with a probability of $\frac{1}{2}$ and a valuation of 0 with a probability of $\frac{1}{2}$, the revenue notably increases with the presence of a "fake" bidder. This increase is even more pronounced in the case of the smart "fake" bidder, as anticipated. We ran the experiment for $T = 1,000,000$ and the increase is even more prominent for both the simple and the smart "fake" bidder.



(a) No fake bidder
Revenue $= 0.519$



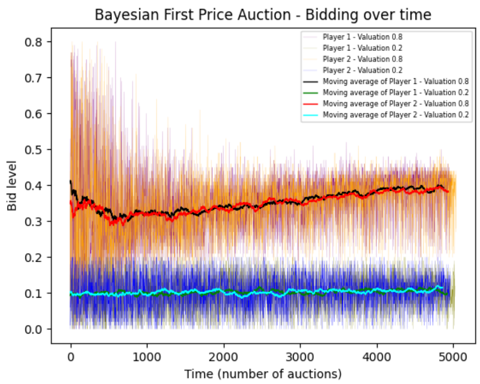
(b) A simple fake bidder
Revenue $= 0.532$



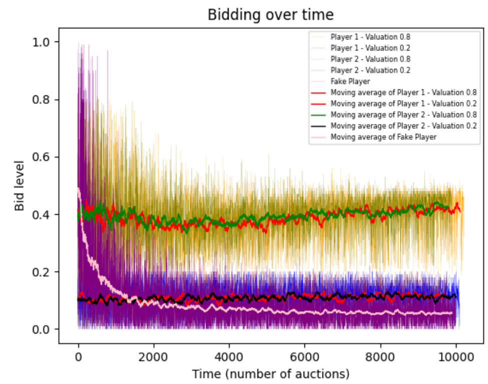
(c) A smart fake bidder
Revenue $= 0.502$

Figure 5.5: Bayesian First-Price Auctions
$v1 = 1$, $v2 = 0.5$ with probability $\frac{1}{2}$ each

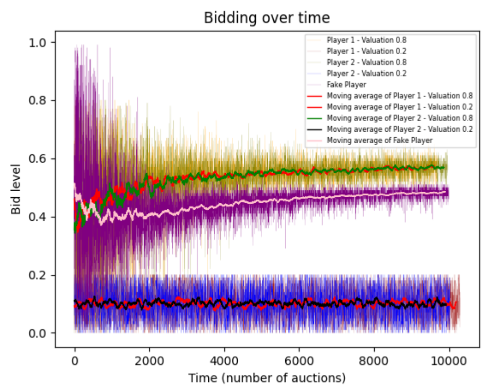
In this scenario, where each bidder has a valuation of 1 with a probability of $\frac{1}{2}$ and a valuation of 0.5 with a probability of $\frac{1}{2}$, the introduction of a “fake” bidder does not significantly enhance revenue. Surprisingly, the smart “fake” bidder does not yield as much revenue as the simpler version. This outcome is likely attributed to the limited time frame of our experiment. To test this assumption, we repeated the experiment for $T = 1,000,000$ times; the revenue generated by the smart “fake” bidder eventually surpassed that of the other two cases. However, the difference in revenue is not as pronounced as in the previous scenario, where the lower possible valuation was 0.



(a) No fake bidder
Revenue = 0.309



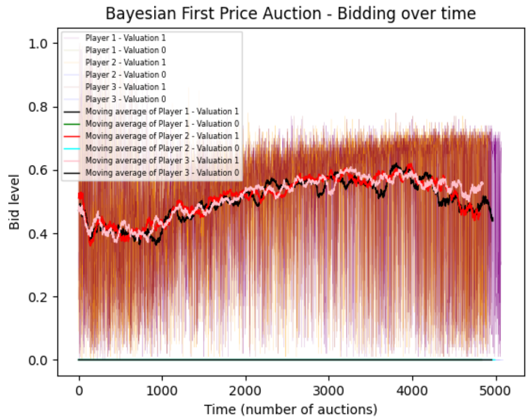
(b) A simple fake bidder
Revenue = 0.326



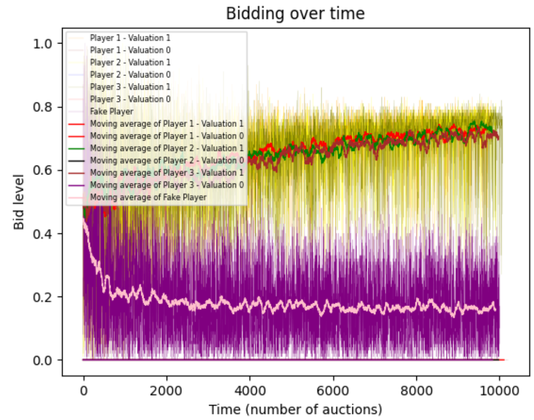
(c) A smart fake bidder
Revenue = 0.388

Figure 5.6: Bayesian First-Price Auction
$v1 = 0.8$, $v2 = 0.2$ with probability $\frac{1}{2}$ each

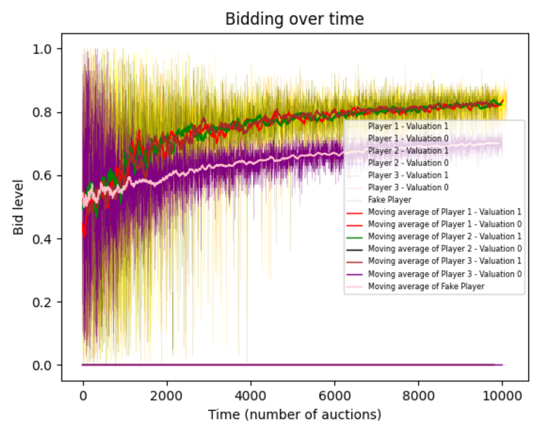
In this case, where each bidder has a valuation of 0.8 with a probability of $\frac{1}{2}$ and a valuation of 0.2 with a probability of $\frac{1}{2}$, the inclusion of a “fake” bidder does lead to an increase in revenue. However, the difference in revenue between auctions with and without a “fake” bidder is not as significant as in the scenario where one valuation was 0 and the other was 1. Instead, the difference is more pronounced compared to the case where one value was 0.5 and the other was 1.



(a) No fake bidder
Revenue = 0.507



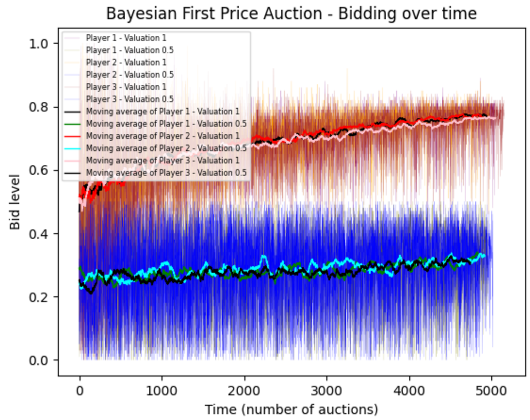
(b) A simple fake bidder
Revenue = 0.596



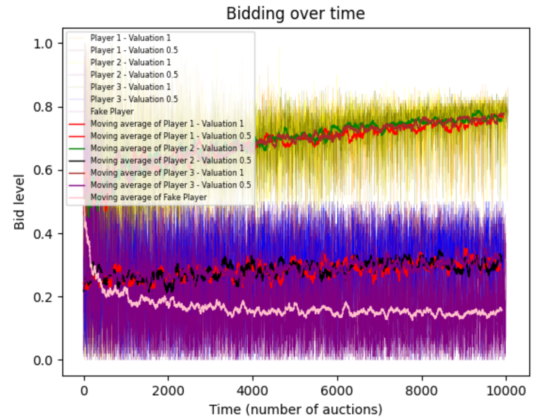
(c) A smart fake bidder
Revenue = 0.656

Figure 5.7: Bayesian First-Price Auction
$v1 = 1$, $v2 = 0$ with probability $\frac{1}{2}$ each

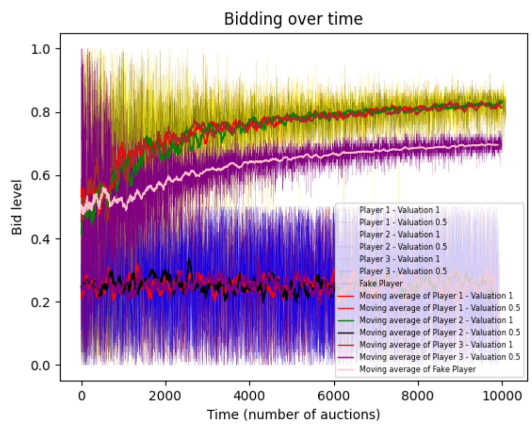
In the scenario involving three bidders, where each bidder has an equal probability of a valuation of either 1 or 0, the presence of a “fake” bidder enhances revenue. However, the increase in revenue is less pronounced than in the corresponding case with two bidders. Nonetheless, the absolute revenue figures in this three-bidder scenario are higher compared to those observed with two bidders.



(a) No fake bidder
Revenue = 0.676



(b) A simple fake bidder
Revenue = 0.675



(c) A smart fake bidder
Revenue = 0.664

Figure 5.8: Bayesian First-Price Auction
$v1 = 1$, $v2 = 0.5$ with probability $\frac{1}{2}$ each

In the scenario with three bidders, where each bidder has an equal probability of a valuation of either 1 or 0.5, the inclusion of a “fake” bidder does not lead to an increase in revenue, similar to the outcome observed with two bidders. To test our hypothesis that the smart “fake” bidder, given sufficient time, would yield higher revenue compared to the simple “fake” bidder (and also compared to the auction without any “fake” bidder), we extended the simulation to $T = 1,000,000$ repetitions. The results confirm our assumption. It is also worth noting that in the two-bidder case, the revenue is lower compared to the three-bidder scenario, which comes as no surprise.

Let’s now consider a Vickrey auction with a reserve price r and two i.i.d. bidders, each having one of two possible valuations, $v1$ and $v2$, where $0 \leq v2 \leq v1 \leq 1$, with each valuation occurring with a probability of $\frac{1}{2}$. The expected revenue, as a function of r, can be computed as follows:

- For $0 \leq r \leq v2$: $\mathbb{E}[Revenue] = \frac{3}{4} \cdot v2 + \frac{1}{4} \cdot v1$.
- For $v2 < r \leq v1$: $\mathbb{E}[Revenue] = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot r + \frac{1}{4} \cdot v1$. In that case, we easily observe that the revenue is maximized when $r = v1$ and is equal to $\frac{3}{4} \cdot v1$.
- For $v1 < r \leq 1$: $\mathbb{E}[Revenue] = 0$.

Thus, the optimal reserve price is equal to $v1$ when:

$$\frac{3}{4} \cdot v2 + \frac{1}{4} \cdot v1 \leq \frac{3}{4} \cdot v1, \text{which holds when } v1 \geq \frac{3}{2} v2.$$

Otherwise, the optimal reserve price can be any value in $[0, v2]$.

We will now calculate the revenue in detail for the case where v1 = 1 and v2 = 0.5, to further clarify this method:

- For $0 \leq r \leq 0.5$: $\mathbb{E}[Revenue] = \frac{3}{4} \cdot 0.5 + \frac{1}{4} \cdot 1 = \frac{5}{8}$.
- For $0.5 \leq r \leq 1$: $\mathbb{E}[Revenue] = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot r + \frac{1}{4} \cdot 1$. The expected revenue in this case is maximized for $r = 1$, where it is equal to $\frac{3}{4} = 0.75$.

Thus, since $\frac{5}{8} < \frac{3}{4}$, the optimal reserve price is equal to 1 and the corresponding expected revenue equal to 0.75. This comes as no surprise, since $1 \geq \frac{3}{2} \cdot 0.5 = 0.75$.

Following the same procedure, we calculated the optimal reserve prices and their corresponding expected revenues in the case of two i.i.d. bidders with valuations v1 = 1 and v2 = 0, as well as v1 = 0.8 and v2 = 0.2, with equal probability each. The previous computational procedure can be easily extended in the case of three bidders, with same valuation distributions as in the case of two bidders. The results of our calculations in comparison with the results from the simulations with the "fake" bidder are presented below:

Valuations	Optimal Reserved Price	Expected Revenue	Simple Fake Bidder	Smart Fake Bidder
v1 = 1 and v2 = 0	1	0.75	0.403	0.501
v1 = 1 and v2 = 0.5	1	0.75	0.532	0.502
v1 = 0.8 and v2 = 0.2	0.8	0.6	0.326	0.388

Table 5.1: Two IID bidders - $T = 10,000$

Valuations	Optimal Reserved Price	Expected Revenue	Simple Fake Bidder	Smart Fake Bidder
v1 = 0 and v2 = 1	1	0.875	0.596	0.656
v1 = 0.5 and v2 = 1	1	0.875	0.675	0.664

Table 5.2: Three IID bidders - $T = 10,000$

Based on our results, the following assumptions appear to hold:

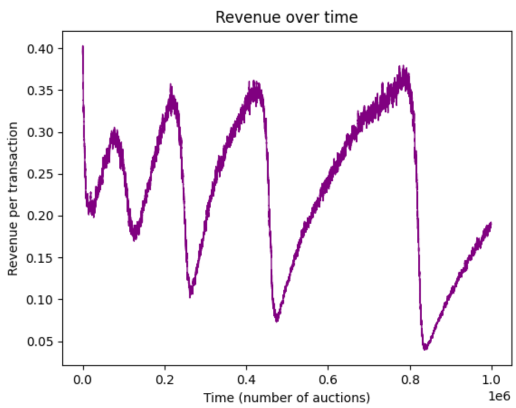
- The higher valuation influences the upper limit of the revenue; specifically, as the valuation increases, the potential revenue also increases.

- The disparity between the higher and lower valuations impacts the revenue difference between scenarios with and without a "fake" bidder; the greater the difference between the two valuations, the more pronounced the effect of the "fake" bidder on revenue.

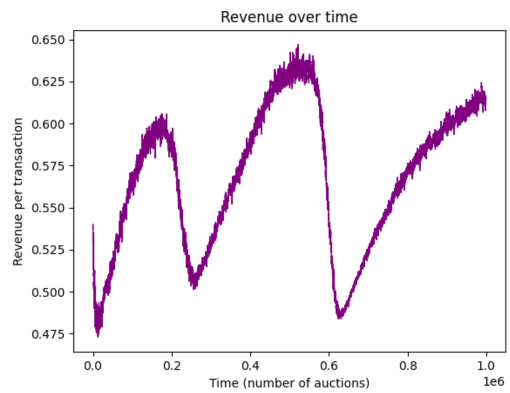
- The introduction of a "fake" bidder does not yield revenue as high as that generated by the optimal reserve price.

5.2.3 Revenue Evolution within the Bayesian Setting

In this section, we visually analyze the evolution of revenue in a Bayesian context. In this scenario, each bidder has two possible valuations, with each valuation occurring with equal probability $\frac{1}{2}$. The analysis encompasses $T = 1,000,000$ repeated auctions and utilizes a sliding window of 1,000. The results are displayed below.



(a) $v1 = 1$ and $v2 = 0$
Revenue = 0.23



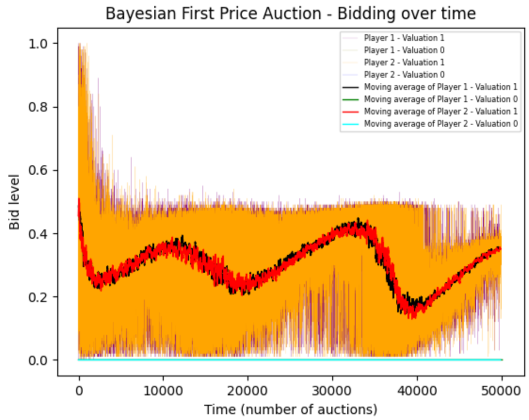
(b) $v1 = 1$ and $v2 = 0.5$
Revenue = 0.57

Figure 5.9: Revenue of Bayesian First-Price Auction over time

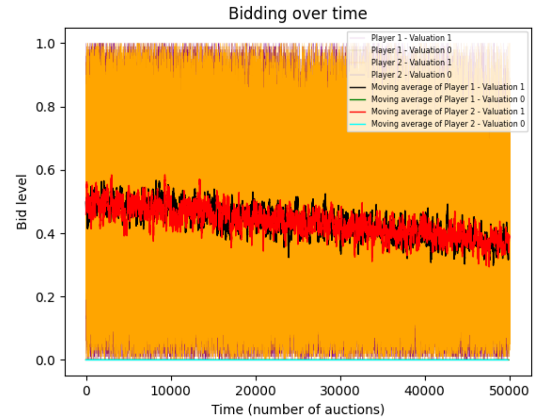
As observed, the revenue fluctuates, with the periods of oscillation growing longer over time. Nevertheless, the average expected revenue remains relatively stable, showing no significant variation as time progresses.

5.2.4 No-Regret Versus No-Swap Regret

In this section we aim to compare the differences in the revenue between bidders that use no-regret algorithms and bidders that use no-swap regret algorithms to simulate their learning process and decision-making regarding upcoming biddings. The auction was repeated $T = 100,000$ times.

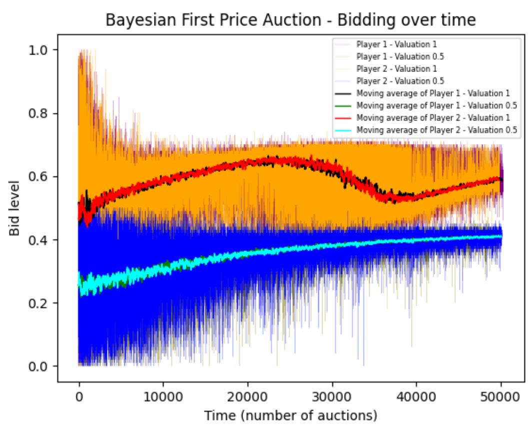


(a) No-Regret
Revenue = 0.239

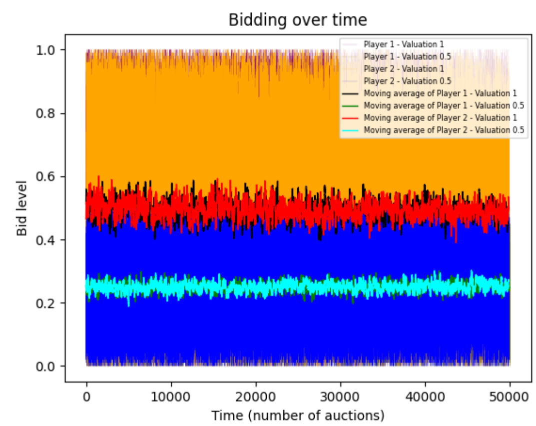


(b) No-Swap Regret
Revenue = 0.365

Figure 5.10: No-Regret and No-Swap Regret
$v1 = 1$ and $v2 = 0$ with probability $\frac{1}{2}$ each



(a) No-Regret
Revenue = 0.543



(b) No-Swap Regret
Revenue = 0.511

Figure 5.11: No-Regret and No-Swap Regret
$v1 = 1$, $v2 = 0.5$ with probability $\frac{1}{2}$ each

As observed, the revenue outcomes are not conclusive. Specifically, when both valuations are equally possible at 1 and 0, the no-swap regret scenario generates higher revenue. Conversely, with equally possible valuations of 1 and 0.5, the no-regret scenario yields greater revenue.

5.2.5 Revenue of Bayesian First-Price Auctions vs Expected Second-Price

Our final, yet pivotal question explores whether revenue in a symmetric Bayesian setting aligns with the expected second price, when bidders are simulated by no-swap regret (or no-regret) software agents. As mentioned in the theory chapters, in a Bayes-Nash equilibrium, there is revenue equivalence between first- and second-price auctions. Therefore, this final question could be reframed as "exploring whether our framework converges to a Bayes-Nash equilibrium". However, it is important to note that while the revenue could be equivalent, this does not necessarily indicate convergence to such an equilibrium.

The tables below summarize the revenue generated by both no-regret and no-swap regret algorithms across various scenarios, incorporating the expected second price for direct comparison.

Valuations	Expected Second Price	No-regret	No-Swap Regret
v1 = 1 and v2 = 0	0.25	0.261	0.401
v1 = 1 and v2 = 0.5	0.625	0.517	0.514
v1 = 0.8 and v2 = 0.2	0.35	0.311	0.364

Table 5.3: Two IID bidders - $T = 10,000$

Valuations	Expected Second Price	No-regret	No-Swap Regret
v1 = 1 and v2 = 0	0.25	0.239	0.365
v1 = 1 and v2 = 0.5	0.625	0.543	0.511
v1 = 0.8 and v2 = 0.2	0.35	0.317	0.35

Table 5.4: Two IID bidders - $T = 100,000$

Valuations	Expected Second Price	No-regret	No-Swap Regret
v1 = 1 and v2 = 0	0.5	0.498	0.522
v1 = 1 and v2 = 0.5	0.75	0.673	0.606

Table 5.5: Three IID bidders - $T = 10,000$

Valuations	Expected Second Price	No-regret	No-Swap Regret
v1 = 1 and v2 = 0	0.5	0.473	0.502
v1 = 1 and v2 = 0.5	0.75	0.704	0.609

Table 5.6: Three IID bidders - $T = 100,000$

As observed, the results are not definite. More specifically:

- In the scenario where each bidder's valuation is either 1 or 0, no-swap regret algorithms appear to generate higher revenue than both no-regret algorithms and the corresponding second-price auction. However, the revenue produced by no-swap regret algorithms tends to decline as the number of iterations increases, preventing any definitive conclusion.

- When each bidder's valuation is either 1 or 0.5, no-swap regret algorithms seem to generate the lowest revenue. Additionally, both first-price auctions involving regret-minimizing agents yield less revenue than the corresponding second-price auction.

- In the case where each bidder's valuation is either 0.8 or 0.2, it appears that, as the number of iterations increases, the revenue generated by no-swap regret algorithms converges to the expected revenue of the corresponding second-price auction.

5.2.6 Summary of Our Experimental Results

In conclusion, our findings indicate that while a "fake" bidder does contribute to increased revenue, it does not generate as much revenue as the optimal reserve price. Furthermore, there is no clear advantage in terms of revenue when using no-regret algorithms over no-swap regret algorithms, or vice versa, as the results vary across different scenarios. Finally, the relationship between the revenue in a Bayesian setting of repeated first-price auctions, where bidders are simulated using no-swap regret (or no-regret) agents, and the expected revenue in the corresponding second-price auction is not straightforward, with varying outcomes observed depending on the case.

Bibliography

[1] M. Shubik, *The Theory of Money and Financial Institutions.* MIT Press, 2004.

[2] V. L. Smith, "Auctions," *Allocation, Information and Markets*, pp. 39–53, 1989.

[3] R. Cassady, *Auctions and Anctioneering.* University of California Press, 1967.

[4] T.-O. Leautier, "Electricity auctions," *Journal of Economics*, vol. 10, pp. 331–358, 2001.

[5] J. Winkler, M. Magosch, and M. Ragwitz, "Effectiveness and efficiency of auctions for supporting renewable electricity – what can we learn from recent experiences?," *Renewable Energy*, vol. 119, pp. 473–489, 2018.

[6] S. Schöne, "Auctions in the electricity market," *Lecture notes in economics and mathematical systems*, pp. 41–203, 2009.

[7] O. Ashenfelter and K. Graddy, "Chapter 26 art auctions," *Handbook of the Economics of Art and Culture*, pp. 909–945, 2006.

[8] V. I. Afualo and J. McMillan, *Auctions of Rights to Public Property.* Palgrave Macmillan UK, 2002.

[9] H. Pleines, "Manipulating politics: Domestic investors in ukrainian privatisation auctions 2000–2004," *Europe-Asia Studies*, vol. 60, pp. 1177–1197, 2008.

[10] O. Hart, R. La Porta Drago, F. Lopez-de Silanes, and J. Moore, "A new bankruptcy procedure that uses multiple auctions," *European Economic Review*, vol. 41, pp. 461–473, 1997.

[11] M. O. Ball, F. Berardino, and M. Hansen, "The use of auctions for allocating airport access rights," *Transportation Research Part A: Policy and Practice*, vol. 114, pp. 186–202, 2018.

[12] D. Lucking-Reiley, "Vickrey auctions in practice: From nineteenth-century philately to twenty-first-century e-commerce," *Journal of Economic Perspectives*, vol. 14, pp. 183–192, 2000.

[13] O. Ashenfelter, "How auctions work for wine and art," *Journal of Economic Perspectives*, vol. 3, pp. 23–36, 1989.

[14] D. G. Gregg and S. Walczak, "E-commerce auction agents and online-auction dynamics," *Electronic Markets*, vol. 13, pp. 242–250, 2003.

[15] W. Wang, Z. Hidvégi, and A. B. Whinston, "Designing mechanisms for e-commerce security: An example from sealed-bid auctions," *International Journal of Electronic Commerce*, vol. 6, pp. 139–156, 2001.

[16] M. R. Vicente, "Determinants of c2c e-commerce: an empirical analysis of the use of online auction websites among europeans," *Applied Economics Letters*, vol. 22, pp. 978–981, 2014.

[17] S. Despotakis, R. Ravi, and A. Sayedi, "First-price auctions in online display advertising working paper first-price auctions in online display advertising," 2019.

[18] A. Ghosh, P. McAfee, K. Papineni, and S. Vassilvitskii, "Bidding for representative allocations for display advertising," *Lecture Notes in Computer Science*, pp. 208–219, 2009.

[19] A. Sayedi, "Real-time bidding in online display advertising," *Marketing Science*, vol. 37, pp. 553–568, 2018.

[20] A. Ghosh and A. Sayedi, "Expressive auctions for externalities in online advertising," *CiteSeer X (The Pennsylvania State University)*, 2010.

[21] T. Payne, E. David, N. R. Jennings, and M. Sharifi, "Auction mechanisms for efficient advertisement selection on public displays," *European Conference on Artificial Intelligence*, pp. 285–289, 2006.

[22] C. Wall, "The english auction: Narratives of dismantlings," *Eighteenth-Century Studies*, vol. 31, pp. 1–25, 1997.

[23] G. Tagliari Evangelista, J. Ferreira Lopes, G. Bruno Fornar, R. Pedroso Oaigen, T. Lopes Gonçalves, T. Esteves de Oliveira, L. Kluwe de Aguiar, and J. O. Jardim Barcellos, "Key factors influencing the sale of bulls in livestock auctions," *Revista Mexicana de Ciencias Pecuarias*, vol. 10, pp. 610–622, 2019.

[24] A. Szyszka, "Charity auctions from an economic perspective — literature review," *Studenckie Prace Prawnicze, Administratywistyczne i Ekonomiczne*, vol. 36, pp. 98–113, 2021.

[25] Y. Kamijo, "Bidding behaviors for a keyword auction in a sealed-bid environment," *Decision Support Systems*, vol. 56, pp. 371–378, 2013.

[26] Z. Katona and M. Sarvary, "The race for sponsored links: Bidding patterns for search advertising," *Marketing Science*, vol. 29, pp. 199–215, 2010.

[27] H. R. Varian, "Position auctions," *International Journal of Industrial Organization*, vol. 25, pp. 1163–1178, 2007.

[28] B. Lebrun, "Existence of an equilibrium in first price auctions," *Economic Theory*, vol. 7, pp. 421–443, 1996.

[29] T. Nedelec, C. Calauzènes, N. El Karoui, and V. Perchet, "Learning in repeated auctions," *Foundations and Trends® in Machine Learning*, vol. 15, 2022.

[30] Q. D. Lã, Y. H. Chew, and B.-H. Soong, "An introduction to game theory," *Springer eBooks*, pp. 3–22, 2016.

[31] B. Lucier, R. Paes Leme, and E. Tardos, "On revenue in the generalized second price auction," *Proceedings of the 21st international conference on World Wide Web*, 2012.

[32] N. Golrezaei, M. Lin, V. Mirrokni, and H. Nazerzadeh, "Boosted second-price auctions for heterogeneous bidders," *SSRN Electronic Journal*, 2017.

[33] S. Bikhchandani, "Reputation in repeated second-price auctions," *Journal of Economic Theory*, vol. 46, pp. 97–119, 1988.

[34] G. Tan and O. Yilankaya, "Equilibria in second price auctions with participation costs," *Journal of Economic Theory*, vol. 130, pp. 205–219, 2006.

[35] Y. Han, Z. Zhou, A. Flores, E. Ordentlich, and T. Weissman, "Learning to bid optimally and efficiently in adversarial first-price auctions," *arXiv (Cornell University)*, 2020.

[36] J. D. Hartline, "Mechanism design and approximation," 2014.

[37] J. Burkett and K. Woodward, "Reserve prices eliminate low revenue equilibria in uniform price auctions," *Games and Economic Behavior*, vol. 121, pp. 297–306, 2020.

[38] J. Li, X. Ni, Y. Yuan, R. Qin, X. Wang, and F.-Y. Wang, "The impact of reserve price on publisher revenue in real-time bidding advertising markets," *2017 IEEE International Conference on Systems, Man and Cybernetics*, 2017.

[39] S. Despotakis, R. Ravi, and A. Sayedi, "First-price auctions in online display advertising," *Journal of Marketing Research*, vol. 58, pp. 888–907, 2021.

[40] S. Goke, G. Y. Weintraub, R. A. Mastromonaco, and S. S. Seljan, "Bidders' responses to auction format change in internet display advertising auctions," *Proceedings of the 23rd ACM Conference on Economics and Computation*, 2022.

[41] R. Paes Leme, B. Sivan, and Y. Teng, "Why do competitive markets converge to first-price auctions?," *Proceedings of The Web Conference 2020*, 2020.

[42] Z. Feng, G. Guruganesh, C. Liaw, A. Mehta, and A. Sethi, "Convergence analysis of no-regret bidding algorithms in repeated auctions," *arXiv (Cornell University)*, 2020.

[43] Y. Kolumbus and N. Nisan, "Auctions between regret-minimizing agents," *Proceedings of the ACM Web Conference 2022*, 2022.

[44] Q. Wang, Z. Yang, X. Deng, and Y. Kong, "Learning to bid in repeated first-price auctions with budgets," *arXiv (Cornell University)*, 2023.

[45] X. Deng, X. Hu, T. Lin, and W. Zheng, "Nash convergence of mean-based learning algorithms in first price auctions," *Proceedings of the ACM Web Conference 2022*, 2022.

[46] E. Hazan, *Introduction to Online Convex Optimization, Second Edition.* MIT Press, 2022.

[47] G. Küchle, *Counterfactuals, rationality and equilibrium concepts in game theory.* Kovač, 2009.

[48] R. W. Rosenthal, "A class of games possessing pure-strategy nash equilibria," *International Journal of Game Theory*, vol. 2, pp. 65–67, 1973.

[49] H. Moulin and J. P. Vial, "Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon," *International Journal of Game Theory*, vol. 7, pp. 201–221, 1978.

[50] D. M. Kreps, "Nash equilibrium," *Game Theory*, pp. 167–177, 1989.

[51] W. Vickrey, *Public economics : selected papers.* Cambridge University Press, 1994.

[52] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," *The Journal of Finance*, vol. 16, pp. 8–37, 1961.

[53] M. Feldman, B. Lucier, and N. Nisan, "Correlated and coarse equilibria of single-item auctions," *Lecture notes in computer science*, pp. 131–144, 2016.

[54] B. Monnot and G. Piliouras, "Limits and limitations of no-regret learning in games," *The Knowledge Engineering Review*, vol. 32, 2017.

[55] M. Braverman, J. Mao, J. Schneider, and M. Weinberg, "Selling to a no-regret buyer," *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018.