

Improving Farmers' and Consumers' Welfare in Agricultural Supply Chains via Data-driven Analytics and Modeling: From Theory to Practice

by

Somya Singhvi

B.S., Cornell University (2015)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020

© Massachusetts Institute of Technology 2020. All rights reserved.

Author
Sloan School of Management
August 6, 2020

Certified by
Retsef Levi
J. Spencer Standish (1945) Professor of Operations Management
Thesis Supervisor

Certified by
Yanchong Zheng
Associate Professor of Operations Management
Thesis Supervisor

Accepted by
Georgia Perakis
William F. Pounds Professor of Management
Co-director, Operations Research Center

Improving Farmers' and Consumers' Welfare in Agricultural Supply Chains via Data-driven Analytics and Modeling: From Theory to Practice

by

Somya Singhvi

Submitted to the Sloan School of Management
on August 6, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

The upstream parts of the agricultural supply chain consists of millions of smallholder farmers who continue to suffer from extreme poverty. The first stream of research in this thesis focuses on online agri-platforms which have been launched to connect geographically isolated markets in many developing countries. This work is in close collaboration with the state government of Karnataka in India which launched the Unified Market Platform (UMP). Leveraging both public data and platform data, a difference-in-differences analysis in Chapter 2 suggests that the implementation of the UMP has significantly increased modal price of certain commodities (5.1%-3.5%), while prices for other commodities have not changed. The analysis provides evidence that logistical challenges, bidding efficiency, market concentration, and price discovery process are important factors explaining the variable impact of UMP on prices. Based on the insights, Chapter 3 describes the design, analysis and field implementation of a new two-stage auction mechanism. From February to May 2019, commodities worth more than \$6 million (USD) had been traded under the new auction. Our empirical analysis suggests that the implementation has yielded a significant 4.7% price increase with an impact on farmer profitability ranging 60%-158%, affecting over 10,000 farmers who traded in the treatment market.

The second stream of research work in the thesis turns to consumer welfare and identifies effective policies to tackle structural challenges of food safety and food security that arise in traditional agricultural markets. In Chapter 4, we develop a new modeling framework to investigate how quality uncertainty, supply chain dispersion, and imperfect testing capabilities jointly engender suppliers' adulteration behavior. The results highlight the limitations of only relying on end-product inspection to deter EMA and advocate a more proactive approach that addresses fundamental structural problems in the supply chain. In Chapter 5, we analyze the issue of artificial shortage, the phenomenon that leads to food security risks where powerful traders strategically withhold inventory of essential commodities to create price surge in the market. The behavioral game-theoretic models developed allow us to examine the effectiveness of common government interventions. The analysis demonstrates the disparate effects of different interventions on artificial shortage; while supply allocation schemes often mitigate shortage, cash subsidy can inadvertently aggravate shortage in the market. Further, using field data from onion markets of India, we structurally estimate that 10% of the total supply is being hoarded by the traders during the lean season.

Thesis Supervisor: Retsef Levi

Title: J. Spencer Standish (1945) Professor of Operations Management

Thesis Supervisor: Yanchong Zheng
Title: Associate Professor of Operations Management

Acknowledgements

First of all, I feel truly blessed to have Retsef Levi and Karen Zheng as my advisors. Your passion for impactful research, innovative ideas and infectious energy is very inspiring. I am profusely grateful for your flexibility and encouragement to let me work on problems that interested me. You taught me everything I know about research and your dedication towards work has left an indelible impression on me. I truly cherish all the fond memories, and eye-opening experiences from our field trips to India. I have learnt a lot from you over the years, and hope to continue doing so in the future.

Next, I would like to express my gratitude towards my thesis committee members: David Shmoys and Nikos Trichakis. Your insightful comments and feedback have significantly improved this manuscript. David, you motivated me to do a PhD in the first place and I want to sincerely thank you again for mentoring me since my undergraduate years at Cornell University. A special thank you to Georgia Perakis. You have been infinitely kind and caring, a pillar of support who stood behind me and Divya during the most difficult times.

There are many professors at Cornell University and MIT who have mentored me over the years. I would like to thank my undergraduate mentors Peter Frazier, Shane Henderson, Andreea Minca, Eoin O'Mahony and Dawn Woodard for guiding me during my initial years as an undergraduate researcher at Cornell University. Sincere thanks also to Dimitris Bertsimas, Daniel Freund, Negin Golrezai, Jonas Jonasson and Joann de Zegher for valuable advice that was very helpful during the academic job market process. Thank you also to Laura Rose and Andrew Carvalho for their timely help in all administrative matters during my time at the ORC.

Collaboration with field partners has been another milestone of my PhD. I will be forever grateful to Manoj Rajan and Rikin Gandhi for believing in me and assisting with field implementation. Sincere thanks to Manik Joshi, Mangala Patil, Nisha Devra, Aditya Sethi, Daniel Quinn and other staff members at ReMS and Digital Green for helping me during my field trips.

Thank you also to Robert Stoner, Chintan Vishnav and Diane Rigos at the Tata Center for Technology and design at MIT for teaching me the right set of tools to conduct field research in India.

My time at MIT would not be the same without the supportive set of friends at the ORC, Ashdown House and Tata Center. I would like to thank Shwetha-Paul, Deeksha-Pradeep, Mukund, In Young, Tamar-Jonathan and Nidhi for being there through thick and thin in the last five years. I will dearly miss our impromptu icecreams, delicious food and memorable trips. Thank you also to Ricardo, Eduardo, Brad, Kriti, Elizabeth, Agni, Rajan, Nishanth, Lennart, Michael, Rim, Daniel, Peter, Andrew, Mathew, Swati-Tushar, Velibor, Hamsa, Mohit, Rajesh, Rakshith, Chinmay and Vaishnavi for the interesting conversations during my time at MIT.

Finally, I owe my deepest gratitude to my family. To my grandparents and parents for their selfless love, unwavering optimism and endless patience. Your immense sacrifices have made this journey possible and no words will ever suffice to express my gratitude. Bhaiya and Bhabhi, thank you for your impeccable advice that has been instrumental in guiding me during testing times. Dhvani, thank you for the surprise ice creams and chocolates that kept me and Divya going, I have found a true friend in you. Thanks also to Ridhima and her family who stood like a rock behind me in the last few months before my defense. Finally, thank you Divya, it is hard to imagine this thesis (and my life) without you. You have been my best friend who supported, encouraged and helped me get through stressful times in the most positive way.

Contents

1	Introduction	17
1.1	Contributions of Chapter 2	20
1.2	Contribution of Chapter 3	20
1.3	Contribution of Chapter 4	21
1.4	Contribution of Chapter 5	22
2	The Impact of Unifying Agricultural Wholesale Markets on Market Prices and Farmers' Profitability	23
2.1	Introduction	23
2.2	Traditional Markets versus UMP	24
2.3	Data and Empirical Approach	26
2.3.1	DID model	26
2.3.2	Robustness tests	27
2.3.3	DID assumption	27
2.3.4	Falsification test	27
2.4	The Impact of UMP	28
2.4.1	Heterogeneous Impacts by Product Quality	29
2.5	Systemic Differences between High-impact and Low-impact Groups	31
2.5.1	Cross-Market Competition	31
2.5.2	In-Market Competition	32
2.5.3	In-Market Concentration	34
2.5.4	Price Discovery Process	35
2.6	Conclusions and Discussion	36
3	Improving Farmers' Income on Online Agri-platforms: Theory and Field Implementation of a Two-stage Auction	37

3.1	Introduction	37
3.1.1	Related Literature and Contributions	40
3.2	Operational Context and Design of the Two-stage Auction	41
3.3	Traders' Bidding Behavior	43
3.3.1	Theoretical Predictions	43
3.3.2	Field Interviews	45
3.4	A Behavioral Auction Model	47
3.5	Implementing and Empirically Evaluating the Impact of the Two-Stage Auction	50
3.5.1	Field Implementation	50
3.5.2	Comparability of the Treatment and Control Markets	52
3.5.3	Parallel Trends Assumption	52
3.5.4	DID Model	53
3.5.5	Robustness Tests	55
3.5.6	Validating the Behavioral Auction Model	57
3.5.7	Additional Insights	59
3.6	Conclusions	62
4	Economically Motivated Adulteration in Farming Supply Chains	64
4.1	Introduction	64
4.2	Modeling the Farming Supply Chain	69
4.3	When Will Farms Adulterate?	75
4.3.1	Preemptive EMA	75
4.3.2	Reactive EMA	78
4.4	The Effect of Supply Chain Dispersion on EMA Risk	83
4.5	The Effect of Quality Uncertainty on EMA Risk	84
4.6	Relating Model Predictions to Empirical Observations	87
4.6.1	Preemptive EMA: Misuse of Antibiotics in Poultry Farming in China	87
4.6.2	Reactive EMA: Melamine-Tainted Infant Formula Scandal	89
4.7	Analyzing Managerial Levers to Mitigate EMA Risk	90
4.8	Conclusions	92
5	Artificial Shortage in Agricultural Supply Chains	95
5.1	Introduction	95

5.1.1	Contributions and Policy Insights	99
5.2	Related Literature	101
5.3	Base Model: Artificial Shortage under No Government Intervention	102
5.4	Modeling Government Interventions	105
5.4.1	Modeling Consumer Welfare under Shortage-induced Budget Adjustment	106
5.5	Analysis of Government Interventions	108
5.5.1	Cash Subsidy (CS)	108
5.5.2	Supply Allocation Schemes	113
5.6	Policy Insights: Which Intervention is Most Effective?	117
5.7	Relating Model Predictions to Empirical Observations	121
5.8	Conclusions	124
A	Appendix of Chapter 2	126
A.1	Data Processing	126
A.2	Cross-market Trading	128
A.3	Dates of Integration	128
A.4	Parallel Trend Test	128
A.5	Robustness Tests	129
A.5.1	Alternative Specifications	129
A.5.2	Multiple Hypothesis Testing	130
A.5.3	Alternative Control Covariates	131
A.6	Falsification Test	132
A.7	Farmer Heterogeneity	132
A.8	Comparing High-impact and Low-impact Groups	135
A.9	Structural Changes in the Markets	137
A.9.1	Spillovers	138
A.10	Normalized Modal Prices of High-Impact and Low-Impact Group	141
B	Appendix of Chapter 3	142
B.1	Operational Feasibility of Dutch Auction	142
B.2	Additional Details of the Behavioral Auction Model	143
B.3	Proofs	144
B.4	Interview Questions	149

B.4.1	General Questions	149
B.4.2	Two-Stage Auction	149
B.4.3	Anticipated Regret	149
B.4.4	Competitive Arousal versus Anchoring	150
B.4.5	Budget Constraints	151
C	Appendix of Chapter 4	152
C.1	Analytical Details on Model Extensions	152
C.1.1	Preemptive EMA When Farms are Averse to Quality Uncertainty	152
C.1.2	EMA Risk When the Quality of All Units from the Same Farm is Perfectly Correlated	153
C.1.3	Reactive EMA with Decision on How Much to Adulterate	154
C.1.4	Alternative Models of Imperfect Testing Sensitivity	155
C.1.5	Alternative Penalty Structures	159
C.2	Investing in Traceability and Testing Frequency to Mitigate EMA Risk	161
C.3	Investing in Testing Capabilities to Mitigate EMA Risk	164
C.4	Proofs of Analytical Results	167
D	Appendix of Chapter 5	195
D.1	Model Notations	195
D.2	Additional Analytical Results	195
D.3	Model Extensions	197
D.3.1	Skewed Budget Distribution	197
D.3.2	Divisible Good	198
D.3.3	Trader with Inventory Holding Cost and Inventory Constraint	200
D.3.4	Unequal Relative Budget Increase in Period 2 Across Consumers	202
D.3.5	Alternative Government Interventions	203
D.3.6	Effect of Government Interventions on Consumers' Budget Adjustment	204
D.3.7	Reference Price Model	205
D.4	Proof of Analytical Results	206

List of Figures

2.1	Estimated average impact (in %) of UMP on the modal prices of different commodities from the Main model (Eq. [2.1]). Green (red) indicates a statistically significant (nonsignificant) impact. Numbers next to the bars are the estimates with standard errors in parentheses.	28
3.1	Rules Under the Two-Stage Auction	42
3.3	Comparing Treatment and Control Markets Pre and Post Implementation	52
3.4	Distribution of Bid Increments in the Second Stage	58
4.1	Sequence of Events for Preemptive and Reactive EMA [-.1in]	71
4.2	Reactive EMA: Farms' Optimal Adulteration Strategies under Perfect and Imperfect Testing	82
4.3	Effect of Quality Uncertainty (p_L) on the Expected Total Amount of Adulterated Output (E_n)	86
4.4	Effect of Traceability (t), Quality Uncertainty (p_L^{\max}), and Testing Frequency (q) on Preemptive EMA Risk	89
4.5	Effect of Traceability (t), High-Quality Unit Price (r_H), and Testing Frequency (q) on Reactive EMA Risk	91
5.1	Normalized Retail Price and Supply of Onions in India's Agricultural Markets . .	96
5.2	Dynamics of the Game	106
5.3	Evaluating Consumer Surplus with Shortage-Induced Budget Adjustment	107
5.4	Revenue Curve and Optimal Quantity for the Trader in Period 2 Under CS	110
5.5	Effect of Supply Allocation Schemes on Period 2 Price-Demand Curve	115

5.6 Effect of Government Budget (\mathbf{b}) on Equilibrium Period 1 Quantity (\mathbf{q}_T^1) under
 Different Government Interventions 118

5.7 Counterfactual Analysis 124

A.2 Number of Karnataka markets that were integrated into UMP since 2014. 128

C.1 Total EMA Risk under Perfect or Imperfect Testing for Preemptive EMA and
 Perfect Testing for Reactive EMA 167

D.1 Equilibrium Period 1 Quantity with Exponential Budget Distribution 199

D.2 Equilibrium Period 1 Quantity with Divisible Good 200

D.3 Equilibrium Period 1 Quantity under Inventory Constraint 201

D.4 Equilibrium Period 1 Quantity under Under Unequal Relative Budget Increase . 203

List of Tables

2.1	Estimated Impact of the UMP by Commodity	29
3.1	Summary of Theoretical Predictions and Results From Field Interviews	47
3.2	Summary Statistics for Tur Trading in the Control and Treatment Markets Pre- Implementation	52
3.3	Estimation Results for the Parallel Trends Test	53
3.4	Estimated Impact of Implementing the Two-stage Auction on Tur Prices	55
3.5	Regression Results of the Robustness Tests	57
3.6	Estimated Impact of the Implementation on the Highest First-Stage Bids	58
3.7	Estimated Impact of the Implementation on the Highest First-Stage Bids: Effect of Anticipated Non-qualification Regret	59
3.8	Estimated Impact of the Implementation on Tur Prices: The Effect of Quality	61
3.9	Estimated Impact of Implementing the Two-stage Auction on the Average and Standard Deviation of First-Stage Bids at the Trader Level	62
4.1	Notations	75
4.2	Effect of Increased Supply Chain Dispersion on EMA Risk	85
4.3	Effect of an Increased Probability of Producing Low-Quality Output (p_L^{\max} or p_L) on EMA Risk	86
4.4	Parameter Values for Guangdong and Shandong Provinces	87
4.5	Parameter Values for Sanlu Group and Bright Dairy	90
5.1	Effects of Increased Budget for Government Intervention on Artificial Shortage	120
5.2	Maximum Likelihood Estimation Results of Candidate Models	123
A.1	Total Number of Primary Markets Within and Outside Karnataka	127

A.2	Summary Statistics of Auction Data from UMP	127
A.3	Results from the Parallel Trend Test	129
A.4	Estimated Impact of UMP on the Prices of Different Commodities	130
A.5	Unadjusted and Adjusted p -Values (Accounting for Multiple Comparisons across Commodities for the Same Outcome)	131
A.6	Unadjusted and Adjusted p -Values (Accounting for Multiple Comparisons across Outcomes for the Same Commodity)	131
A.7	Estimated Impact of UMP: Specifications with Different Control Covariates . . .	132
A.8	Estimated Impact in 1,000 Replications of Falsification Tests	133
A.9	Estimated Differential Impacts of UMP by Farmer Characteristics	134
A.10	Correlation Matrix among Market Features for H2–H4	136
A.11	Systemic Differences between High-impact and Low-impact Groups	136
A.12	Change in Market Structure over Time	138
A.13	Change in the Fraction of Quantity Won by Home Traders over Time	138
A.14	Change in the Fraction of Quantity Purchased at Home Market by Cross-Market Traders over Time	139
A.15	Change in Quantity Distribution among Karnataka Markets over Time	140
C.1	Total EMA Risk When Farms Engage in Both Preemptive and Reactive EMA . .	166
D.1	Model Notations	195

Chapter 1

Introduction

Agricultural supply chains in developing countries share a unique set of characteristics. These supply chains are characterized by a large number of smallholder farmers in the upstream as the main producers, an opaque network of powerful traders in the middle who buy from farmers and eventually sell to consumers, and in the downstream an expansive market of consumers who have little control on the quality, volume, and price of the products available. Such a fragmented, nontransparent supply chain gives rise to several serious challenges. In the upstream, the phenomenon of severe poverty continues to persist among smallholder farmers in part due to imperfect competition and inefficient price discovery processes in traditional agricultural markets (Bergquist 2019, Goyal 2010). This prevalence of extreme poverty among smallholder farmers, and opaqueness in the upstream of the supply chain, also leads to *food safety* concerns for consumers in the downstream. In particular, economically motivated adulteration (EMA) by smallholder farmers, who work under extreme financial distress, is a major cause of *food safety* risks in the current times (Everstine et al. 2013). Finally, high market concentration in agricultural markets also leads to *food security* concerns as timely availability of essential food items at affordable prices is threatened in the markets. Artificial shortage by monopolistic traders of essential commodities, such as onions and lentils, to raise market prices has led to *food security* concerns in many countries across the world (Jain 2017, Hindustan Times 2019).

Because of the significant impact from these challenges, public organizations have responded in a variety of ways to protect farmer and consumer welfare. These initiatives have focused on developing new information technologies and sciences, new regulatory initiatives, as well as, designing subsidy schemes for consumers. For instance, to address the challenge of imperfect competition and inefficient price discovery processes in traditional agricultural markets, substan-

tial investments have been made to connect geographically distributed agri-markets through a single online agri-platform. Various countries have updated agri-laws to launch and enable such online agri-platforms that aim to transform traditional markets, with prominent examples such as the Ethiopia Commodity Exchange (ECX) and the eNational Agriculture Market (eNAM) in India (Ababa 2017, Nirmal 2017). Similarly, in order to mitigate EMA risks, regulatory initiatives, such as the Food Safety Modernization Act in the U.S., have been passed in many countries to improve testing methods and food sampling strategies (Johnson 2014). Further, to protect consumers' welfare during artificial shortages, governments across the world have spent a significant amount of resources on intervention programs. The interventions range from direct imports and/or inspections that increase supply in the markets, to cash subsidy aimed at increasing consumers' purchasing power during shortage (Chaki 2013, Business Standard 2015).

However, despite significant efforts, the success of these initiatives have been limited so far. Although it is postulated that building online agri-platforms can benefit farmers by increasing market competition, and enabling transparency, qualitative evidence suggests mixed results (Ababa 2017, Nirmal 2017). Similarly, regulatory focus towards sampling end products and inspecting downstream manufacturing facilities have not deterred EMA incidences in recent times across many countries (Tanzina and Shoeb 2016, Doyle et al. 2013). Even after substantial expenditure in consumer subsidies, instances of artificial shortage continue to significantly affect consumer welfare (Jain 2017, Hindustan Times 2019). Why have these initiatives not met with much success and how can we develop effective implementable solutions are open research questions of significant value to policy makers across the globe.

A key aspect that is understudied in the existing literature is the role of supply chain, operational and behavioral factors. The existing academic research in this domain mainly focuses on IT aspects and economic dynamics, and often aims at replicating successful examples from more affluent settings. For example, design of online agri-platforms is primarily done from an economic and information systems perspective. Similarly, most current research and practices related to food safety largely focus on testing in the downstream, a topic predominantly studied by biologists, chemists, and medical researchers. Further, practical implementation of many of the proposed solutions is difficult since solutions that are most effective in more developed settings (e.g., e-commerce in developed economies) can be operationally infeasible in resource constrained environments such as agri-systems of developing countries. This is in large part due to the distinctive operational and behavioral characteristics underlying these markets and

supply chain. Without carefully accounting for such operational and behavioral considerations, theoretically appealing solutions could seriously backfire.

This thesis uses data-driven analytics and modeling to identify practically-implementable and empirically-grounded solutions to some of the key challenges in agri-systems and makes contributions to both research and practice. It sheds light on the following important questions: (i) What is the impact of online agri-platforms on farmer welfare and how can their design be improved further to increase market and information access? (ii) How can we design policies to tackle structural issues of food safety and food security that are detrimental to consumer welfare and arise regularly in traditional agri-markets? Building on close collaborations with the state Government of Karnataka in India, the research work has yielded a substantial positive societal impact. In particular, profits for more than 10,000 smallholder lentil farmers have increased by approximately 150% from a field intervention on a large-scale online agri-platform run by the Karnataka government, and led to a number of practically important insights for a wide range of audience including commercial entities, policy makers, and nonprofits in the food and agriculture space.

The research work in this thesis also makes important methodological contributions and uses a multi-method approach that combines field research, empirical analysis, analytics and theoretical modeling to derive key insights. The analysis demonstrates that a systemic approach that, (i) is grounded in the field; (ii) incorporates key operations, behavioral and supply chain perspectives; (iii) and uses data-driven analytics and modeling is essential to optimize the design of current agri-systems. For example, Chapter 2 and Chapter 3, which investigate the impact and design of online-agri platforms, provide rigorous empirical evidence that the success of online agri-platforms critically depends on systemic supply chain, logistics and behavioral considerations that affect trades in the physical markets. The analysis in Chapter 4 and Chapter 5 introduces new behavioral and supply chain perspectives to the topics of *food safety* and *food security* that are traditionally studied by the disciplines of sciences and economics. Results from the thesis demonstrate instead that having a systemic understanding of the structural and environmental factors in the underlying supply chain is instrumental to identify fundamental weaknesses in the current system and to develop practical remedies. The key contributions from each of the chapters are described briefly next.

1.1 Contributions of Chapter 2

The first stream of research work focuses on online agri-platforms which have been launched in multiple developing countries to connect geographically isolated markets as a means to lift smallholder farmers out of poverty. This body of work is based on close collaboration with the state government of Karnataka in India which launched Unified Market Platform (UMP) to connect all of the regulated agri-markets in the state. Chapter 2 provides the first empirical assessment on the impact of the Unified Market Platform (UMP) – on market prices and farmers’ profitability. UMP was created in 2014 to unify all trades in the agricultural wholesale markets of the state to be carried out within a single platform. By November 2019, 62.8 million metric tons of commodities valued at \$21.7 billion (USD) have been traded on UMP. Employing a difference-in-differences method, we demonstrate that the impact of UMP on modal prices varies substantially across commodities. In particular, the implementation of UMP has yielded an average 5.1%, 3.6%, and 3.5% increase in the modal prices of paddy, groundnut, and maize. Furthermore, UMP has generated a greater benefit for farmers who produce higher-quality commodities. Given low profit margins of smallholder farmers (2%-9%), the range of profit improvement is significant (36%-159%). In contrast, UMP has no statistically significant impact on the modal prices of cotton, green gram, or tur. Utilizing detailed market data from UMP, we analyze how features related to logistical challenges, bidding efficiency, in-market concentration, and the price discovery process differ between commodities with and without a significant price increase due to UMP. These analyses lead to several policy insights regarding the design of similar agri-platforms in developing countries.

1.2 Contribution of Chapter 3

A key challenge identified from the analysis in Chapter 2 is that of limited competition in many agricultural markets. In order to tackle this challenge, Chapter 3 describes work conducted in close collaboration with the state government of Karnataka, India, to *design, implement, and assess* the impact of a new two-stage auction format on UMP. The design of the two-stage auction is informed by operational constraints, and guided by theory-informed, semi-structured interviews with a majority of the traders in the pilot market. The interviews suggest anticipated regret from non-qualification and anchoring as two important behavioral factors likely to affect the traders’ bidding strategies in a two-stage auction. A new behavioral auction model is

developed to capture these factors and determine *when* the two-stage auction can generate a higher revenue for farmers than the traditional single-stage, first-price, sealed-bid auction. The new auction mechanism was implemented on the UMP for a major market of lentils in February 2019. By June 2019, commodities worth more than \$6 million (USD) had been traded under the new auction. A difference-in-differences analysis demonstrates that the implementation has yielded a significant 4.7% price increase with an impact on farmer profitability ranging 60%–158%, affecting over 10,000 farmers who traded in the treatment market. The results from this chapter offer tangible insights on how innovative price discovery mechanisms could be enabled by online agri-platforms in resource constrained environments. Importantly, the success of these designs critically depends on systemic supply chain, process design and behavioral considerations that affect trades in the physical agri-markets.

1.3 Contribution of Chapter 4

The second stream of research in this thesis turns to consumer welfare and focuses on designing policies that tackle structural issues which arise in traditional agricultural markets. In particular, ensuring food safety and security in a highly-fragmented global food supply chain with monopolistic intermediaries remains a challenging, but an extremely important problem for governments across the world. Chapter 4 investigates food safety risks by developing a modeling framework to examine farms' strategic adulteration behavior and the resulting Economically Motivated Adulteration (EMA) risk in farming supply chains. We study both "preemptive EMA," where farms engage in adulteration to decrease the likelihood of producing low-quality output, and "reactive EMA," where adulteration is done to increase the perceived quality of the output. We fully characterize the farms' equilibrium adulteration behavior in both types of EMA and analyze how quality uncertainty, supply chain dispersion, traceability, and testing sensitivity (in detecting adulteration) jointly impact the equilibrium adulteration behavior. We determine when greater supply chain dispersion leads to a higher EMA risk and how this result depends on traceability and testing sensitivity. Furthermore, we caution that investing in quality without also enhancing testing capabilities may inadvertently increase EMA risk. Our results highlight the limitation of only relying on end product inspection to deter EMA. We leverage our analyses to offer tangible insights that can help companies and regulators to more proactively address EMA risk in food products.

1.4 Contribution of Chapter 5

Chapter 5 focuses on the challenge price surge of essential commodities despite inventory availability, due to *artificial shortage*, that presents a serious threat to food security in many countries. To protect consumers' welfare, governments intervene *reactively* with either (i) Cash Subsidy, to increase consumers' purchasing power by directly transferring cash, or (ii) Supply Allocation, to increase product availability by importing the commodity from foreign markets and selling it at subsidized rates. This chapter develops a new behavioral game-theoretic model to examine the supply chain and market dynamics that engender artificial shortage, as well as to analyze the effectiveness of various government interventions in improving consumer welfare. We analyze a three-stage dynamic game between the government and the trader. We fully characterize the market equilibrium and the resulting consumer welfare under the base scenario of no government intervention, as well as under each of the interventions being studied. The analysis demonstrates the disparate effects of different interventions on artificial shortage; while supply allocation schemes often mitigate shortage, cash subsidy can inadvertently aggravate shortage in the market. Further, empirical analysis with actual data on onion prices in India shows that the proposed model explains the data well and provides specific estimates on the implied artificial shortage. A counterfactual analysis quantifies the potential impacts of government interventions on market outcomes. The analysis shows that *reactive* government interventions with supply allocation schemes can have a *preemptive* effect to reduce the trader's incentive to create artificial shortage. While cash subsidy schemes have recently gained wide popularity in many countries, we caution governments to carefully consider the strategic response of different stakeholders in the supply chain when implementing cash subsidy schemes.

Chapter 2

The Impact of Unifying Agricultural Wholesale Markets on Market Prices and Farmers' Profitability

2.1 Introduction

Agriculture plays a significant role in the economies of most developing countries. As [Mondiale \(2008\)](#) notes, “Of the developing world’s 5.5 billion people, 3 billion live in rural areas, nearly half of humanity. Of these rural inhabitants, an estimated 2.5 billion are in households involved in agriculture, and 1.5 billion are in smallholder households.” Sadly, smallholder farmers in developing countries persistently struggle with poverty, in part due to unfavorable market outcomes for these farmers. Prior studies have examined various factors affecting farmers’ revenue in traditional markets of developing economies. These include imperfect competition ([Bergquist 2019](#)), provision of market information or the lack thereof ([Jensen 2007](#), [Parker et al. 2016](#), [Fafchamps and Minten 2012](#), [Aker 2010](#)), logistical infrastructure ([Casaburi et al. 2013](#)), and limited market access ([Goyal 2010](#)).

To address these challenges, one approach that has been attracting substantial investment is to connect geographically distributed agri-markets through a single online agri-platform. The hope is that integrating geographically distant markets within a common platform can increase market competition, enable transparency of the price discovery process, and ultimately, improve farmers’ profitability. For example, The World Bank has invested \$4.2 billion between 2003 and 2010 to develop infrastructure for information and communication technologies in the

developing world (IEG 2011). Various countries have launched online agri-platforms to transform traditional markets, with prominent examples such as the Ethiopia Commodity Exchange (ECX) and the eNational Agriculture Market (eNAM) in India. While it is postulated that building online agri-platforms can benefit farmers for the aforementioned reasons, rigorous empirical evidence is limited. Furthermore, qualitative evidence has suggested mixed results for some of the existing platforms (Ababa 2017, Nirmal 2017). To fill this gap, this chapter offers the first econometric analysis that evaluates the impact of launching such a platform – the Unified Market Platform (UMP) in Karnataka, India – on farmers' profitability.

UMP was established in 2014 by the state government of Karnataka to unify all transactions occurring in the state's regulated agricultural wholesale markets to be carried out within a single online platform. By November 2019, 162 out of the 164 regulated markets across 30 districts in the state have been integrated to UMP, and approximately 62.8 million metric tons of commodities valued at \$21.7 billion (USD) have been traded on the platform. Therefore, it is of significant value to empirically evaluate whether and how much the implementation of this state-wide agri-platform has impacted market prices and farmers' profitability. In addition, we utilize the analysis to shed light on systemic features that should be carefully considered in the design and implementation of other agri-platforms around the world.

2.2 Traditional Markets versus UMP

India's agricultural regulations in many states require that the trading of a predefined set of agricultural commodities be conducted in regulated agricultural markets called "mandis." The process starts with farmers bringing their commodities to a commission agent of their choice in their local mandi. Traders in the mandi visit commission agent shops, examine the quality of the commodities, and engage in auctions or direct negotiations to purchase the commodities of interest. Each trade is typically for a single "lot" from one farmer. Once the trade is finalized, commodities in the lot are weighed, and the winning trader pays the total price plus a commission to the agent, who later pays the farmer.

While the regulations were enacted to protect farmers' welfare, the resulting market structure and trade process have led to poor outcomes for farmers (Acharya 2004). First, lack of transportation and storage capabilities limit farmers' sales channels to the local mandis nearby. Second, since traders need to apply for a separate license for each mandi, there are typically a

small number of traders participating in each mandi. Furthermore, the traders' demand (from their customers) predominantly occurs during harvest seasons when supply quantity is high, and hence, traders engage in active trading mostly during those times. Third, the price-setting process in the mandis is done through hand-written tender slips and is not documented. It is subject to collusions among traders and also often involves private negotiations between the commission agents and the traders. Taken together, the restricted market access, their weak market power, and the nontransparent price discovery process all contribute to low sale prices and poor revenue for the farmers.

Realizing these challenges, the state government of Karnataka established the Rashtriya e Market Services Private Limited (ReMS) in 2014 and tasked this organization to integrate and digitize all mandis in the state through a single online platform – UMP. Under this reform, a number of changes were made to the traditional process (Aggarwal et al. 2017). First, the open outcry ascending auction is replaced by the online first-price sealed-bid auction. Traders must submit their (private) bids for all the lots they want to purchase on UMP by a preannounced cutoff time. Once the submission window is closed, all bids for the same lot are compared by the computer and the highest bidder is declared the winner. Second, all lots arriving to any of the integrated mandis are recorded on UMP and visible to all traders. Furthermore, the government enacted a single-license system so that traders need only one license to trade in all mandis within the state. Therefore, traders can now bid for lots that are put up for sale in other mandis. The process of generating transport permits and bills has also been digitized to facilitate the post-trade processes for traders. Finally, in order to increase price transparency for farmers, the government: (i) installed computer kiosks where farmers can check prices in major mandis across the state; (ii) started sending SMS messages to farmers informing them of the winning bid for their lots.

The government believes that these changes can benefit farmers through two main mechanisms – increased competition among the traders and improved transparency. The hope is that both in-market and cross-market competition would increase because (i) traders can participate in cross-market trading with a single license, and (ii) moving to sealed-bid auctions online and automating many of the post-trade processes increases traders' efficiency. In the meantime, price transparency benefits farmers because it helps to increase farmers' bargaining power compared to the traditional process. Nevertheless, the key question is, have these expected mechanisms been effective in improving prices for farmers?

2.3 Data and Empirical Approach

We employ a difference-in-differences (DID) approach for our analysis (Eq. [2.1]). The 158 markets in Karnataka, once integrated into UMP, are taken as treatment markets, and all markets outside of Karnataka are taken as control markets. The analysis focuses on six commodities for which (i) we have sufficient numbers of both treatment and control markets, and (ii) a common linear pre-trend of prices between the treatment and control markets prior to UMP's implementation cannot be rejected in the parallel trend test. These commodities are cotton, green gram, groundnut, maize, paddy, and tur. See Appendix A.1 for more details.

We utilize both public data from the Government of India and lot-level data from UMP in our analysis. There are three major data sources: (i) daily modal, maximum, minimum prices and supply quantity data for multiple commodities in all regulated mandis across India from 2012 to 2017 published by the Government of India; (ii) district and state level demographic, socioeconomic, and rainfall data from 2012 to 2017 published by government agencies; (iii) dates of UMP implementation across Karnataka mandis and lot-level auction data on UMP from 2016 to 2018. We aggregate the daily price and quantity data to the weekly level for the analysis because different markets are open on different days of the week. Data source (ii) is used to construct various covariates to control for potential differences across markets, including monthly rainfall from current to 6 months prior, yearly total production, yearly yield, all at the district level, and per capita GDP at the state level. We map each market to the associated district and state to match these covariates to the market level. The lot-level data from UMP is used to analyze systemic features that may differ across different commodities. See Appendix A.1 for more details.

2.3.1 DID model

We use a DID approach to estimate the impact of UMP on the prices of various commodities. Note that different markets in Karnataka were integrated into UMP at different, exogenously-determined dates (see Appendix A.3). All Karnataka markets, once integrated into UMP, are considered as treatment markets. All non-Karnataka markets are considered as control markets. Specifically, we estimate the following model for each commodity separately:

$$\log(P_{m,t}) = \gamma I_{m,t} + \delta \log(Q_{m,t}) + \Theta X_{m,t} + \alpha_m + \beta_t + \epsilon_{m,t}. \quad (2.1)$$

The dependent variable $\log(P_{m,t})$ is the logarithm of the modal price, maximum price, or minimum price observed in market m at week t . The key independent variable is the implementation dummy: $I_{m,t} = 1$ if market m has been integrated into UMP at week t and 0 otherwise. The variable $Q_{m,t}$ is the total quantity in market m at week t , and $X_{m,t}$ denote the vector of all control covariates discussed earlier. We control for market fixed effects (α_m) and week fixed effects (β_t). $\epsilon_{m,t}$ is the idiosyncratic error term. The coefficient of interest is γ . A positive and significant value of γ indicates that the implementation of UMP has led to a significant increase in the commodity's modal, maximum, or minimum price.

2.3.2 Robustness tests

We perform a number of robustness analyses to strengthen our results. In particular, we consider (i) five alternative model specifications; (ii) p value adjustments to account for multiple hypothesis testing; and (iii) three alternative specifications with different control covariates. We consider our main result to be robust if the direction and statistical significance of the coefficient for the implementation dummy are consistent between the main model and these robustness tests. We confirm that this is the case for all six commodities (see Appendix A.1).

2.3.3 DID assumption

Following Jensen (2007), we perform the following parallel trend test for each commodity using the data prior to the first market being integrated on UMP:

$$\log(P_{m,t}) = \gamma_1 t + \gamma_2 I_m^K t + \delta \log(Q_{m,t}) + \Theta X_{m,t} + \alpha_m + \epsilon_{m,t}. \quad (2.2)$$

The variable t denotes the number of weeks since the start of the data. The dummy variable $I_m^K = 1$ if market m is a Karnataka market and 0 otherwise. The remaining variables follow from Eq. [2.1]. A nonsignificant γ_2 indicates that we cannot reject the null hypothesis that prices in the treatment and control markets follow the same linear pre-trend (see Appendix A.4).

2.3.4 Falsification test

We follow Karplus et al. (2018) to perform a falsification test. In particular, we reestimate model [2.1] 1000 times with the data from the pre-treatment period and assuming randomly

selected placebo dates as the implementation dates each time. We confirm that the empirical distribution of the estimated effects from this falsification test is close to zero and significantly smaller than the estimated impact with the true implementation dates for paddy, groundnut, and maize, and it is not statistically significantly different from zero for cotton, tur, or green gram (see Appendix A.6).

2.4 The Impact of UMP

Figure 2.1 illustrates the estimated average impact of UMP on the modal prices of the six commodities, ordered by the magnitude of the impact from top to bottom. The statistics are summarized in the “Main model” column in Table A.4. We observe that the implementation of UMP has yielded statistically significant, positive impacts on the modal prices of paddy, groundnut, and maize – a 5.1%, 3.6%, and 3.5% increase. Given low profit margins for smallholder farmers (2%-9%), these price increases imply 36%-159% improvement in the profit margins for over 2 million farmers who traded on UMP (Vissa 2017).¹ However, the analysis also shows that the implementation of UMP has *not* generated statistically significant impacts on the modal prices of cotton, tur, or green gram. These results remain valid in a number of robustness analyses (see Appendix A.5 for more details).

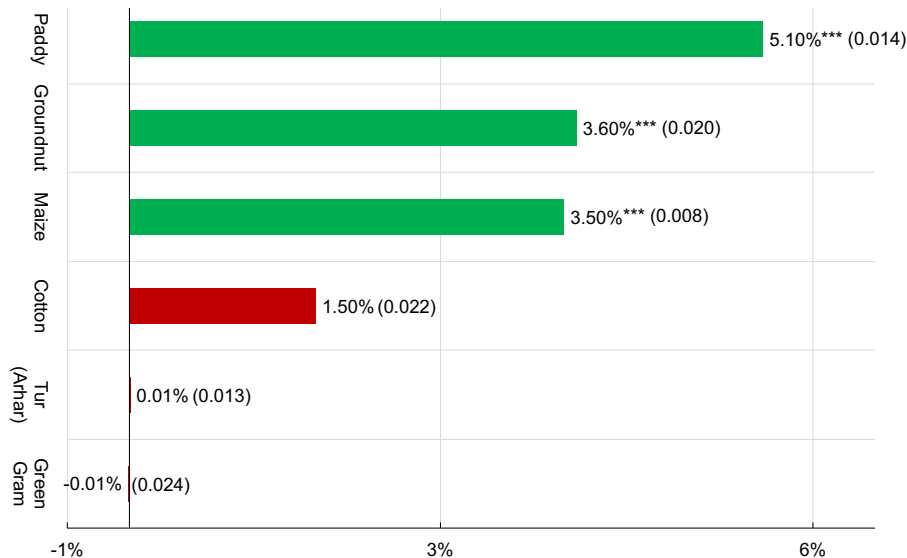


Figure 2.1: Estimated average impact (in %) of UMP on the modal prices of different commodities from the Main model (Eq. [2.1]). Green (red) indicates a statistically significant (nonsignificant) impact. Numbers next to the bars are the estimates with standard errors in parentheses.

¹The assumption underlying this derivation is that the farmers’ costs and production quantities have not changed significantly due to UMP.

Commodity	Main model	Maximum price	Minimum price	Fertilizer usage	
				γ_0	γ_H
Paddy	0.051*** (0.013)	0.054*** (0.013)	0.049*** (0.011)	0.015 (0.012)	0.063*** (0.022)
Observations	131,903	129,504	129,532	124098	–
Groundnut	0.036*** (0.014)	0.048*** (0.010)	-0.028 (0.031)	0.013 (0.016)	0.054** (0.021)
Observations	34,944	34,837	34,848	33,542	
Maize	0.035*** (0.007)	0.024*** (0.006)	0.046*** (0.010)	0.039*** (0.012)	-0.009 (0.013)
Observations	89,478	87,106	87,105	87,685	
Cotton	0.015 (0.012)	0.033*** (0.010)	-0.048 (0.045)	0.005 (0.015)	0.020 (0.023)
Observations	46,055	46,051	46,053	40,813	
Tur	0.008 (0.016)	0.004 (0.015)	0.017 (0.020)	0.012 (0.021)	-0.007 (0.031)
Observations	52,953	50,377	50,359	51,283	
Green Gram	-0.009 (0.013)	-0.006 (0.013)	0.007 (0.019)	-0.013 (0.014)	0.011 (0.026)
Observations	51,742	50,822	50,836	47,243	

Table 2.1: Estimated Impact of the UMP by Commodity

Notes. For all but the “Fertilizer usage” model, we report the γ estimates for the implementation dummy ($I_{m,t}$) in Eq. [2.1]. For the “Fertilizer usage” model, we report the estimates of γ_0 and γ_H in Eq. [A.1]. Standard errors (in parentheses) are clustered at the market level. In all models, we control for district-level yearly production, yearly yield, monthly rainfall, and state-level per capita GDP, as well as market and week fixed effects.

2.4.1 Heterogeneous Impacts by Product Quality

One natural question is whether farmers who produce higher-quality products benefit more from an online agri-platform such as the UMP.² To investigate this question, we perform two additional analyses. First, we examine the impact of UMP on the maximum price and minimum price of the six commodities. Compared to the results on modal prices, we observe an additional impact of UMP on the maximum price of cotton but do not observe a significant impact on the minimum price of groundnut (Table A.4 columns 2–3).

Second, we use farmers’ fertilizer usage per unit of farmland as a proxy for the respective product quality and analyze whether the impact of UMP differs for farmers with high vs. low usage.³ To do so, we first identify Karnataka markets located in districts with fertilizer usage

²Since we do not have data directly on product quality, we rely on multiple sources to understand for which commodities prices are likely to be sensitive to quality. These include our field interactions with the traders and domain experts, academic references, and analyzing within-day price variation across lots in the UMP data.

³The presumption that higher fertilizer usage indicates better overall quality of the commodities grown in the district is based on evidence in the agricultural sciences (Yousaf et al. 2017), and it is reasonable in the setting

above the median usage across all districts in the state. We then estimate the impact of UMP on the modal prices in these markets as compared to the remaining markets (Eq. [A.1] in *Materials and Methods*).

In particular, we estimate the following model for each commodity:

$$\begin{aligned} \log(P_{m,t}) = & \gamma_0 I_{m,t} + \gamma_H I_{m,t} \times F_m^H + \delta \log(Q_{m,t}) + \Theta X_{m,t} \\ & + \alpha_m + \beta_t + \epsilon_{m,t}. \end{aligned} \tag{2.3}$$

The additional variable F_m^H is an indicator variable for high fertilizer usage. That is, $F_m^H = 1$ if the fertilizer usage per unit of farmland in the district where market m is located is above the median of the distribution of fertilizer usage across all districts in the state, and $F_m^H = 0$ otherwise. A positive and statistically significant value of γ_H indicates that farmers in districts with above-the-median fertilizer usage benefit from a larger price increase due to UMP compared to the rest of the farmers.

The estimated differential impact is presented in Table A.4 column γ_H . We observe a statistically significant differential impact for paddy and groundnut. Specifically, paddy and groundnut farmers with above-the-median fertilizer usage gain a greater price increase than those with below-the-median fertilizer usage. Note however that we do not find differential impact with respect to fertilizer usage for maize, despite a significant overall price increase for this commodity.

Our field interactions with the traders and domain experts, as well as some documented evidence ([Ethridge and Davis 1982](#), [Nautiyal 2002](#)), indicate that paddy, groundnut, and cotton are commodities whose prices are sensitive to quality. In contrast, price for maize is less sensitive to its quality as maize is predominantly used as cattle feed in the country ([Singh 2014](#)). We find some evidence in line with these expectations in the UMP data. Specifically, we should expect to see more price variation across lots on a given day for commodities whose price is more sensitive to quality. Indeed, the coefficient of variation (CV) of the prices across lots within any given day and market is substantially higher for paddy, groundnut, and cotton than for maize. The average CV of prices across all days and markets is 0.141, 0.135, and 0.118 for paddy, groundnut, and cotton, versus 0.043 for maize. A regression analysis confirms that the CV of prices for the former three commodities is significantly higher than that of maize. Taken

of resource-constrained smallholder farmers where overuse of fertilizers is unlikely.

together, the above analyses suggest that the implementation of UMP may have generated greater benefits for farmers who produce higher-quality products. See Appendix A.7 for further details.

The disparate impacts of UMP across different commodities motivate us to investigate what systemic factors potentially contribute to these differences. This investigation is informed by our extensive field visits and interviews with farmers, commission agents, traders, and mandi officials in Karnataka. Hereafter, we refer to paddy, groundnut, and maize as the “high-impact” group (i.e., commodities for which UMP has yielded a statistically significant, positive impact), and we refer to cotton, tur, and green gram as the “low-impact” group (i.e., commodities for which UMP has not generated a significant impact). The analysis suggests four factors that distinguish the high-impact group from the low-impact group: logistical challenges for cross-market trading, the increased efficiency of bidding under UMP, the level of in-market concentration for a commodity, and the price discovery process used. The first three factors relate to how the implementation of UMP may have increased competition among traders (or not), and the last factor relates to increased bargaining power for farmers due to price transparency. We observe mild correlations among these factors, suggesting that they each capture some distinctive aspect of differences between the high-impact and low-impact groups (see Appendix A.8).

2.5 Systemic Differences between High-impact and Low-impact Groups

2.5.1 Cross-Market Competition

Transportation and logistics costs are believed to be significant barriers to effectively integrating agricultural markets in developing countries (Casaburi et al. 2013). Our field visits to the mandis reveal that logistical infrastructure is still lacking to support cross-market trading. Specifically, traders are fully responsible for processing and transporting all the lots they have won within the same day, regardless of where the lots are located. Traders need to hire additional laborers, representatives, and transporters to handle these tasks, adding costs and coordination challenges across mandis. Thus, we postulate that cross-market trades would be more likely to occur for commodities whose major markets are closely located from each other, as compared to commodities whose major markets are geographically spread out. Because increasing market competition – in part by enabling cross-market trades – is conceived as a key mechanism to

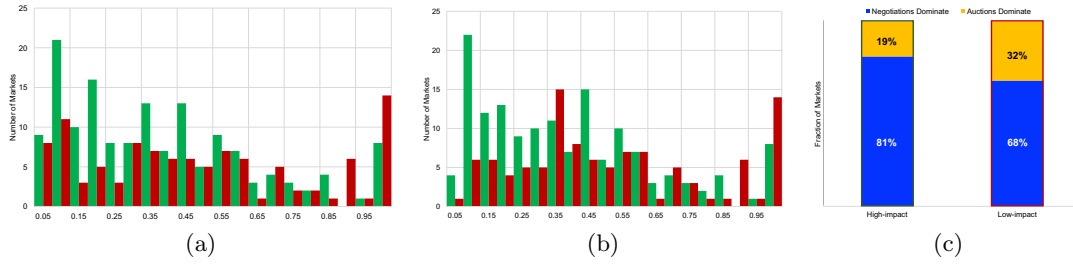


Figure 2.2: Fig. 2.2a ($n = 258$) and Fig. 2.2b ($n = 258$) present the distribution of the average HHI-a and average HHI for the high-impact group (green) and the low-impact group (red) across markets. Fig. 2.2c ($n = 461$) presents the fraction of markets in which negotiations (blue) or auctions (yellow) dominate. Left bar considers markets for the high-impact group; right bar considers markets for the low-impact group.

improve prices for farmers, we make the following hypothesis.

- H1.** The major markets for commodities in the high-impact group are more closely located than those for the low-impact group.

Evidence

To test H1, we compare the pairwise distances among the major markets by quantity between commodities in the high-impact group versus those in the low-impact group. The set of major markets for a commodity are the largest markets where 90% of all quantities are traded. We focus on major markets by quantity because cross-market trading, if any, most likely occurs among markets with large quantities. A t test confirms that the major markets for commodities in the high-impact group (paddy, groundnut, and maize) are indeed closer to each other ($t = -5.495$, $p < 0.0001$). The policy recommendation from this analysis is that, the benefits of online agri-platforms can be enhanced by building logistical infrastructure conducive for cross-market trades. Possible options include providing cross-market traders with services of post-trade sorting and processing, arranging third-party logistics providers (with long-term contracts) to facilitate transportation, and other similar pooling services that can leverage economies of scale, and hence, lower the logistics costs associated with cross-market trading for the traders.

2.5.2 In-Market Competition

Our field interviews with traders reveal that a key benefit from UMP to them is the increased efficiency in trading and bidding due to the online platform. In traditional mandis, traders need to visit each commission agent shop to submit their bids or negotiate in a sequential manner. For commodities whose supply is distributed across many commission agents, this process is

very time-consuming. As a result, traders often can only visit a limited number of agents within the day, which in turn limits the number of bids each lot receives. In contrast, under UMP, traders can quickly visit many agents to inspect the lots and privately record their bids on a tender slip before submitting all bids online all at once. This increased efficiency could lead to an increase in the number of bids per lot (and hence, the winning price) under UMP. Note that this effect would be stronger for commodities whose supply is more distributed across agents, and hence, the efficiency gain would be more prominent. We summarize the above logic in the following hypothesis.

H2. Commodities in the high-impact group have more distributed supply among commission agents than those in the low-impact group.

Evidence

We use the Herfindahl-Hirschman Index (HHI) to quantify the level of supply concentration among commission agents for different commodities, denoted as HHI-a (Rhoades 1993). In particular, the HHI-a for a commodity in a given market is defined as the sum of squares of each commission agent's supply share of that commodity in the market, where an agent's supply share is the fraction of total market quantity sold at the agent's shop. A larger value of HHI-a implies a higher level of supply concentration among the agents. Figure 2.2a presents the distribution of the average HHI-a for the high-impact group (green) and the low-impact group (red) across all markets trading these commodities. We confirm with a t test that commodities in the high-impact group (paddy, groundnut, and maize) are associated with a lower HHI-a (i.e., a larger number of agents splitting the total quantity), supporting H2 ($t = -2.981, p = 0.003$).⁴

We also utilize the detailed lot-level auction data from the UMP to examine whether the extent of competition in the markets has changed post the implementation. We observe that the average number of bids received by each lot has significantly increased for paddy, groundnut, maize, and cotton; however, it has decreased for tur and green gram (see Appendix A.9).⁵ These results are consistent with the expectation that a larger increase in the level of competition among traders can result in a stronger benefit from UMP. The policy implication is that, the benefits of online agri-platforms can be enhanced further by reducing search costs for traders

⁴This result is also confirmed by a regression analysis that clusters standard errors at the market level. Similar analyses are done for H3 and H4. See Appendix A.8 for more details.

⁵We do observe a limited extent of cross-market trading (2%-12% of the traders across commodities) that may correlate with our results on the average number of bids per lot. See Appendix A.2 for more details.

in the markets. For example, pooling lots of similar quality and/or quantity can allow traders to identify prospective lots more efficiently both within and across markets.

2.5.3 In-Market Concentration

A key drawback in traditional mandis that has been extensively reported is the presence of collusion and price manipulation by traders (Aggarwal et al. 2017). Due to the traditionally manual and undocumented auction process, it was easy for traders to collude among themselves or with commission agents and change bid prices in the tender slips. It is easier for such collusion and price manipulation to occur when market power is concentrated on a small number of traders. One intended benefit of UMP is to deter these collusive behaviors by digitally recording bids, automating winner determination, and disseminating winner information to farmers. Given this logic, we postulate two opposing effects from in-market concentration. First, if market concentration is low, then the extent of collusion prior to UMP's implementation may already be low, and hence, UMP would have a small impact for commodities with low market concentration. Conversely, if market power is highly concentrated among a few dominant traders, then there remain risks of these traders forming bidding rings to manipulate prices even under UMP. As a result, UMP would have a limited impact for commodities with high market concentration. We thus propose two competing hypotheses to be examined with the data.

H3a. There is higher concentration of in-market power among traders for commodities in the high-impact group than for those in the low-impact group.

H3b. There is lower concentration of in-market power among traders for commodities in the high-impact group than for those in the low-impact group.

Evidence

We again use the HHI to quantify the level of market concentration for different commodities. Specifically, the HHI for a commodity in a given market is defined as the sum of squares of each trader's market share of that commodity in the market. In our context, a trader's market share corresponds to the fraction of total market quantity bought by the trader. A larger value of HHI implies a higher level of market concentration. Figure 2.2b presents the distribution of the average HHI for the high-impact group (green) and the low-impact group (red) across all markets trading these commodities. A t test shows that commodities in the high-impact group (paddy, groundnut, and maize) are associated with a *lower* HHI (i.e., a larger number of traders

splitting the total quantity; $t = -3.678$, $p = 0.0003$). This result supports H3b instead of H3a. The policy implication is that, the benefits of online agri-platforms can be enhanced further by reducing entry barriers to attract new traders, or by adopting alternative price discovery mechanisms to intensify within-market competition.

2.5.4 Price Discovery Process

The lack of price transparency in the traditional system wherein farmers had to rely on traders for price information was a key issue negatively affecting market outcomes for farmers. In order to increase transparency, the government has installed computer kiosks with price information and also started sending price information through SMS messages to farmers. By allowing farmers to access price information in real-time, farmers' bargaining power vis-a-vis traders is expected to increase. The effect of such increased bargaining power on final prices would depend on the price discovery process used. With direct negotiation, farmers could potentially use the price information to their advantage and bargain with traders more proactively. In contrast, the scope of bargaining is limited when prices are determined via first-price sealed-bid auctions (without reserve prices). Therefore, we posit that for commodities in the high-impact group, direct negotiation may be used more prevalently than auctions, leading to the following hypothesis.

H4. A larger fraction of quantity is traded through direct negotiations as opposed to auctions for commodities in the high-impact group than for those in the low-impact group.

Evidence

Figure 2.2c presents the proportion of markets in which a larger fraction of quantity is traded through negotiations (blue) versus auctions (yellow), separating the high-impact group (left bar) and the low-impact group (right bar). We observe that negotiations are indeed used more frequently for commodities in the high-impact group (paddy, groundnut, and maize). A t test that compares the fraction of quantity traded through negotiations between the high-impact and low-impact groups confirm this observation, supporting H4 ($t = 3.822$, $p = 0.0002$). The policy implication from this result is that, online agri-platforms using auctions would benefit from additional means to increase the bargaining power of farmers. For instance, farmers can be recommended reserve prices for their lots based on prevailing market rates.

2.6 Conclusions and Discussion

While online agri-platforms provide the desirable infrastructure to enable potential integration of distant agri-markets, the results from this chapter highlight that their success critically depends on systemic supply chain logistics and process design considerations that affect trades in the physical markets. For example, providing integrated logistics can encourage and facilitate cross-market trading; pooling lots of similar quality can further reduce traders' search costs and enhance bidding efficiency; optimizing the auction design can further increase market competition and strengthen the farmers' bargaining power. These practical insights are relevant and applicable to the design and optimization of other similar agri-platforms beyond the case of UMP. In addition, the results also show that an integrated agri-platform such as the UMP generates greater benefits for farmers who produce high-quality products.

Three limitations in the current study are worth discussing. First, traders do not necessarily submit the cross-market bids themselves, but often use "delegates" in other markets to bid on their behalves. Since cross-market bidding through "delegates" is not identifiable in the UMP data, we can only provide a lower bound on cross-market trading. Despite this limitation, we find that directional results on cross-market trading are in line with our hypothesis on logistical challenges (see Appendix A.2). Second, to evaluate H2–H4, we use the UMP data to measure potential differences in market structure across commodities, assuming that these structural features remain similar before and after UMP's implementation. We have to do so because detailed market structure data are not available for Karnataka markets prior to UMP's implementation. In Appendix A.9, we discuss mild changes in some of the market features since UMP's implementation. Similarly, data on these market features are not available for the control markets. Hence, we cannot directly analyze the moderating effects of these features on the impact of UMP. Third, we do not have data on product quality and thus, rely on the analyses on maximum/minimum price and fertilizer usage to shed light on how the benefit of UMP may vary by product quality. As the government continues to scale their quality assaying efforts, future studies can utilize the resulting data to more directly investigate the role of product quality in affecting the realized benefit of an integrated agri-platform such as the UMP.

Chapter 3

Improving Farmers' Income on Online Agri-platforms: Theory and Field Implementation of a Two-stage Auction

3.1 Introduction

The phenomenon of severe poverty continues to persist among smallholder farmers in developing countries, in part due to unfavorable market outcomes for these farmers ([Mondiale 2008](#)). Prior studies suggest that imperfect competition and inefficient price discovery processes in traditional agricultural markets are important factors affecting farmers' income ([Bergquist 2019](#), [Goyal 2010](#)). To tackle these challenges, one prevalent intervention that has been attracting substantial investment is to connect geographically isolated markets via an online platform. In such a platform, various aspects of the price discovery process are digitized and automated, including winner determination and declaration in auctions, as well as dissemination of price information. The hope is that these platforms could increase market competition, enable transparency of the price discovery process, and ultimately, improve farmers' income.

Multiple countries have launched such online agri-platforms, for example, the commodity exchange platforms in Ethiopia, Kenya, Nairobi, and Uganda, as well as the eNational Agriculture Market (eNAM) launched by the central government of India, and the Unified Market Platform (UMP) implemented in the state of Karnataka, India. The price discovery mechanisms

used in these platforms vary widely from ascending auctions, first-price sealed-bid auctions, to warehouse-based negotiation, mostly following a digital version of the traditional (pre-platform) mechanisms. It remains an open question as to which format may be most beneficial for farmers (Strzebicki 2015, Deichmann et al. 2016). This is further highlighted by the mixed empirical evidence regarding the impacts of these platforms on farmers' income. For example, trades on the Ethiopia Commodity Exchange are dominated by two export crops, coffee and sesame seeds, while farmers of other crops are largely left out (Ababa 2017, Hernandez et al. 2017). In India, only 14% of all farmers in India have registered on eNAM, and over half of these registered farmers are reported to not have benefited from the platform (Business Line 2019).

This chapter describes work done in close collaboration with the state government of Karnataka to increase farmers' revenue by redesigning the price discovery mechanism on the UMP. UMP was created in 2014 by the state government to unify all trades in agricultural wholesale markets in Karnataka to be carried out within a single platform. By November 2019, approximately 62.8 million metric tons of commodities valued at \$21.7 billion (USD) had been traded across 162 markets on the UMP. The major types of commodities traded on the UMP include grains, lentils, oil seeds, and cash crops (e.g., cotton, areca nut). Trades occur on a daily basis through first-price, sealed-bid auctions on the UMP.

While the well-known revenue equivalence theorem (Krishna 2009) implies that the first-price sealed-bid auction would achieve equivalent expected revenue as many other theoretically appealing auction mechanisms (including the second-price sealed-bid auction), the theorem assumes that bidders have a quasi-linear utility function that is unaffected by the information feedback in the auction. In contrast, a large body of empirical evidence shows that bidders are often subject to behavioral regularities induced by information cues. One widely-observed regularity is *anticipated regret*. The effect of anticipated regret on bidding behavior has been observed extensively in laboratory experiments of single-stage auctions (Filiz-Ozbay and Ozbay 2007, Engelbrecht-Wiggans and Katok 2007, 2008). Bidders are shown to bid higher prices if they know that the winning bid will be disclosed to them when they lose (versus no disclosure). The hypothesized reason behind this phenomenon is that bidders may anticipate feeling loser's regret ex post (i.e., they could have won and made a profit), and such anticipation affects their bidding strategies ex ante. These empirical findings motivate us to investigate how the auction mechanism used on the UMP can be redesigned to *intensify* anticipated regret of the traders and increase revenue for the farmers.

We postulate that two-stage auctions – a format commonly used in real estate markets and for mergers and acquisitions, albeit due to entirely different motivation– can intensify *anticipated regret* and thereby increase farmers' revenue. A two-stage auction works as follows. In the first stage, all traders bid as in the single-stage, first-price, sealed-bid auction. At the end of the first stage, the top k traders (based on their first-stage bids) qualify to participate in the second-stage. Some information related to the first-stage bids (e.g., the highest first-stage bid) is disclosed to the qualified traders who are given an opportunity to increase their bids in the second stage. Finally, the top bidder at the end of the second stage is declared as the winner. A two-stage auction design is expected to intensify anticipated regret of the traders for two reasons. First, traders are exposed to *repetitive regret* (Coricelli et al. 2005, Roese et al. 2009) as information feedback of both not qualifying for the second stage and losing the auction is made salient. Second, anticipated regret is less intense for low-probability events (e.g., winning in a single-stage, first-price, sealed bid auction) compared to events of higher probability (e.g., qualifying for the second stage in a two-stage auction; Robinson and Botzen 2019).¹

Since changing the auction mechanism on the UMP can have a profound impact on the farmers' revenue, we engage in a careful design process to examine when implementing the two-stage auction mechanism will indeed benefit the farmers. To do so, we first develop hypotheses on the traders' bidding behavior in a two-stage auction based on existing theories in the literature. These hypotheses guide the design of semi-structured field interviews with a group of professional traders. The interviews provide evidence that the traders are subject to anticipated regret and an additional behavioral factor – anchoring. Motivated by these field insights, a behavioral auction model is developed to analyze when a two-stage auction can yield a higher revenue for the farmer than a single-stage, first-price, sealed-bid auction. The analysis demonstrates that the two-stage auction would most likely benefit farmers for commodities of constrained supply or high economic values. As a result, tur (pigeon pea) is chosen as the commodity to implement the two-stage auction due to its limited supply as compared to its consistently high consumption across India.

Working with the state government of Karnataka, the two-stage auction was implemented on the UMP in February 2019 for a major market of tur. By the end of May 2019, more than

¹We do not consider other common auction mechanisms, such as second-price sealed-bid or English auction, since prior research has shown that anticipated regret plays little role in affecting bidders' bidding behaviors under these mechanisms (Filiz-Ozbay and Ozbay 2007). There may exist other forms of auctions that can potentially intensify anticipated loser's regret of the traders, e.g., Dutch auction. The focus on the two-stage auction is driven by considerations of operational constraints in the markets. See Appendix §B.1 for further discussions.

8,000 metric tons of tur worth over \$6 million (USD) had been traded under the new auction in the treatment market. To evaluate the impact of the implementation on farmers' revenue, a difference-in-differences analysis is performed utilizing lot-level bidding data for all tur trades in the treatment market and a comparable control market (a total of 172,342 bids). The empirical analysis demonstrates that the implementation has yielded a statistically significant, 4.7% average price increase in the treatment market. The price increase translates into an average revenue gain of \$452,000 (USD) for over 10,000 farmers in a matter of three months, representing an average 15% increase in their monthly income. Given low profit margins for these farmers (4%–7%), the range of profit improvement is substantial (60%–158%). These positive outcomes are particularly significant and encouraging given the lack of price gain for tur farmers from UMP's original implementation in 2014 identified in Chapter 2.

Finally, the bidding and field interview data allows us to validate the relevance of anticipated regret and anchoring captured in the behavioral auction model. First, with respect to anticipated regret, we observe a 2.2 percentage point additional increase (or a 73% increase) in the bids from high-regret traders versus low-regret ones, as compared to bids in the control market. Second, with respect to anchoring, the bid increments observed in the second stage throughout the implementation never exceeds 0.75% of the highest first-stage bid, with an average of 0.2%.

3.1.1 Related Literature and Contributions

This chapter addresses the crucial topic of improving smallholder farmers' welfare and makes important contributions to both research and practice. Our work is related to two streams of operations management literature. The first stream is a growing body of research focusing on improving farmer and social welfare in agricultural supply chains (see [Bouchery et al. 2017](#), [Kalkanci et al. 2018](#), for broader reviews). While a majority of these studies focus on the production stage (e.g., [Dawande et al. 2013](#), [Alizamir et al. 2019](#), [Boyabati et al. 2019](#), [Federgruen et al. 2019](#), [Hu et al. 2019](#)), a few recent papers ([Parker et al. 2016](#), [Liao et al. 2019](#)) examine the market side of agricultural supply chains. This chapter differs from the literature by examining auction design in the new context of online agri-platforms. The chapter demonstrates the importance of adopting a behavior-centric, field-based, data-driven approach to *operationalize* and *improve* the design of online agri-platforms. The approach combines field interviews, behavioral game-theoretic modeling, and econometric analysis. Working closely with the local government, the analysis informs the design and implementation of a large-scale field pilot of a new two-stage

auction. The field implementation provides strong empirical evidence that effective auction design that carefully accounts for the salient behavioral and operational factors in the physical markets can yield significant benefits to smallholder farmers trading on the platform.

The second stream of related literature studies the effects of bidders' behavioral factors on their bidding strategies (see [Elmaghraby and Katok 2018](#), for a recent review). Most of the research here uses laboratory experiments to examine various behavioral factors such as regret, anchoring, risk aversion, learning, and competitive arousal. This chapter is most closely related to the papers that study the effects of regret and anchoring on bidding behavior (e.g., [Ariely and Simonson 2003](#), [Ku et al. 2006](#), [Filiz-Ozbay and Ozbay 2007](#), [Engelbrecht-Wiggans and Katok 2008](#)). We add to this literature in two aspects. First, we develop a new behavioral auction model that *jointly* captures the effects of anticipated regret and anchoring on bidding behavior in a two-stage auction. The modeling analysis provides prescriptive insights into *when* a two-stage auction can generate higher revenue than a single-stage first-price sealed-bid auction in practice. These insights are generalizable to other markets and products beyond the specific ones considered in this chapter. Second, while prior works mainly rely on laboratory experiments as the empirical basis, the development of the proposed model is informed by semi-structured field interviews with professional traders, and the empirical validation of the model is based on actual auction data from a large-scale field implementation.

The remainder of the chapter is organized as follows. [§3.2](#) presents the operational backgrounds of the UMP and introduces the two-stage auction. [§3.3](#) examines the traders' bidding behavior in a two-stage auction based on existing theories in the literature and theory-informed, semi-structured interviews in the field. [§3.4](#) develops and analyzes a behavioral auction model motivated by field insights. [§3.5](#) describes the implementation of the two-stage auction in a major lentils market, empirically evaluates its impact on farmers' revenue, and validates the relevance of the behavioral factors with field data. [§3.6](#) concludes the chapter.

3.2 Operational Context and Design of the Two-stage Auction

To fully appreciate the operational context, it is important to understand the current trading process on the UMP. Trades happen every day in the physical markets. At the start of a day, farmers bring their commodities to a commission agent of their choice in a local market. All lots arriving to any of the markets are posted on the UMP and are visible to all traders. Throughout

the day, traders or their representatives visit commission agent shops to identify lots they want to purchase and the prices to bid. Afterwards, they must submit their (private) bids for all the lots they want to purchase on the UMP before a pre-announced cutoff time (currently 2:30 p.m.). Once the bidding window is closed, the highest bidder for each lot is declared as the winner. The winner pays the farmer the highest bid, and the winning bid is announced publicly. That is, the current auction design on the UMP is a single-stage, first-price, sealed-bid auction at an individual lot level.

Prior research has shown in laboratory experiments that the disclosure of winning bids in single-stage auctions can engender anticipated regret and result in higher bids by the bidders (Filiz-Ozbay and Ozbay 2007, Engelbrecht-Wiggans and Katok 2008). Specifically, when bidders know that the winning bid will be disclosed (versus no disclosure), they may anticipate feeling loser's regret ex post (i.e., they could have won and made a profit). Such anticipation can motivate them to bid more aggressively ex ante. Building on these behavioral insights, we consider a two-stage auction design which is postulated to intensify anticipated regret of the traders. Figure 3.1 presents the timeline and rules of the proposed two-stage auction mechanism. Specifically, in the first stage, all traders bid as in the current first-price sealed-bid auction. At the end of the first stage, the current highest bid of each lot is disclosed to the top k traders (based on their first-stage bids) who have bid on the lot. These traders are then offered an opportunity to increase their bids over the highest first-stage bid in the second stage, which must be completed within a much shorter time window than the first stage. The top bidder at the end of the second stage is declared as the winner.² In this design, the first stage serves as a qualification stage. Only the subset of qualified traders can bid in the second stage, and if they do, they are required to match or outbid the highest first-stage bid. Hence, the second stage is essentially another first-price sealed-bid auction with a reserve price.

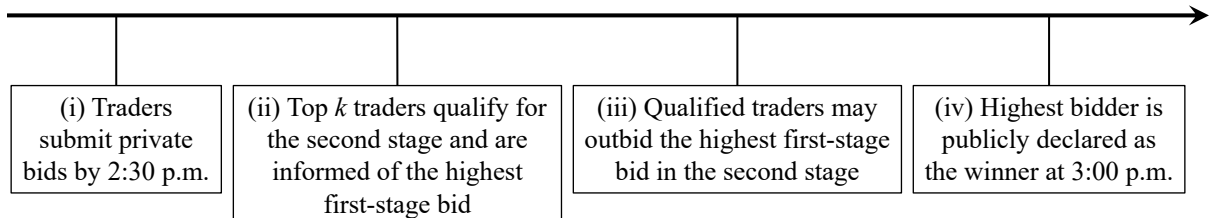


Figure 3.1: Rules Under the Two-Stage Auction

²If a lot receives fewer than k bids in the first stage or if no one increases their bids in the second stage, then the top bidder in the first stage is declared as the winner of the lot.

3.3 Traders' Bidding Behavior

Changing the price discovery mechanism on the UMP can have a profound impact on farmers' revenue. Since two-stage auctions have never been used in agricultural markets, we must carefully evaluate its viability to enhancing farmers' revenue. Therefore, the first step in the design process is to gain a good understanding of the traders' bidding behavior in the two-stage auction. This analysis contains three parts: (i) investigate relevant theoretical predictions of the traders' bidding behavior based on extant literature, (ii) conduct semi-structured interviews to test whether these theoretical predictions apply to professional traders in the field, and (iii) develop and analyze a behaviorally-justifiable auction model based on the field insights. The model analysis confirms the revenue benefit of implementing the two-stage auction (as compared to the single-stage, first-price, sealed-bid auction) for the commodity of tur. Furthermore, it determines more generally the conditions under which a two-stage auction will likely benefit farmers.

3.3.1 Theoretical Predictions

We identify from the literature four relevant theories that may influence bidding behavior in two-stage auctions: anticipated regret, competitive arousal, anchoring, and costly information acquisition.³

Anticipated (non-qualification) regret.

The effect of anticipated regret on bidding behavior has been observed extensively in laboratory experiments of single-stage auctions (Filiz-Ozbay and Ozbay 2007, Engelbrecht-Wiggans and Katok 2008). Bidders are shown to bid higher prices if they know that the winning bid will be disclosed to them when they lose (versus no disclosure). This is because bidders anticipate feeling regret ex post that they could have won and made a profit. To avoid such regretful emotions, bidders may bid more aggressively ex ante. Based on evidence in clinical psychology (Coricelli et al. 2005, Roese et al. 2009), we expect a priori that a two-stage auction would intensify such regretful emotions anticipated by the traders due to non-qualification. Specifically, traders are exposed to *repetitive regret* because information feedback of *both* not qualifying for the second

³Another commonly discussed factor in the literature is risk aversion. We do not test this factor because (i) there is no a priori expectation that risk aversion would lead to more or less aggressive bidding in two-stage vs. single-stage auctions; and (ii) we had to control the length of the interview to appreciate the traders' time constraints.

stage *and* losing the auction is made salient. In addition, anticipated regret is less intense for events that are of low probability compared to events that are of higher probability (Robinson and Botzen 2019). For any given bid, *winning* in a single-stage first-price sealed-bid auction has a lower probability in comparison to *qualifying* for the second stage in a two-stage auction. As a result, non-qualification regret in a two-stage auction is likely to be more intense than the loser's regret in a first-price sealed-bid auction. Anticipating this intensified non-qualification regret may motivate the traders to further increase their first-stage bids, thereby benefiting the farmers.

Competitive arousal versus anchoring.

According to the competitive arousal theory, intensifying the competitive situation (e.g., heightening rivalry and time pressure) can motivate bidders to overbid so as to win the competition, even if doing so is costly (Malhotra 2010, Adam et al. 2015). The a priori expectation from this theory is that introducing a second stage, in which there are only a few rivals and traders need to bid in a short time frame, could intensify the traders' feeling about the competitive situation. As a result, qualified traders will likely bid aggressively in the second stage to ensure that they win the qualified lots, thus improving the farmers' revenue. In contrast, the theory of anchoring suggests instead that the highest first-stage bid disclosed to the qualified traders would act as a salient anchor for their second-stage bids (Holst et al. 2015). This anchoring behavior could limit the extent that these traders increase their bids in the second stage, thus restricting the potential revenue benefit of a two-stage auction.

Costly information acquisition – quality uncertainty.

Multi-stage auctions are shown to benefit the auctioneer in situations when the valuation of the asset is highly uncertain (Ye 2007) or when bidders have interdependent values (Milgrom and Weber 1982). In these situations, potentially interested buyers can gain additional information about the asset's value after the first-stage bidding. This in turn allows the second-stage bids to better capture the true value of the asset, and hence, increase the auctioneer's revenue (Ye 2007, Quint and Hendricks 2013, Golrezaei and Nazerzadeh 2016, Milgrom and Weber 1982). In our context, one aspect of information that substantially affects traders' valuations of a lot is the potentially uncertain quality of the lot. When traders have to bid on hundreds of lots within a limited time frame, acquiring accurate quality information can be costly. Furthermore,

the unknown objective quality of a lot may lead to interdependent valuations among traders. Thus, adding the qualification stage can allow the traders to obtain more accurate signals about the quality of the lots they qualify to bid on in the second stage, e.g., by reinspecting the lots or inferring quality from the highest first-stage bid.

3.3.2 Field Interviews

Guided by the theories discussed in §3.3.1, we design and conduct a series of semi-structured interviews with a group of tur traders in a major lentils market to test whether the theoretical predictions indeed apply to professional traders.⁴ The design of the interview follows well-established principles in the social sciences (Lune and Berg 2016, , see Appendix B.4). In total, we interviewed 13 traders (out of 19 major traders in the market). These traders have on average 16 years of experience in agricultural trading and have won over 68% of all the lots (over 69% of the total quantity) traded in the market between January 2018 and February 2019. Therefore, the sample represents a majority of the traders whose bids substantially influence the final market prices. Each interview took about an hour to complete.

Anticipated non-qualification regret.

To examine whether traders are influenced by non-qualification regret as hypothesized in §3.3.1, the traders were presented a hypothetical scenario of a two-stage auction and asked to bid in the first stage during the interviews. Next, they were asked to imagine that they were not qualified for the second stage and to rate on a 9-point scale the intensity of different emotions (anger, envy, irritation, regret, and sadness) that they would feel given this outcome. A higher score indicates a stronger feeling of the corresponding emotion, with 1 meaning no feeling at all and 9 meaning an extremely strong feeling. The design of these questions follows from Filiz-Ozbay and Ozbay (2007), who study the role of anticipated regret in single-stage auctions. Figure 3.2 presents the distribution of intensity of different emotions stated by the traders. We observe that a considerable number of traders indeed feel substantial regret for not being qualified for the second stage. This observation supports the conjecture that (at least some of) the traders' bidding behavior in a two-stage auction would likely be affected by anticipated non-qualification regret.

⁴We do not test other behavioral theories in the interviews because: (i) we had to be mindful of the traders' time constraints; (ii) other factors, such as risk aversion (Kahneman and Tversky 2013), are unlikely to become less/more salient due to the launch of two-stage auction.

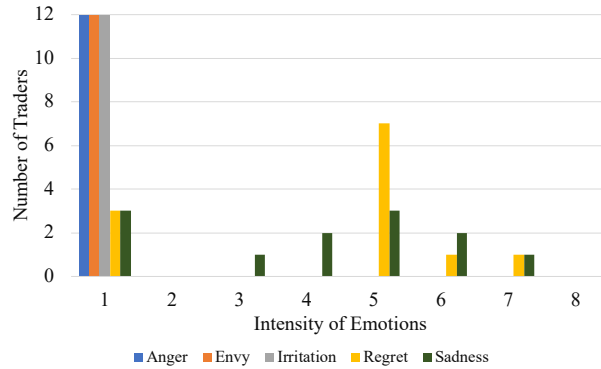


Figure 3.2: Distribution of the stated intensity of different emotions if not qualified for the second stage. A higher score indicates a stronger feeling of the corresponding emotion: 1 = no feeling at all, and 9 = an extremely strong feeling.

Competitive arousal versus anchoring.

In §3.3.1, we identify two theories that make competing predictions on traders' bidding behavior in the second-stage. In particular, while the competitive arousal theory suggests that traders are likely to bid more aggressively in the second stage (due to, e.g., a reduced number of rivals), the theory of anchoring suggests instead that the second-stage bids would be very close to the highest first-stage bid. To test these competing hypotheses, the traders were asked to hypothetically bid for a lot they qualify in the second stage across different scenarios with varying numbers of qualified traders. Contrary to the competitive arousal theory, the traders did not bid higher when fewer traders qualified. In fact, two traders bid lower when there were fewer competitors. Instead, the traders exhibit strong anchoring on the highest first-stage bid disclosed to them and outbid this value by no more than 0.01%. Therefore, results from the interviews indicate that anchoring, rather than competitive arousal, would more likely play a role in affecting the traders' bidding behavior in our context.

Costly information acquisition – quality uncertainty.

As discussed in §3.3.1, if substantial uncertainty about the lots' quality exists, then introducing the second stage can allow the traders to obtain more accurate quality signals for the lots they qualify for the second-stage bidding. Contrary to this hypothesis, however, the traders interviewed explain that there is little uncertainty about the lots' quality once they have physically inspected the lots before the first-stage bidding, and there is no need to reinspect the lots in the second stage. This is because the primary factors determining the quality of tur are lot-level features, such as seed size, moisture level, and foreign particles in the lot, and the traders are all

Table 3.1: Summary of Theoretical Predictions and Results From Field Interviews

Theory	Theoretical Prediction	Field Interview Insights
Anticipated non-qualification regret	Increased first-stage bids	Find evidence: A majority of traders exhibit regretful emotions due to non qualification
Competitive arousal vs anchoring	Competitive arousal predicts more aggressive bidding in the second stage; anchoring predicts limited bid increase in the second stage	Evidence for anchoring: Second-stage bids at most 0.01% higher than the highest first-stage bids
Costly information acquisition-quality uncertainty	Improved quality signal after the first stage could increase second-stage bids	Does not apply: No quality uncertainty given inspection before the first stage

highly experienced in quality inspection. In addition, beyond these observable quality features, the remaining differences in the traders' valuations for a lot are primarily driven by idiosyncratic factors related to each trader's specific buyers or demand conditions. These results suggest that auction models considering costly information acquisition or interdependent valuations are not a good candidate for characterizing the traders' bidding behavior in our context.

3.4 A Behavioral Auction Model

Table 3.1 summarizes the results from the field interviews and compares them to the theoretical predictions from the literature. Results from the field interviews reveal two salient behavioral factors that affect traders' bidding behavior in the two-stage auction: (i) anticipated non-qualification regret affects their bidding in the first stage, and (ii) anchoring on the highest first-stage bid affects their bidding in the second stage. These insights motivate the development of a simple model that jointly captures these factors to characterize the traders' bidding behavior in the two-stage auction. The main goal is to develop a behaviorally plausible model to examine *when* a two-stage auction outperforms a single-stage first-price sealed-bid auction to generate higher revenue for the farmers.

Specifically, the model considers three traders bidding for a single lot and the top two traders in the first stage qualify for the second stage.⁵ Because quality can be easily verified

⁵Although the traders bid for multiple lots every day, they commented in the interviews that their bidding for each lot is independent because they do not have any budget constraints and could always sell all the lots they win due to high demand from their buyers. Hence, we focus on modeling the auction of a single lot.

by the traders and is an important factor determining the bid prices, the trader's valuation of a lot is modeled as $q + V$, where q is the commonly observed quality of the lot (known to all three traders), and V is each trader's private value. This private value captures idiosyncratic factors related to the trader's specific buyers or demand conditions that may affect his bidding (see §3.3.1).⁶ Each trader observes his own private value v and knows that the other traders' private values are independently and uniformly distributed on $[0, 1]$. Trader i 's bidding strategy can be characterized by a mapping $B_i(q, v) : [q, q + 1] \rightarrow \mathcal{R}^+$, which specifies his bid given his valuation $q + v$. Following the auction literature, the analysis is focused on characterizing a symmetric equilibrium in which all traders' equilibrium bidding strategies follow the same structure, denoted as $B^*(q, v)$. Given this setup, the following lemma presents the symmetric equilibrium in a single-stage first-price sealed-bid auction with loser's regret.⁷ All proofs are deferred to the online appendix.

Lemma 3.4.1. Under a first-price, sealed-bid auction, there exists a symmetric Bayesian Nash equilibrium in which the traders' bidding strategy is given by $B_{FP}^*(v, q) = q + \frac{2 + 2\lambda_L}{3 + 2\lambda_L}v$.

The next step is to characterize the traders' equilibrium bidding strategy in the two-stage auction. First recall from §3.3.1 that traders exhibit strong anchoring on the highest first-stage bid, denoted by $q + H$. To capture this anchoring behavior, the assumption is that the top-ranked trader bids $q + H$ again, and the second-rank trader bids $q + H$ if he makes a positive profit with this bid. The analysis characterizes the expected payoff of a trader as a function of his first-stage bid $q + b$. There are three possibilities. If the trader is ranked first, then he wins and earns a revenue of $(v - b)$ with probability $1/2$ (1) if the second-rank trader bids again (does not bid) in the second stage. If the trader is ranked second, then he wins and earns $(v - H)$ with probability $1/2$ if he bids again. If the trader does not qualify, then he feels non-qualification regret (§3.3.2), with the magnitude of the regret proportional to the lost revenue from non-qualification, i.e., $(v - H)^+$. This modeling choice follows from the literature (Filiz-Ozbay and Ozbay 2007, Engelbrecht-Wiggans and Katok 2007, 2008). The following theorem characterizes the traders' symmetric equilibrium bidding strategy in the first stage of the two-stage auction.

⁶Athey et al. (2011) make a similar assumption in their study of timber auctions.

⁷The literature has shown that the presence of loser's regret is salient in single-stage auctions while winner's regret is not (Filiz-Ozbay and Ozbay 2007). We posit that the loser's regret may be relevant in our context as well. This is because only the winning bid for each lot is disclosed to the traders on the UMP under the current auction format. Thus, we incorporate the loser's regret in the single-stage first-price sealed-bid auction model to achieve a fairer comparison with the behavioral auction model we develop. The results remain unchanged if we do not incorporate this regret.

Theorem 3.4.2. In the two-stage auction, there exists a symmetric equilibrium in which the bidding strategy of a trader with a regret intensity $\lambda > 1/2$ can be characterized as follows: $b(0) = q$ and $\partial b(v)/\partial v = 2(1 + \lambda) \int_v^1 (v - b(x))^+ dx / [(v - b(v))^2 + 2b(v)^2]$.

Comparing Theorem 3.4.2 to Lemma 3.4.1, the following result can be shown.

Proposition 3.4.3. When the traders exhibit anticipated regret and anchoring, there exists a threshold $\bar{\lambda}_H$ such that if $\lambda \geq \bar{\lambda}_H$, then the expected revenue under a two-stage auction will be higher than that under a single-stage first-price sealed-bid auction.

The model analysis shows that the two-stage auction with anticipated non-qualification regret and anchoring leads to a higher expected revenue for the farmers than the single-stage first-price sealed-bid auction if the intensity of the traders' anticipated non-qualification regret is sufficiently strong. This result is in sharp contrast to prior results that the expected revenue from a single-stage first-price sealed-bid auction with loser's regret is *always* higher than that without loser's regret (Engelbrecht-Wiggans and Katok 2008). The key underlying reason is that anticipated non-qualification regret and anchoring have opposing effects on the traders' bidding behavior. On the one hand, since a trader would win with a positive probability as long as he qualifies for the second stage (as opposed to having to be the highest bidder to win in a single-stage first-price auction), he would bid lower in the first stage than in the first-price auction, and anchoring would lead to a lower winning bid. On the other hand, anticipated non-qualification regret motivates a trader to bid higher in the first stage than in the first-price auction. Therefore, the two-stage auction would generate a higher revenue for the farmers than the single-stage first-price auction if the effect of anticipated non-qualification regret on the traders' bidding behavior is stronger than that of anchoring.

An immediate implication from the model analysis is that implementing the two-stage auction would be beneficial for those commodities for which the traders tend to exhibit strong anticipated non-qualification regret. This is likely to be the case for commodities whose supply is constrained relative to market demand or those that are of high economic values. This result strengthens our confidence that the two-stage auction would benefit tur farmers, because tur, as well as other lentils, have limited supply relative to the consistently high consumption across the country due to their importance in the mostly vegetarian Indian diets. In addition, the analysis suggests that high-value cash crops such as areca nut, cotton, and turmeric are also good candidates for implementing the two-stage auction.

3.5 Implementing and Empirically Evaluating the Impact of the Two-Stage Auction

The two-stage auction design is akin to conducting multiple first-price sealed-bid auctions, and thus, has a strong implementation appeal to the government. Transitioning from the traditional markets to the UMP involved substantial technical investments and on-the-ground effort to train and build trust with key stakeholders including farmers and traders. Therefore, to be mindful of resource constraints, and more importantly, to preserve the hard earned trust from farmers, any change to the auction mechanism had to be incremental vis-a-vis the current first-price, sealed-bid auction. The new auction design minimizes disruption to the current trading and post-auction processes, is easy to explain to farmers and traders, and requires small technical changes on the platform to accommodate resource constraints.

3.5.1 Field Implementation

The market in which the interviews were conducted was chosen to be the treatment market for implementing the two-stage auction. This choice is driven by two key reasons. First, there exists another similar market that can serve as the control market to allow for empirical evaluation of the impact of the implementation on market prices. Second, the field interview data and the lot-level auction data from the implementation can later be combined to validate the behavioral auction model developed. To ensure a successful launch, extensive training and discussions were carried out with the market participants weeks before the implementation. Particularly to get the traders' buy-in, the two-stage auction design was "marketed" to them as giving them an additional chance to win the lots that they would have otherwise lost. A number of implementation details resulted from interactions with the traders during the training phase. First, the traders requested that the length of the second stage should be no more than half an hour, because they had to perform a series of post-auction tasks by the end of the day (e.g., weighing, cleaning, packing, and transporting the lots, transferring payments to the farmers). Second, considering the potential average number of lots that a trader would be qualified to bid again (based on historical data) and the feasibility of doing so within half an hour, only the top 3 bidders for each lot would be qualified, i.e., $k = 3$ in Figure 3.1. Third, the traders requested that for lots they qualify, their own rank among the top three should also be disclosed (in addition to the highest first-stage bid) to allow for more informed bidding in the second stage.

After accounting for these feedback, the final implementation was determined as follows: (i) the cutoff time for the first-stage bidding would remain at 2:30 p.m. as in the current process; (ii) at that time, the traders would receive mobile text messages that inform them for which lots they qualify for the second stage; (iii) the second-stage bidding would begin at 2:35 p.m.; (iv) the traders would observe on the UMP the highest first-stage bids and their own rank for the lots they qualify, and they have until 3 p.m. to increase their bids; (v) at 3 p.m., the winners for *all* lots in the market would be declared. A single winner declaration time for all lots is chosen to simplify the coordination of post-auction activities.⁸ The implementation was launched on February 22, 2019.

To empirically evaluate the impact of implementing the two-stage auction on farmers' revenue, we adopt a difference-in-differences (DID) approach (Imbens and Wooldridge 2009). The control market is selected based on the following criteria: (i) it must be a major market of tur, as is the treatment market; (ii) there are a similar number of traders in the two markets; (iii) the prices in the two markets pre implementation must satisfy the parallel trends assumption; (iv) the two markets are sufficiently far apart (223 kilometers away) that no traders or farmers would switch to the other market for trading, as confirmed by market officials; (v) domain knowledge from the collaborators further confirms that the two markets are similar in terms of infrastructure, the quality of the lots, and the characteristics of the traders.

The empirical analysis utilizes lot-level auction data for the control and treatment markets from one month prior to the implementation till May 31, 2019.⁹ The data contains all of the bids that were placed on the UMP within this time window in both markets, including all bids placed in both stages after the two-stage auction was implemented in the treatment market (a total of 172,342 bids). Figure 3.3a presents the distribution of the winning bids in the control and treatment markets pre-implementation (left) and post-implementation (right). Figure 3.3b presents the time series of the average weekly prices in the control and treatment markets during the time window of our analysis. The vertical line indicates the date on which the two-stage auction was implemented in the treatment market. It can be observed from both the figures that the difference in market prices between the two markets widens after the implementation. These observations provide preliminary evidence that implementing the two-stage auction has

⁸Since the second-stage bidding is done within a short time, we do not expect increased risk of collusion from the two-stage auction.

⁹We use this cutoff date because market officials confirm that the 2019 season for tur ends by the last week of May in the state. Aggregate supply data obtained from the UMP confirms that approximately 93% of the total quantity traded in 2019 was before June.

resulted in better prices for farmers in the treatment market. We next present the details of the formal impact assessment.

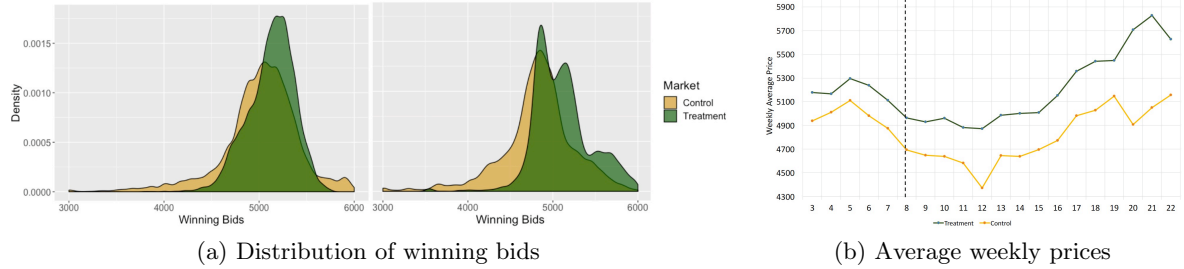


Figure 3.3: Comparing Treatment and Control Markets Pre and Post Implementation

Figure 3.3a presents the distribution of winning bids at the lot level in the control and treatment markets pre (left) and post (right) implementation. Figure 3.3b presents the time series of the average weekly prices in the control and treatment markets. The vertical line indicates the date on which the two-stage auction was implemented.

3.5.2 Comparability of the Treatment and Control Markets

We first verify the comparability of the two markets along a few important market features. Table 3.2 presents the summary statistics related to tur trading in the control and treatment markets. We observe that the two markets are comparable in both overall supply (e.g., total number of lots arriving in the market) and demand (e.g., number of traders, average number of bids per lot, and average price) features prior to implementation. Furthermore, domain knowledge from the collaborators also assures that the two markets are both major markets of tur with similar characteristics in terms of market infrastructure, the quality of the lots, and the characteristics of the traders.

Table 3.2: Summary Statistics for Tur Trading in the Control and Treatment Markets Pre-Implementation

	Total no. of lots	Total no. of traders	Average no. of bids per lot	Average price (rupees/100kg)
Control market	11,327	45	5.21	4,987
Treatment market	9,642	43	5.75	5,126

3.5.3 Parallel Trends Assumption

The key identifying assumption for a DID approach is that the difference in market prices between the treatment and control markets would have remained constant over time in the absence of the treatment; i.e., the price trends are parallel between the two markets. Following

established methods in the literature (Autor 2003, Angrist and Pischke 2008), we verify whether this parallel trend assumption is satisfied by estimating the following model:

$$\log(P_{m,t,l}) = \left(\sum_{w=1}^5 \gamma_w M_{w,l} \right) + \delta_1 Q_{m,t,l} + \delta_2 N_{m,t} + (\text{market, date fixed effects}) + \epsilon_{m,t,l}. \quad (3.1)$$

The dependent variable $\log(P_{m,t,l})$ is the logarithm of the winning bid for lot l in market m on day t . $Q_{m,t,l}$ is the size of lot l in market m on day t . $N_{m,t}$ is the total number of lots arriving in market m on day t . $M_{w,l}$ is an indicator variable that is equal to 1 if lot l is traded w weeks prior to the implementation in the treatment market and 0 otherwise. $\epsilon_{m,t,l}$ is the idiosyncratic error term. Given this specification, the parallel trend assumption is satisfied if a majority of the estimates of γ_w are not statistically significant (Cui et al. 2018, Ganesh et al. 2019). The estimation results (Table 3.3) show that four out of the five coefficients are not statistically significant.

Table 3.3: Estimation Results for the Parallel Trends Test

Variable	Estimate
$I_{m,t,l} \times M_{1,l}$	-0.005 (0.004)
$I_{m,t,l} \times M_{2,l}$	0.008 (0.009)
$I_{m,t,l} \times M_{3,l}$	-0.005 (0.007)
$I_{m,t,l} \times M_{4,l}$	-0.017** (0.005)
$I_{m,t,l} \times M_{5,l}$	-0.009 (0.005)
$Q_{m,t,l}$	0.002 (0.002)
$N_{m,t}$	0.00004 (0.00002)
Fixed effects	Y
Observations	16,371

Notes. We report the coefficient estimates of γ_w for $w = 1, \dots, 5$ in Equation (3.1). Standard errors (in parentheses) are clustered at the market and date level. ***: $p < 0.01$; **: $p < 0.05$. We control for market and date fixed effects in the model.

3.5.4 DID Model

We estimate the following model as our main DID specification:

$$\log(P_{m,t,l}) = \gamma I_{m,t,l} + \delta_1 Q_{m,t,l} + \delta_2 N_{m,t} + (\text{trader, date fixed effects}) + \epsilon_{m,t,l}. \quad (3.2)$$

$I_{m,t,l}$ is the treatment dummy: $I_{m,t,l} = 1$ for all lots in the treatment market on all dates after the implementation, and $I_{m,t,l} = 0$ otherwise. The variables $\log(P_{m,t,l})$, $Q_{m,t,l}$, $N_{m,t}$, and $\epsilon_{m,t,l}$ are defined the same as in Equation (3.1). We control for the size of the lot ($Q_{m,t,l}$) and the total number of lots arriving in the market m on day t ($N_{m,t}$) in Equation (3.2) because they could affect the price of a lot.¹⁰ Since farmers' decision on when and how much to sell are significantly affected by production side factors, the treatment is unlikely to affect either of these control variables in the short term. Nevertheless, we consider an alternative specification in §3.5.5 without these control variables to test the robustness of our results. The model controls for trader and date fixed effects. The coefficient of interest is γ . A positive and significant value of γ indicates that the implementation of the two-stage auction has led to a significant price increase in the treatment market. Standard errors are clustered at the market and date level to account for potential correlation across observations (Bertrand et al. 2004).

In addition to Equation (3.2), the following model is estimated to examine whether the treatment effect changes over time as the market participants gain experience with the new auction design:

$$\log(P_{m,t,l}) = I_{m,t,l} \times \left(\sum_{t'=1}^4 \gamma_{t'} M_{t',l} \right) + \delta_1 Q_{m,t,l} + \delta_2 N_{m,t} + (\text{trader, date fixed effects}) + \epsilon_{m,t,l} \quad (3.3)$$

The variable $M_{t',l}$ is an indicator variable that is equal to 1 if lot l is traded in the t^{th} month after implementation and 0 otherwise. All the remaining variables follow from Equation (3.2).

Table 3.4 presents the estimated average impact of implementing the two-stage auction on tur prices in the treatment market based on Equations (3.2) and (3.3). Observe from column 2 that the implementation has yielded a statistically significant 4.7% increase in tur prices in the treatment market. In addition, column 3 shows that the positive impact of the implementation tends to increase over time (1.6% in February to 6.7% in May). The estimated 4.7% average price increase translates into a revenue gain of \$452,000 (USD) for over 10,000 tur farmers who traded in the treatment market in a matter of three months, i.e., a 15% increase in their monthly income.¹¹ Given low profit margins for these farmers (4%-7%), the range of profit improvement

¹⁰The total number of lots arriving in the market is strongly and positively correlated with the total supply quantity and the total number of active traders in the market. Therefore, we can only include one of them as a control variable. We confirm that our results remain the same regardless of which one of these three variables we control for.

¹¹We use the total number of lots sold in the treatment market as a proxy for the total number of farmers who traded in the market. Since the same farmer may visit and sell in the market for multiple days, the stated 15% income raise is a conservative estimate.

Table 3.4: Estimated Impact of Implementing the Two-stage Auction on Tur Prices

Variable	Equation (3.2)	Equation (3.3)
$I_{m,t,l}$	0.047*** (0.009)	–
$I_{m,t,l} \times M_{1,l}$ (February)	–	0.016** (0.008)
$I_{m,t,l} \times M_{2,l}$ (March)	–	0.056*** (0.009)
$I_{m,t,l} \times M_{3,l}$ (April)	–	0.049*** (0.011)
$I_{m,t,l} \times M_{4,l}$ (May)	–	0.067*** (0.014)
$Q_{m,t,l}$	0.0006 (0.0008)	0.0006 (0.0008)
$N_{m,t}$	0.00002 (0.00003)	0.00001 (0.00003)
Fixed effects	Y	Y
Observations	30,867	30,867

Notes. “–” means the variable is not present in the model. Standard errors (in parentheses) are clustered at the market and date level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. We control for trader and date fixed effects in both models.

is substantial (60%-158%). These positive results are particularly significant given that the launch of the UMP itself had not significantly improved tur prices prior to the implementation of the two-stage auction.

3.5.5 Robustness Tests

We perform a number of robustness analyses to further strengthen our results. In particular, we consider (i) commodity and time placebo tests; (ii) filtering on the subset of traders who were active both pre and post implementation; (iii) alternative control variables; (iv) alternative clustering of standard errors; and (v) additional data from the 2020 season. We consider our main result to be robust if the direction and statistical significance of the coefficient for the implementation dummy are consistent between the main model (Equation (3.2)) and these robustness tests.

Placebo tests.

We perform two placebo tests to ensure that the estimated impact discussed above is indeed a result of the field implementation. First, we reestimate Equation (3.2) using data for a placebo commodity (Bengal gram, another lentil) that was traded in the same markets during the same

time period of the implementation but followed the original single-stage first-price sealed-bid auction. Second, we reestimate Equation (3.2) using tur trading data from a placebo time period (the same months in 2018). If the field implementation is the main driver of the observed impact, then we should not observe any statistically significant effect in these two tests. The regression results indeed show no statistically significant effect in either test (Table 3.5, columns 2 and 3).

Active traders only.

We observe that not all the winning traders are active in *both* pre- and post-treatment periods. If traders who won only in pre- or post-treatment periods had substantially different valuation than other traders, the estimated impact could be confounded by these valuation differences. To mitigate this potential confound, we focus on the subset of traders who have won at least one lot in both pre- and post-treatment periods and reestimate Equation (3.2). The regression results show a similar statistically significant effect as in the main DID model (Table 3.5, column 4).

Alternative control variables.

Here, we reestimate Equation (3.2) using two additional specifications with different sets of control variables. First, we reestimate Equation (3.2) with a market specific time trend. This specification allows the control and treatment markets to follow different time trends (Angrist and Pischke 2008). Second, we remove $Q_{m,t,l}$ and $N_{m,t}$ from Equation (3.2) and reestimate the model. The regression results for both the specifications show a similar statistically significant effect as in the main DID model (Table 3.5, columns 5 and 6).

Alternative clustering of standard errors.

We consider alternative standard error clustering by doing one way market-level clustering to account for within-market correlations and reestimate Equation (3.2). The regression results show a similar statistically significant effect as in the main DID model (Table 3.5, column 7).

Additional data.

We reestimate Equation (3.2) after including additional data from the 2020 season. Due to COVID-19, the 2020 season is available only for January and February of 2020 because the markets were shut down starting March 2020. Nevertheless, the regression results continue

to show a positive and statistically significant impact from the implementation for both years (Table 3.5, column 8). This result provides preliminary evidence that the revenue benefit of the two-stage auction sustains over time.

Table 3.5: Regression Results of the Robustness Tests

Variable	Commodity placebo	Time placebo	Active traders only	Market-level trend	No control variables	Alternative clustering	Additional Data
$I_{m,t,l}$	0.008 (0.006)	0.004 (0.004)	0.048*** (0.008)	0.013** (0.007)	0.044*** (0.003)	0.047*** (0.003)	–
$I_{m,t,l} \times Y_{2019}$	–	–	–	–	–	–	0.023*** (0.006)
$I_{m,t,l} \times Y_{2020}$	–	–	–	–	–	–	0.033* (0.019)
Control variables	Y	Y	Y	Y	N	Y	Y
Fixed effects	Y	Y	Y	Y	Y	Y	Y
Observations	14,966	33,155	28,542	30,867	30,867	30,867	77,382

Notes. Column 2 presents results from estimating Eq. (3.2) based on trading data of a placebo commodity (Bengal gram) in the same period of the implementation. Column 3 presents results from estimating the same model based on tur trading data from February to May in 2018. Column 4 presents results from estimating the same model based on filtering on active traders. Column 5 presents results with additional control for market level trends. Column 6 presents results with no control variables except market and date fixed effects. Column 8 presents the results with additional data from 2020 and estimates the impact for 2019 and 2020 separately. Standard errors (in parentheses) are clustered at the market and date level in all specifications except for column 7 (clustered at the market level). We control for trader and date fixed effects in all specifications. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$.

3.5.6 Validating the Behavioral Auction Model

The empirical analysis thus far demonstrates that the field implementation has led to a significant revenue benefit for farmers in the treatment market. The next step in the analysis uses the auction and interview data to provide evidence that the two behavioral factors captured in the auction model – anchoring and anticipated non-qualification regret – indeed play a significant role in affecting the traders' bidding behavior.

Anchoring.

Across all lots that received second-stage bids, the bid increment in the second stage never exceeds 0.75% of the highest first-stage bid, with an average of 0.2% (10 rupees/100 kg compared to the average price of about 5,000 rupees/100kg; see Table 3.2 and Figure 3.4). This observation provides strong evidence of anchoring behavior by the traders and suggests that the positive

impact of the implementation mainly comes from more competitive bidding by the traders in the first stage. To investigate this postulation, we reestimate Equation (3.2) using a lot's highest bid in the first stage as the dependent variable instead. Since there is only one stage of bidding in the control market, the dependent variable for the lots in the control market remains the same as before. Table 3.6 presents the results. We observe that the implementation dummy has a significantly positive coefficient ($\gamma = 0.046$) that is of similar magnitude as the estimated impact in our main model (Table 3.4, column 2).

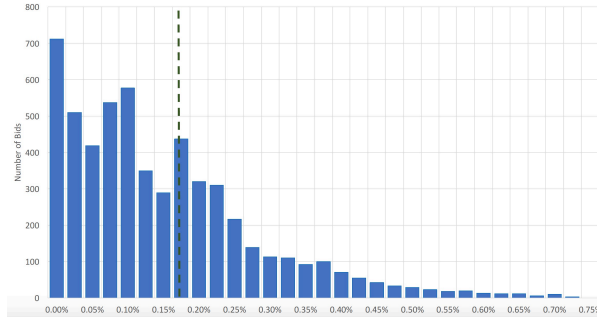


Figure 3.4: Distribution of Bid Increments in the Second Stage

Second-stage bid increment is calculated as a qualified trader's second-stage bid minus the highest bid in the first stage for the associated lot. The vertical line indicates the average second-stage bid increment across all lots and traders. The figure only accounts for those lots that indeed received second-stage bids.

Table 3.6: Estimated Impact of the Implementation on the Highest First-Stage Bids

Variable	Estimate
$I_{m,t,l}$	0.046*** (0.008)
$Q_{m,t,l}$	0.001 (0.001)
$N_{m,t}$	-0.0000 (0.0000)
Fixed effects	Y
Observations	30,867

Notes. Standard errors (in parentheses) are clustered at the market and date level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. We control for trader and date fixed effects in the model.

Anticipated non-qualification regret.

To test the role of anticipated non-qualification regret, the traders' stated intensity of regret felt in the case of non-qualification in the interviews (§3.3.2) was used to divide the traders into two groups. The traders who do not exhibit any regret (stated score for regret = 1) are defined as the "low-regret" group. The remaining traders (stated score for regret ≥ 5) are defined as

the “high-regret” group. Next, the following model is estimated to test whether there is an even larger price increase (relative to the control market) among those lots for which the top-rank trader in the first stage is in the high-regret group.

$$\log(P_{m,t,l}^1) = \gamma_0 I_{m,t,l} + \gamma_H I_{m,t,l} \times Z_{H,l} + \delta_1 Q_{m,t,l} + \delta_2 N_{m,t} + (\text{market, date fixed effects}) + \epsilon_{m,t,l}. \quad (3.4)$$

The variable $Z_{H,l}$ is an indicator variable that is equal to 1 if the top-rank trader in the first stage for lot l is in the high-regret group and 0 otherwise. The dependent variable is the logarithm of the highest first-stage bid for lot l in market m on day t . The remaining variables are defined as before. Table 3.7 presents the estimation results. The estimate of γ_H is 0.022 (s.e. = 0.007, $p < 0.01$). That is, there is a 2.2 percentage point additional increase (or a 73% increase) in the highest first-stage bid from high-regret traders versus low-regret ones. This result provides empirical support for the significant role that anticipated non-qualification regret plays in affecting the traders' bidding behavior in the two-stage auction.

Table 3.7: Estimated Impact of the Implementation on the Highest First-Stage Bids: Effect of Anticipated Non-qualification Regret

Variable	Estimate
$I_{m,t,l}$	0.030*** (0.003)
$I_{m,t,l} \times Z_{H,l}$	0.022*** (0.007)
$Q_{m,t,l}$	0.002 (0.002)
$N_{m,t}$	0.00003** (0.00000)
Fixed effects	Y
Observations	25,435

Notes. Standard errors (in parentheses) are clustered at the market and date level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. We control for trader and date fixed effects in the model.

3.5.7 Additional Insights

We finally discuss two additional insights regarding the positive impact of implementing the two-stage auction on farmers' revenue.

The effect of product quality.

Quality is an important factor that determines prices for agricultural commodities. Thus, it is of great interest to examine whether the impact of the implementation on tur prices depends on a

lot's quality. Since the quality of a lot is not directly observable, we consider the following proxy measure. For a given lot l , we calculate its normalized number of bids, $B_{m,t,l}^n$, as the number of bids it received divided by the maximum number of bids received by any lot in the same market on the same day. Our field interactions with the traders indicate that higher-quality lots naturally attract more bids. Hence, $B_{m,t,l}^n$ would be closer to 1 (0) for lots of high (low) quality. The normalization allows us to do a fair comparison across days and markets. We then estimate the following two models.

$$\begin{aligned} \log(P_{m,t,l}) = & \gamma_0 I_{m,t,l} + \gamma_1 I_{m,t,l} B_{m,t,l}^n + \gamma_2 I_{m,t,l} (B_{m,t,l}^n)^2 + \delta_1 Q_{m,t,l} + \delta_2 N_{m,t} \\ & + (\text{trader, date fixed effects}) + \epsilon_{m,t,l}, \end{aligned} \quad (3.5)$$

$$\begin{aligned} \log(P_{m,t,l}) = & \gamma_0 I_{m,t,l} + I_{m,t,l} \times \left(\sum_{i \in H,M} \gamma_i B_{i,l} \right) + \delta_1 Q_{m,t,l} + \delta_2 N_{m,t} \\ & + (\text{trader, date fixed effects}) + \epsilon_{m,t,l}. \end{aligned} \quad (3.6)$$

In Equation (3.6), $B_{H,l}$ is an indicator variable that is equal to 1 if $B_{m,t,l}^n$ is greater than the third quartile of the distribution of $B_{m,t,l}^n$ in the same market on the same day, and 0 otherwise. Similarly, $B_{M,l}$ is an indicator variable that is equal to 1 if $B_{m,t,l}^n$ is between the first and the third quartiles of the same distribution, and 0 otherwise. That is, we divide all lots arriving at a given market on a given day into three groups based on a first and third quartile split. The baseline group in Equation (3.6) represents the low-quality lots, and $B_{M,l}$ ($B_{H,l}$) represents the medium-quality (high-quality) lots.¹² All other variables remain the same as before. Our method of evaluating possible differential impacts of the implementation follows from the literature (Parker et al. 2016, Zhang et al. 2018).

The regression results are presented in Table 3.8 and demonstrate that the implementation has resulted in a stronger positive impact on the prices of higher quality lots. First, we observe a significantly positive coefficient for $B_{m,t,l}^n$, our proxy measure of lot quality, in column 2 of Table 3.8. Furthermore, the coefficient for $(B_{m,t,l}^n)^2$ is significantly negative. Thus, lots receiving a larger number of bids benefit more from the implementation, although this effect has a (naturally) diminishing return. Second, we observe from column 3 of Table 3.8 that the prices of high-quality (medium-quality) lots have increased 4 percentage points (3 percentage

¹²Our results remain the same under the following alternative groupings: (i) a four-group split based on the first quartile, median, and third quartile; (ii) a three-group split based on the 10th and 90th percentiles; and (iii) a median split.

points) more relative to the prices of low-quality lots (which have increased by 1.7%) due to the implementation. These results highlight that in addition to generating price gains for the farmers, implementing the two-stage auction also provides them with stronger incentives to enhance the quality of the lots they bring to the market.

Table 3.8: Estimated Impact of the Implementation on Tur Prices: The Effect of Quality

Variable	Equation (3.5)	Equation (3.6)
$I_{m,t,l}$	0.047*** (0.009)	0.017*** (0.009)
$I_{m,t,l} \times B_{m,t,l}^n$	2.100*** (0.562)	–
$I_{m,t,l} \times (B_{m,t,l}^n)^2$	-1.110*** (0.043)	–
$I_{m,t,l} \times B_{H,l}$	–	0.040*** (0.003)
$I_{m,t,l} \times B_{M,l}$	–	0.030*** (0.001)
$Q_{m,t,l}$	0.001 (0.001)	0.001 (0.001)
$N_{m,t}$	-0.000 (0.000)	-0.000 (0.000)
Fixed effects	Y	Y
Observations	30,867	30,867

Notes. “–” means the variable is not present in the model. Standard errors (in parentheses) are clustered at the market and date level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. We control for trader and date fixed effects in both models.

Changes in bids over time.

Here we examine how the average and standard deviation of (daily) first-stage bids placed by individual traders in the treatment market may have changed over time, as compared to the control market. In particular, we estimate the following model.

$$\log(M_{b,t}) = I_{b,t} \times \left(\sum_{t'=1}^4 \gamma_{t'} M_{t',b} \right) + R_{m,t} + (\text{trader, month fixed effects}) + \epsilon_{b,t}. \quad (3.7)$$

The dependent variable $\log(M_{b,t})$ is the logarithm of the average first-stage bids placed by trader b on day t . $I_{b,t}$ is the treatment dummy: $I_{b,t} = 1$ for all traders in the treatment market on all dates after the implementation, and $I_{b,t} = 0$ otherwise. As in Equation (3.3), the variable $M_{t',b}$ is an indicator variable that is equal to 1 if the observation from trader b is in the t^{th} month

after implementation and 0 otherwise. $R_{m,t}$ is the number of traders in market m on date t . $\epsilon_{b,t}$ is the idiosyncratic error term. We control for trader and month fixed effects in this model.¹³ Standard errors are clustered at the market level. We also estimate a similar model as Equation (3.7) but use the standard deviation of the daily first-stage bids submitted by each individual trader as the dependent variable.

Table 3.9: Estimated Impact of Implementing the Two-stage Auction on the Average and Standard Deviation of First-Stage Bids at the Trader Level

Variable	Average	Standard Deviation
$I_{m,t,l} \times M_{1,l}$ (February)	-0.029** (0.009)	-0.026 (0.041)
$I_{m,t,l} \times M_{2,l}$ (March)	0.028*** (0.002)	-0.058*** (0.013)
$I_{m,t,l} \times M_{3,l}$ (April)	0.015*** (0.004)	-0.099*** (0.012)
$I_{m,t,l} \times M_{4,l}$ (May)	0.036*** (0.003)	-0.145*** (0.013)
$N_{m,t}$	0.002* (0.001)	0.010*** (0.0003)
Observations	2,749	2,749

Notes. Standard errors (in parentheses) are clustered at the market level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. We control for trader and month fixed effects in both models.

Table 3.9 presents the estimation results. Consistent with our main results, the traders in the treatment market had in general placed significantly higher bids than those in the control market since the implementation of the two-stage auction, and this difference persists over time (column 2). Interestingly, we also observe from column 3 that the implementation has led to a significant decrease in the standard deviation of the bids placed by the traders on the same day, and this decrease is more substantial in later months (14.5% reduction in May vs. 5.8% reduction in March). To summarize, the implementation of the two-stage auction has both increased the average bid price and decreased the bid dispersion of individual traders in the treatment market.

3.6 Conclusions

This chapter introduces a behavior-centric, field-based, data-driven methodology to design effective auction mechanisms to enhance farmers' revenue in online agri-platforms. By collaborating closely with the state government of Karnataka, we *design, analyze, and implement* a new two-

¹³Since each observation is at the date level, we can no longer control for date fixed effects.

stage auction on the state's agri-platform for a major lentils market. A difference-in-differences analysis, based on lot-level auction data demonstrates significant revenue gain for over 10,000 smallholder farmers traded in the market in a matter of three months. Encouraged by the positive outcomes and guided by the research insights, the state government has selected a set of suitable commodities and markets to next implement the two-stage auction on a larger scale. We believe that the methods introduced in this chapter can provide generally applicable knowledge to researchers and platform designers as they continue to enhance the design of agri-platforms to improve the livelihood of smallholder farmers.

Chapter 4

Economically Motivated Adulteration in Farming Supply Chains

4.1 Introduction

Food adulteration is a serious threat to public health and a major concern to most governments in both developed and developing countries. Food adulteration can occur in a broad range of scenarios. Unintentional adulteration often occurs as a result of negligence or incompetence, for example, bacterial contamination due to bad hygiene practices. In some scenarios, food adulteration is intentional, motivated by malicious intent to harm the public food system (e.g., bioterrorism). In many other scenarios, intentional adulteration is driven by economic motives and often referred to as *economically motivated adulteration* (EMA). The U.S. Food and Drug Administration defines EMA as the “fraudulent, intentional substitution or addition of a substance in a product for the purpose of increasing the apparent value of the product or reducing the cost of its production, i.e., for economic gain” (Johnson 2014). We particularly focus on EMA that causes harm to human health. Over the last several decades there were many publicly-known EMA incidents of food products around the world, and a majority of them originated from developing countries. For example, consumption of melamine-tainted infant formula and milk led to six infant deaths and nearly 300,000 young children severely sickened in China in 2008 (Everstine et al. 2013). Melamine was added to the milk by farmers and collectors to increase the perceived protein content of the milk. In another example, outbreaks

of avian flu led to extensive use of antibiotics and other illegal drugs in poultry farming in Asia. In particular, the 2012 KFC “instant chicken scandal” in China revealed that the chickens used by KFC were treated with as many as 18 illegal antibiotics on the farms (Pi et al. 2014). In both of these examples, the source of adulteration was in the upstream parts of the supply chains, specifically farms and collectors.

In this chapter, we develop a new modeling framework to examine strategic adulteration behaviors of farms (and/or collectors) and the resulting EMA risk in a farming supply chain. Our models simultaneously capture various major drivers for EMA, including the uncertainty or variability of the quality of the farms’ output, the dispersion and traceability of the supply chain, and the testing capabilities present in the supply chain. We address two main research questions: (i) What are the farms’ optimal or equilibrium adulteration strategies under different EMA scenarios? (ii) How do the above drivers jointly impact the farms’ adulteration behaviors? We validate the models with real cases and field data to ensure that the models are grounded in practice and consistent with empirical evidence. In addition, we analyze a few managerial levers, such as investing in traceability and testing capabilities, that a manufacturer can use to mitigate EMA risk in the supply chain. We leverage the analysis of our models to derive important and unique insights that can be used to help both regulators and commercial entities to better prioritize and address EMA risk more proactively. We next elaborate on the major EMA risk drivers captured in our models.

The first factor is the uncertainty or variability of the quality of a farm’s output. Quality uncertainty can result from issues inherent to the production process; e.g., the quality of milk produced from a cow, typically measured by its compositional characteristics such as protein and fat content, depends on the health of the cow. Quality uncertainty can also be the result of external factors; e.g., epidemics like avian flu affect the quality, captured by the health and weight, of chickens raised in a farm. Quality uncertainty can be a major cause of EMA in markets with quality-based pricing, i.e., where farms receive a better selling price if the products appear to have higher quality. We divide food adulteration driven by quality uncertainty into two distinct scenarios. The first scenario is called “preemptive EMA” where adulteration occurs *before* the uncertain quality of the products is realized. The primary goal of preemptive EMA is to decrease the likelihood of producing low-quality output. For example, farms overuse antibiotics to prevent producing sick or underweight (i.e., low-quality) animals. This is a serious concern in pork, poultry, and seafood farming in various countries including Bangladesh, China,

India, and Vietnam (Doyle et al. 2013). The second scenario is called “reactive EMA” where adulteration occurs *after* the uncertain quality of the products is realized. The primary goal of reactive EMA is to increase the perceived quality of low-quality products and create fake high-quality ones. For example, the intentional adulteration of raw milk with melamine emerged due to price pressure for low-protein milk in China (Sharma and Rou 2014). Similarly, farms in India adulterated milk with urea to increase its perceived solids-not-fat (SNF) content and attract higher prices (Tanzina and Shoeb 2016).

Another factor that may contribute to EMA is the dispersion of a farming supply chain. We define supply chain dispersion as the extent to which agricultural products are sourced from a distributed network of farms, each producing a small fraction of the total quantity (Huang et al. 2017b). Dispersed farming supply chains with hundreds or thousands of smallholder farms are prevalent in many developing countries (Narrod et al. 2008, Chen et al. 2014). Supply chain dispersion may increase EMA risk for at least two possible reasons. First, with a dispersed network of farms, it is difficult (if not impossible) to inspect every farm or to trace every unit of supply back to the producing farm. Instead, manufacturers (in the best case) only inspect the aggregated supply after the products from all farms are mixed. Due to limited traceability, even if the manufacturer detects adulteration in the mixed supply, it often cannot identify the problematic farms nor effectively impose penalties to deter the farms from adulterating. In practice, some firms try to improve traceability by storing samples from (some of) the farms before mixing the supply (Flynn and Zhao 2014). We model traceability in light of such practices. Second, smallholder farms rely on the revenue from selling their products to sustain their families. Hence, when they face quality uncertainty and the associated price pressure, they are likely to become aggressive and engage in adulteration to ensure their only means of income.¹

Adding to the complexity of a dispersed farming supply chain, the manufacturer’s testing capability in terms of test sensitivity is another factor affecting EMA risk in food products. We model both perfect and imperfect testing scenarios. Perfect testing corresponds to scenarios where a known adulterant is being tested and accurate methods exist to examine whether the residue amount (if any) exceeds the maximum allowable limit defined by food safety standards. For example, there are highly sensitive methods to detect certain antibiotics in food products (Pikkemaat 2009, Mungroo and Neethirajan 2014). However, since the space of possible adul-

¹In Appendix C.1.1, we show that farms’ aversion to quality uncertainty motivates them to adulterate even more.

terants is practically unlimited, it is difficult (if at all possible) to develop sensitive tests for every adulterant. We thus model imperfect testing to capture scenarios where the adulterant is less studied or even unknown. In these cases, the ability to detect adulteration often depends on the relative amount of adulterants being used. For example, the presence of a large quantity of adulterants may change the characteristics of the product (e.g., its texture, smell, or color) or may lead to adverse symptoms when people consume the product, signaling adulteration. In fact, the melamine-tainted infant formula scandal broke out precisely because the quantity of melamine added to the milk became so high that many children got ill, thus alerting the authorities.

Related literature and contributions: This chapter addresses a timely and crucial topic on food safety and makes important contributions to both research and practice. From a research standpoint, we develop a new modeling framework that realistically captures various major risk drivers of EMA and validate the model predictions with field data. Our work is related to three streams of literatures: supply chain risk management, quality management, and socially responsible operations. The supply chain risk management literature has mainly focused on the topic of supply *disruption* risks, typically modeled as exogenous shocks to the system (e.g., [Sheffi 2005](#), [Dada et al. 2007](#), [Van Mieghem 2007](#), [Tomlin 2009](#), [Babich 2010](#), [Wang et al. 2010](#), [Simchi-Levi et al. 2014](#)). A common conclusion in this literature is that dual or multi-sourcing helps to mitigate disruption risks and establish supply chain resilience. We instead focus on *quality* risks that stem from both exogenous uncertainty and endogenous actions within the supply chain. In this regard, we are closely related to works in the quality management and socially responsible operations literatures that examine opportunistic or unethical supplier behavior. We refer readers to [Nagurney and Li \(2016\)](#), [Atasu \(2016\)](#), and [Bouchery et al. \(2017\)](#) for comprehensive reviews on these two streams of literatures. We discuss here a subgroup of papers most relevant to ours. For example, [Babich and Tang \(2012\)](#) analyze deferred payment and inspection mechanisms as tools for a manufacturer to deter product adulteration by its supplier. [Cho et al. \(2015\)](#) study how a brand-name company may use quality and price levers to combat counterfeiters and show that either strategy could be ineffective when facing a deceptive counterfeiter. [Mu et al. \(2014, 2016\)](#) focus on milk supply chains with monopolistic or competing collection intermediaries and examine incentive schemes to induce better milk quality with minimal testing.

Our models differentiate from this group of papers in two key aspects. First, we explicitly

model exogenous quality uncertainty and distinguish it from endogenous adulteration decisions, whereas prior works consider adulteration as equivalent to producing low-quality products at a lower cost. Our approach is essential to adequately capture how quality uncertainty combined with quality-based pricing motivates adulteration. Second, while prior works treat detection of low-quality products as exogenous, we model imperfect testing where the probability of detecting adulteration is endogenously influenced by the farms' adulteration decisions. This endogeneity substantially complicates our analysis because we must consider the strategic interactions among farms and the interdependency of their adulteration decisions under quality uncertainty. In the case of reactive EMA with imperfect testing, we must solve for the farms' equilibrium adulteration strategies in a game of asymmetric information (as the number of realized low-quality units at each farm is the farm's private information).

The socially responsible operations literature studies supplier behaviors that are socially undesirable but do not directly impact the product's functional quality. For example, [Wang et al. \(2016\)](#) studies a regulator's optimal reward and inspection policy to motivate a company to voluntarily disclose a privately-observed, stochastically-occurring environmental hazard. [Chen and Lee \(2017\)](#) analyze how a company can employ contingency payment, supplier certification, and process audit to prevent unethical actions by a supplier. [Letizia and Hendrikse \(2016\)](#) and [Orsdemir et al. \(2016\)](#) examine how horizontal or vertical integration of supply chain parties may enhance responsible supplier practices. [Plambeck and Taylor \(2016\)](#) analyze a "backfiring condition" under which a supplier can effectively evade a buyer's audit and discuss its implications for motivating supplier responsibility. [Fang and Cho \(2016\)](#) and [Caro et al. \(2018\)](#) study joint or shared audits with a collective penalty and analyze when such audits lead to better supplier compliance compared to independent audits. [Cho et al. \(2018\)](#) examine a company's choice of inspection policy and wholesale price to combat a supplier's use of child labor. We differ from this body of research in the following ways. This literature mainly considers settings of a supply chain with one or two suppliers where the action of the supplier(s) may or may not be endogenized.² We instead focus on a network of suppliers each simultaneously making strategic decisions on adulteration. In addition, the two-step testing process and imperfect testing scenario in our context further complicate our analysis of strategic interactions among the

²Two recent exceptions are [Huang et al. \(2017a\)](#) and [Zhang et al. \(2017\)](#). The former studies a three-tier supply chain with one player in each tier and only the most upstream player could incur a responsibility risk. The latter examines the problem of curbing conflict minerals in a three-tier supply network consisting of manufacturers, smelters, and mines. They analyze the manufacturers' decisions to be compliant or not and the smelters' decisions to be certified or not in a deterministic setting.

suppliers, whereas audits of environmental and social issues typically act upon each individual supplier separately.

Due to these distinctive modeling features, our analysis yields new practical insights for addressing EMA risk in food more proactively. For example, we show that supply chain dispersion plays a critical role in affecting EMA risk. We determine when greater supply chain dispersion leads to a higher risk and how this result depends on traceability and testing capabilities. We demonstrate that supply chain dispersion could increase EMA risk beyond its impact on decreased traceability. In particular, even in a fully traceable supply chain, both preemptive and reactive EMA risk are higher in a more dispersed supply chain when testing is imperfect. Furthermore, increasing testing frequency has a limited effect on reducing EMA risk in a dispersed supply chain. Our results highlight the limitation of only relying on end product inspection to deter EMA unless highly robust testing methods can be developed (e.g., methods that can detect adulteration even without knowing what the adulterant is). Another new insight emerged from our analysis is that merely investing in quality improvement without enhancing testing capabilities at the same time may backfire and inadvertently increase EMA risk. This can occur when testing is imperfect, and hence, farms feel “safe” to adulterate to a moderate level without worrying about being caught later. Taken together, our results underscore the importance and necessity of applying a systemic, supply chain perspective to enable proactive management of EMA risk in food products. Such a perspective is missing in most current practices, which mainly rely on inspection and information at the product and individual company level to manage risk.

4.2 Modeling the Farming Supply Chain

We consider a farming supply chain in which a manufacturer (she) procures a total supply chain output of k units of an agricultural product from n homogeneous farms (he). Based on farming supply chain data we collected for multiple industries in China, we observe that when a manufacturer sources from multiple farms, the output level of each farm is very similar. We thus focus on homogeneous farms. Let m denote the number of output units from each farm, i.e., $m = k/n$. The quality of a farm’s output is uncertain ex ante. Specifically, the quality of each unit is low with probability p_L and high with probability $(1 - p_L)$.³ Hence, the total number

³In Appendix C.1.2, we analyze a case where the quality of all units from the same farm is perfectly correlated and obtain similar insights.

of low-quality units at a farm follows a binomial distribution with parameters m (number of units) and p_L (probability of each unit being low-quality)⁴.

We study the effect of supply chain dispersion by keeping the total supply chain output k constant and changing n . This effectively varies the fraction of the total supply chain output supplied by each farm. In particular, a larger n (i.e., a smaller m) represents higher supply chain dispersion, i.e., a supply chain with a larger number of farms each supplying a smaller fraction of the total output. Another parameter of interest is p_L , the probability of a unit being low-quality. Changing p_L allows us to study how a farm's adulteration behavior is affected by an increased or a decreased chance of producing low-quality output. A high p_L means there is a greater chance for the farm to produce low-quality output (e.g., during epidemics).

We study both preemptive and reactive EMA. Under preemptive EMA, farms make adulteration decisions *before* the uncertain quality of their output is realized. Preemptive EMA reduces the value of p_L , i.e., decreases the chance of producing low-quality output. We measure preemptive EMA risk in the supply chain based on the amount of adulterants added by the farms. Conversely, under reactive EMA, farms make adulteration decisions *after* the uncertain quality of their output is realized. Reactive EMA increases the perceived quality of the output and creates fake, high-quality units. Farms can condition their reactive EMA decisions on the realized number of low-quality units. Therefore, we measure reactive EMA risk in the supply chain by both the probability of an individual farm adulterating his output and the expected total number of adulterated units within the total supply chain output.

Figure 4.1 illustrates the sequence of events in our model for both preemptive and reactive EMA. The key differences between the two scenarios are in the first two steps. For preemptive EMA, the model dynamics are as follows: (i) Each farm simultaneously and individually decides the amount of adulterants to add to reduce p_L . (ii) The uncertain quality of the output is realized. (iii) The manufacturer purchases from all farms and pays each farm based on the average quality of his output. (iv) The manufacturer stores output samples from t randomly-chosen farms, where $t \in [0, n]$. (v) The manufacturer aggregates the output from all farms; with probability q , she tests the aggregated supply for adulteration. (vi) If adulteration is detected in the aggregated supply, then the manufacturer tests each of the stored samples. (vii) If a farm is found to have adulterated his output, then he is charged a penalty of cm , where c is the per-unit

⁴In examples such as milking farms where the output is fluid, one unit of output corresponds to all the milk produced by one cow; i.e., quality uncertainty acts at the cow level. We normalize the milk volume produced by each cow to 1.

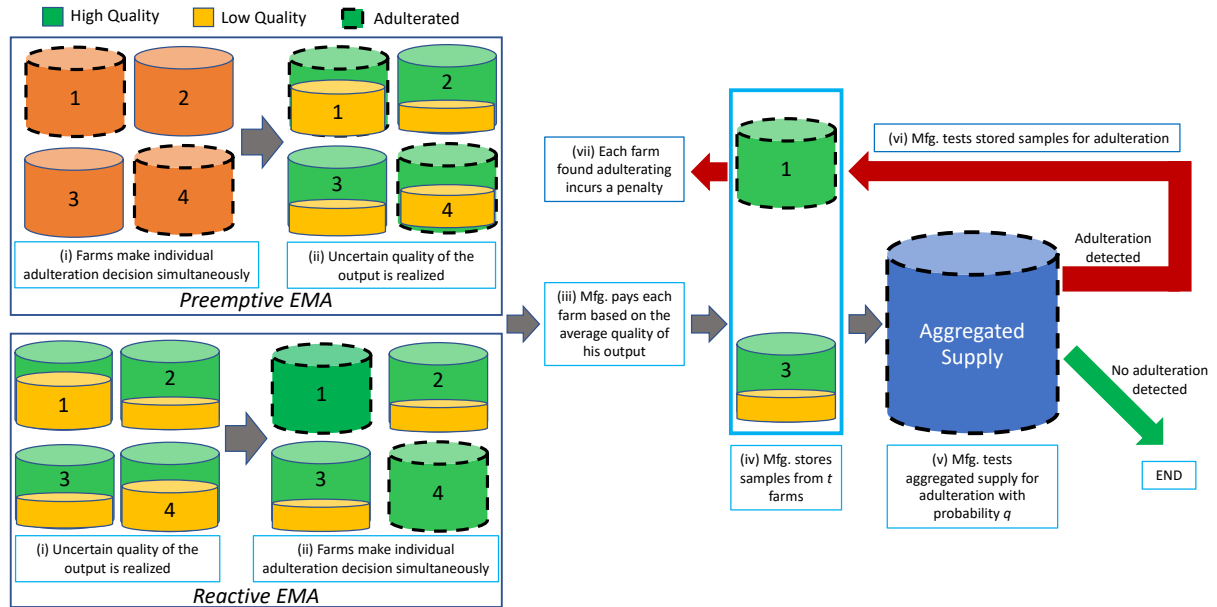


Figure 4.1: Sequence of Events for Preemptive and Reactive EMA

penalty. For reactive EMA, the first two steps in the model dynamics are instead as follows: (i) The uncertain quality of the output is realized. Each farm privately observes the number of low-quality units he produces. (ii) Each farm simultaneously and individually decides whether or not to adulterate all of the realized low-quality units to create fake high-quality ones. The remaining steps proceed exactly the same as in preemptive EMA.⁵

To map steps (i) and (ii) to practice, first consider for preemptive EMA the example of excessive or illegal use of antibiotics in poultry farming. Here, quality refers to the weight and health status of the grown-up chickens at the time of sale as these parameters largely determine the selling price. Step (i) in Figure 4.1 corresponds to a farm adding (or not) antibiotics to the feeds of all chickens, and step (ii) corresponds to some chickens growing healthily while others suffering from disease. For reactive EMA, consider the example of milking farms adding melamine to artificially increase the perceived protein content in milk. Here, quality refers to the compositional characteristics (e.g., protein and fat content) of the milk, which again substantially influence the selling price. Step (i) in Figure 4.1 corresponds to some cows producing milk with good protein content while others producing milk with lower protein level, and step

⁵Mu et al. (2014, 2016) also consider adulteration in milk supply chains with individual and mixed testing. The major differences between our model and theirs are twofold: First, we distinguish between exogenous quality uncertainty and endogenous adulteration decisions, whereas they model adulteration as farms producing (deterministically) low-quality output. Second, we model both perfect and imperfect testing scenarios, whereas they only analyze perfect testing.

(ii) corresponds to the farm adding (or not) melamine to the milk.⁶ Note that the “farm” and “manufacturer” notation in our model represents more generally a player in the upstream (e.g., a farm or a collector) and downstream (e.g., a manufacturer or a wholesaler) of the supply chain.

We note a key difference in the farm’s adulteration decision between preemptive and reactive EMA. In preemptive EMA, the farm chooses *how much* to adulterate, i.e., the amount of adulterants to add, and the adulterants are applied to all units. In reactive EMA, the farm chooses *whether or not* to adulterate; in the case of adulterating, he adulterates all of the realized low-quality units.⁷ In addition, we remark that we do not treat preemptive and reactive EMA as mutually exclusive choices by the farms. There could be situations when farms engage in both types of EMA. We consider the separate analysis of either scenario as a critical first step toward analyzing situations where both types of EMA could be relevant.⁸ When we analyze the manufacturer’s testing capability as a lever to deter EMA in §4.7, we allow the farms to engage in both types of EMA.

We next explain steps (iii)–(vii) in Figure 4.1 in more detail. First, in step (iii), we model quality-based pricing—a very common payment scheme in many agricultural industries (Bennett et al. 2001). In particular, the price for a high-quality (low-quality) unit is r_H (r_L) with $r_H > r_L$. This price difference is an important economic motive for farms to engage in adulteration. Note that the manufacturer does not need to test the quality of every unit to determine the price. Instead, our model captures linear pricing based on average quality.⁹ In the example of preemptive EMA and poultry farming, our model corresponds to a process where all chickens from a farm are weighed and the farm gets paid based on the total weight. Similarly, in the example of reactive EMA and milk farming, a farm typically aggregates the milk from all the

⁶Agricultural studies show that even in the absence of any equipment, milking farmers can predict the key compositional features (e.g., protein and fat content) of raw milk with high accuracy based on a number of features known to the farmers, e.g., the health of the cow, the quality of the feed, climate, the age and breed of the cow (Gale and Hu 2009, Wongpom et al. 2017).

⁷In Appendix C.1.3, we allow farms to adulterate a fraction of their low-quality units under reactive EMA and obtain similar results.

⁸The two types of EMA could be linked in our analysis by allowing p_L to be dependent on the farms’ preemptive EMA decision. Under the reasonable assumption of farmers being short-term oriented (e.g., Antle 1987, Chintapalli and Tang 2017, Hu et al. 2017), their preemptive and reactive EMA decisions are effectively decoupled. That is, when farmers are making preemptive EMA decisions (which would impact p_L), they are not likely to have the foresight to account for every possible realized number of low-quality units in the future and the corresponding reactive EMA decision. As such, the analyses in this chapter can be directly applied to examine farms’ adulteration behaviors and the resulting total EMA risk when both types of EMA occur.

⁹Mathematically, let n_H and n_L denote the number of high-quality and low-quality units from a farm, and recall that m is the total number of units at each farm. The price based on the average quality of the farm’s output can be modeled as $r_H(n_H/m) + r_L(n_L/m)$. This formulation essentially linearizes the two extreme price points onto the average quality level. Therefore, the revenue to the farm from selling all m units is equal to $r_H n_H + r_L n_L$, equivalent to each high-quality (low-quality) unit selling at r_H (r_L).

cows before bringing to the manufacturer. Hence, the manufacturer simply tests the (average) protein and fat content of the mixed milk and pays a price accordingly.

Second, in step (iv), we capture the *traceability* of the supply chain with the number of randomly-chosen farms, $t \in [0, n]$, from which the manufacturer stores samples. In a fully traceable supply chain ($t = n$), samples from all farms are stored and hence, can be tested for adulteration if needed. Conversely, in a partially traceable supply chain ($t < n$), only some of the farms' samples are stored.¹⁰ Therefore, the traceability of the supply chain significantly impacts the manufacturer's ability to identify all of the adulterating farms. Only those farms that are traced could potentially incur a penalty from adulterating. In practice, some dairy (poultry) companies store samples of milk (meat) before aggregating the supply for processing. If a quality problem is detected in the output, the companies test the samples to identify problematic farms (Flynn and Zhao 2014, Zhang and Bhatt 2014).

Third, in steps (v) and (vi), we model the manufacturer's two-step process of testing for adulteration. Specifically, with some probability q (which captures inspection frequency), the manufacturer tests the aggregated supply for adulteration. Only if adulteration is detected in the aggregated supply will the manufacturer then test the stored samples. This two-step process is common in practice due to its cost effectiveness and potential resource constraints of the manufacturer (Draaiyer 2002, Mu et al. 2016). Note that tests for adulteration differ substantially from quality tests for pricing purposes. In particular, quality-based pricing typically considers a small number of quality parameters (e.g., weight of a chicken, or protein and fat content in milk) that can be measured with simple tests in real time. For example, checking the weight of chickens is simple and quick. Similarly, the Gerber butterfat test widely used in the dairy industry involves a simple procedure and takes about 10 minutes to complete (FAO 2009). In contrast, tests for adulteration often require more advanced technologies, take longer time, and are more expensive. For example, the HPLC-UV test for targeted detection of adulterants uses specialized equipment and is labor intensive and costly (Handford et al. 2016). In practice, many of these tests are done in formal testing labs. Thus, to do these tests requires the manufacturer to send samples to a lab and wait for up to a few weeks' time to get the results. However, the manufacturer often needs to pay the farms at the time of sale (or shortly after). Due to these

¹⁰Storage of food/agricultural samples for an extensive period of time is difficult and costly due to perishability and the substantial facility investments needed (relative to the generally low margin of agricultural products). Furthermore, it is challenging (if not impractical) to build full traceability in a dispersed supply chain with a large number of small farms. These features partly explain why the traceability of agricultural supply chains is typically lower than that of pharmaceutical and electronics supply chains.

differences, the predominant majority of food manufacturers do not perform adulteration tests all the time in current practices. Furthermore, the two-step process also captures cases where adulteration is detected more downstream in the supply chain when products are inspected by a third party (e.g., government agencies) or when consumers develop adverse symptoms due to consumption of adulterated products.

An important factor we capture related to testing is the sensitivity of the test in detecting adulteration. We model both perfect and imperfect testing. With perfect testing (e.g., advanced tests targeted for certain antibiotics), the test is very sensitive and can accurately detect whether the residue amount of an adulterant (if any) in the food exceeds the maximum limit allowed by food safety standards. With imperfect testing, the sensitivity of the test depends on the relative amount of adulterants in the total supply chain output. The larger this relative amount, the more likely that adulteration can be detected. We capture this dependency by modeling the detection probability to be linearly increasing in the relative amount of adulterants in the total output.¹¹

Finally in step (vii), we model the penalty for an adulterating farm to be cm . This penalty structure is motivated by current food safety regulations in China regarding highly toxic adulterants (e.g., melamine, malachite green, or other legally controlled/banned compounds; [Handford et al. 2016](#)). For these adulterants, regulations and standards are defined based on whether or not the residue amount exceeds the maximum allowable limit (e.g., [PRC-MIIT 2011](#), [FAO 2017](#), [CFDA 2018](#)). As long as this limit is exceeded, monetary fines (or prison time) are primarily determined based on the sales revenue associated with the products ([PRC-NPC 2015](#)).¹² Furthermore, modeling the penalty to be proportional to the size of the farm can be interpreted as capturing the cost of lost future business if a farm is caught adulterating. This interpretation is particularly relevant in situations where smallholder farmers are rarely fined or legally prosecuted due to political sensitivity (e.g., in India).

Table 4.1 summarizes the key parameters in our model. In the next section, we examine the farms' adulteration behaviors and the resulting EMA risk in four different settings: preemptive or reactive EMA with perfect or imperfect testing. All proofs are deferred to the online appendix. Additional results presented in the online appendix are referenced as O.X.

¹¹In Appendix C.1.4, we analyze settings where we relax the linearity assumption under imperfect testing.

¹²For less toxic adulterants whose harm to human health gradually increases with quantity (e.g., using non-hygienic water to adulterate food; [Handford et al. 2016](#)), it may be reasonable to model the per-unit penalty to increase in the amount of adulterants used. In Appendix C.1.5, we show that all of our results remain the same with this alternative formulation.

Variable	Definition
k	total supply chain output
n	total number of farms in the supply chain
m	Number of units supplied by each farm; $m = k/n$
p_L	Probability that a unit of output is of low quality
r_H	Per-unit price paid by the manufacturer to the farm for high-quality output
r_L	Per-unit price paid by the manufacturer to the farm for low-quality output
t	Number of farms whose samples are stored by the manufacturer; $t \in [0, n]$
q	Probability that the manufacturer tests the aggregated supply for adulteration
c	Per-unit penalty charged to a farm that is found adulterating

Table 4.1: Notations

4.3 When Will Farms Adulterate?

4.3.1 Preemptive EMA

Let $x \in [0, 1]$ be a farm's adulteration decision, where $x = 0$ means the farm does not adulterate his output and $x = 1$ means the farm adulterates with the maximum dosage. In light of our earlier discussions, a decision of $x = 0$ represents the farm using the adulterant (e.g., veterinary drugs) within the legal limit. A decision of $x > 0$ represents using the adulterant beyond the legal limit (as opposed to using a positive quantity per se), and thus, it will cause harm to human health. We assume that adding more adulterants beyond the maximum dosage can no longer decrease the likelihood of producing low-quality output. Therefore, x can be interpreted as the relative quantity of adulterants used by the farm as compared to the maximum dosage. Let $h(x)$ denote the resulting probability that a unit of output is low-quality given x . We assume that $h(x)$ is convex decreasing in x with $h(0) = p_L^{\max}$, $h(1) = p_L^{\min}$, and $p_L^{\max} > p_L^{\min}$. The notations p_L^{\max} and p_L^{\min} represent the largest and smallest probability that a unit of output is low-quality given the farm's adulteration decision. With a large (small) p_L^{\max} , the farms face a higher (lower) risk of producing low-quality units without adulteration. During epidemics, for example, we can expect p_L^{\max} to be large. The convexity of $h(x)$ implies that the effectiveness of the adulteration to reduce p_L is marginally decreasing. In what follows, we characterize the farms' optimal preemptive EMA strategies under perfect and imperfect testing separately.

Perfect testing.

Given a farm's adulteration decision x , the probability that each unit of his output is of low quality is $h(x)$. Hence, the expected total number of low-quality units is $mh(x)$. If the farm

chooses not to adulterate (i.e., $x = 0$), then he earns an expected payoff of $r_H m(1 - p_L^{\max}) + r_L m p_L^{\max}$. If instead the farm chooses to adulterate, then he earns an expected revenue of $r_H m(1 - h(x)) + r_L m h(x)$ but would incur a penalty of cm if he is caught by the manufacturer. We now analyze the probability that an adulterating farm will be caught by the manufacturer.

Recall from §4.2 that the manufacturer employs a two-step testing process. Suppose in the case that the manufacturer tests the aggregated supply (which happens with probability q), she randomly picks one unit of the aggregated supply (e.g., a chicken) to test for adulteration. With perfect testing, the manufacturer detects adulteration as long as any one farm has adulterated. Let n_a be the total number of farms that have adulterated their output. The chance that the manufacturer picks an adulterated unit is thus n_a/n . If the manufacturer detects adulteration in the first step, then she further tests the t samples. Since samples are taken from randomly-chosen farms, the chance that an adulterating farm's sample has been stored is t/n . Taken together, the probability that an adulterating farm will eventually be caught by the manufacturer is thus $q(n_a/n)(t/n)$.

We make two important observations from the analysis thus far. First, since testing is perfect, the farm faces the same level of penalty for any $x > 0$. In addition, his expected revenue without considering the potential penalty is increasing in x . Hence, if the farm decides to adulterate, it is in his best interest to adulterate with the maximum dosage, i.e., choosing $x = 1$. As a result, an adulterating farm's expected payoff is equal to $r_H m(1 - p_L^{\min}) + r_L m p_L^{\min} - q(n_a/n)(t/n)cm$. Second, a farm's adulteration decision depends on how many other farms are also adulterating (i.e., the value of n_a). Therefore, the farms' adulteration decisions impact each other's payoffs, and we solve for a Nash equilibrium (NE) in this static game of complete information (Fudenberg and Tirole 1991, , Chapter I). Theorem 4.3.1 below characterizes the set of Nash equilibria in the game.

Theorem 4.3.1. For preemptive EMA with perfect testing, the total number of adulterating farms in any Nash equilibrium of the game is characterized by

$$n_a^* = \max \left\{ 0, \min \left\{ n, \left\lceil \frac{(r_H - r_L)(p_L^{\max} - p_L^{\min})n^2}{cqt} - 1 \right\rceil \right\} \right\}, \quad (4.1)$$

where $\lceil \cdot \rceil$ denotes the smallest integer greater than the argument. That is, any subset of n_a^* farms adulterating with the maximum dosage while the rest $(n - n_a^*)$ farms not adulterating

constitutes a Nash equilibrium of the game.¹³

The farm's decision on whether or not to adulterate is driven by the tradeoff between the expected revenue gain and the potential penalty from adulteration. By adulterating, the farm increases his expected revenue by $(r_H - r_L)(p_L^{\max} - p_L^{\min})$ per unit of output, while in the meantime facing a penalty of $q(n_a/n)(t/n)c$ per unit. The equilibrium value n_a^* defines the threshold number such that none of the adulterating farms find it profitable to not adulterate, and none of the non-adulterating farms find it profitable to adulterate.

Imperfect testing.

The key difference under imperfect versus perfect testing is that the detection of adulteration depends on the relative amount of adulterants in the output. Here, we model this dependency to be linearly increasing. Index the farms by $i = 1, \dots, n$ and let $x_i \in [0, 1]$ be farm i 's adulteration decision. Due to imperfect testing, the amount of adulterants added (i.e., x_i) affects the chance that an adulterating farm will be caught. As a result, it is not necessarily optimal for the farm to adulterate to the maximum dosage (unlike in perfect testing). We first derive the chance of an adulterating farm eventually being caught by the manufacturer. Similar to §4.3.1, suppose with probability q , the manufacturer picks one unit (e.g., a chicken) from the aggregated supply to test for adulteration. In this case, the probability that the manufacturer detects adulteration in the aggregated supply can be derived as follows:

$$\begin{aligned} \mathbb{P}(\text{detection}) &= \sum_j [\mathbb{P}(\text{detection} | \text{farm } j\text{'s output is picked}) \mathbb{P}(\text{farm } j\text{'s output is picked})] \\ &= \sum_j [x_j(1/n)] = \sum_j x_j/n. \end{aligned}$$

The conditional probability in the bracket is equal to x_j because (i) under preemptive EMA, the added adulterants affect all units of output at a farm, and (ii) the detection sensitivity is linearly increasing in the relative amount of adulterants added. If the manufacturer detects adulteration in the aggregated supply, then she will test the t samples. An adulterating farm's sample will be stored with probability t/n . Finally, the chance that the manufacturer detects adulteration in the sample again depends on the amount of adulterants added, x_i . To summarize, the chance that an adulterating farm i will eventually be caught by the manufacturer is equal to $q(\sum_j x_j/n)(t/n)x_i$. Therefore, farm i 's expected payoff given his adulteration decision x_i can

¹³We assume that when a farm is indifferent between adulterating or not, he chooses not to adulterate.

be characterized as follows.

$$\pi^{PV}(x_i, x_{-i}) = r_H m(1 - h(x_i)) + r_L m h(x_i) - q \left(\frac{x_i + \sum_{-i} x_{-i}}{n} \right) (t/n) x_i c m, \quad (4.2)$$

where $-i$ denotes all farms other than i . Since all the farms are homogeneous, we focus on analyzing symmetric NE of the game; i.e., NE in which the equilibrium strategy x_i^{PV*} has the same structure for all i . This approach is common in game-theoretic analysis with homogeneous players (e.g., Che 1993, Lee et al. 1997, Wang and Zender 2002, Golosov et al. 2014). The next theorem characterizes the unique symmetric NE of the game. We drop the subscript i to simplify notation.

Theorem 4.3.2. For preemptive EMA with imperfect testing, there exists a unique symmetric NE in which x^{PV*} is determined as follows.

- (a) If $c < \frac{-h'(1)(r_H - r_L)}{q(t/n)((n+1)/n)}$, then all farms adulterate to the maximum level; i.e., $x^{PV*} = 1$.
- (b) If $c \geq \frac{-h'(1)(r_H - r_L)}{q(t/n)((n+1)/n)}$, then the farms adulterate to some extent; i.e., $x^{PV*} \in (0, 1)$ and is the solution to the following equation: $-h'(x)/x = q(t/n)((n+1)/n)c/(r_H - r_L)$.

When the per-unit expected penalty is sufficiently small compared to the per-unit expected revenue gain, all farms adulterate to the maximum level as under perfect testing. However, when the per-unit expected penalty is large, farms in equilibrium adulterate to some extent but not to the maximum level (Theorem 4.3.2(b)). This is because both the expected revenue gain and the chance of being caught increase as farms increase the amount of adulterants.

4.3.2 Reactive EMA

Recall from §4.2 that under reactive EMA, the farms decide whether or not to adulterate his low-quality units to create fake high-quality ones *after* the uncertain quality of his output is realized. In addition, the total number of low-quality units at a farm follows a binomial distribution with parameters m (number of units) and p_L (probability of each unit being low-quality). This distribution can be well approximated by a normal distribution with mean mp_L and variance $mp_L(1 - p_L)$ if $mp_L \geq 5$ and $m(1 - p_L) \geq 5$ (Ross 2005, Stirzaker 1999, Clemens and Inderfurth 2015). To ensure tractability, we perform our analysis with this normal approximation. Let $f(x, m, p)$ denote the probability density function (PDF) of a normal distribution with mean

mp and variance $mp(1-p)$ evaluated at x . We now characterize the farms' optimal reactive EMA strategies under perfect and imperfect testing separately.

Perfect testing.

Let n_L be the realized number of low-quality units at a farm. If the farm does not adulterate, then he earns $r_H(m - n_L) + r_L n_L$ based on the average quality of his output. If the farm decides to adulterate, then he earns a revenue of $r_H m$ but would incur a penalty of cm if he is caught by the manufacturer. Note that if a farm adulterates, then the adulterants will be present in his output, the manufacturer's aggregated supply, and the stored sample. For example, if a dairy farm adds melamine to his raw milk, then melamine will be present in the milk from this farm, the aggregated pool of milk at the manufacturer, and any sample stored from this farm. Under perfect testing, the manufacturer detects adulteration in the aggregated supply as long as any farm has adulterated and the manufacturer tests the aggregated supply (the latter occurs with probability q). If further the sample of an adulterating farm is stored (this occurs with probability t/n), then the farm will be caught and incur the penalty. Hence, the probability of a farm's adulteration being detected is $q(t/n)$, and the expected payoff of an adulterating farm is equal to $r_H m - q(t/n)cm$. Observe that the farms' adulteration decisions do not affect each other's payoffs and can be solved independently. A farm chooses whether or not to adulterate to maximize his expected payoff. The following theorem describes the optimal adulteration strategy for a farm with n_L units of low-quality output.

Theorem 4.3.3. For reactive EMA with perfect testing, the optimal adulteration strategy for a farm is a threshold strategy: He does not adulterate if $n_L \in [0, \beta^{RP}]$, and he adulterates if $n_L \in (\beta^{RP}, m]$. In addition, β^{RP} is decreasing in the price difference between high- and low-quality output ($r_H - r_L$) and increasing in testing frequency (q), the number of stored samples (t), and per-unit penalty (c).

The threshold strategy in Theorem 4.3.3 follows from a tradeoff between revenue gain and the potential penalty from adulteration. Since the revenue gain increases with n_L whereas the expected penalty is independent of n_L , the farm finds it beneficial to adulterate when n_L is sufficiently large. The more frequent the manufacturer tests the aggregated supply or the better traceability in the supply chain (higher q and t), the higher chance that an adulterating farm will be caught, and hence, the less likely a farm is to adulterate. Similarly, a smaller per-unit penalty (c) and a greater price difference between high- and low-quality output ($r_H - r_L$)

make the penalty from adulteration less severe and the gain more attractive, thus motivating a farm to adulterate more often. This last point is in line with qualitative evidence that many adulteration incidents occurred when there was external price pressure for low-quality products. For example, the Indian government found that milk adulteration was significantly more severe in Maharashtra where low-fat milk was rejected at collection centers versus in Gujarat where it was accepted at a lower price (Deshmukh 2011). We warrant that simply reducing the price difference is not likely a viable way to deter adulteration, especially if farms rely on the price premium for higher quality to truly invest in quality. Instead, we advocate that fairer risk sharing between the manufacturer and the farms is essential (e.g., offering protective prices in light of quality uncertainty common in agricultural production).

Imperfect testing.

In this case, the chance that the manufacturer detects adulteration when testing the aggregated supply is modeled as linearly increasing in the fraction of adulterated units within the total output. Therefore, the farms' (simultaneous) adulteration decisions affect each other's payoffs through this detection probability. Since the realized number of low-quality units at a farm is the farm's private information, we model the farms' strategic interactions as a static game of incomplete information and solve for the Bayesian Nash equilibrium (BNE) of the game (Fudenberg and Tirole 1991, , Chapter III).

Formally, let $n_{L,i}$ denote the realized number of low-quality units at farm i ($i = 1, \dots, n$). Let $a_i(n_{L,i}) : \{1, \dots, m\} \rightarrow \{0, 1\}$ be farm i 's adulteration strategy, where a value of 1 (0) means adulterating (not adulterating). That is, farm i 's adulteration strategy specifies for each realized number of low-quality units, whether or not the farm adulterates these low-quality output. From farm i 's perspective, given all other farms' adulteration strategies $a_{-i}(n_{L,-i})$, the expected total number of adulterated units from these farms is equal to $\mathbb{E}_{n_{L,-i}} [\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]$. The expectation is taken on the (uncertain) number of low-quality units at the other farms, which is not observable by farm i .

We next derive the chance of farm i being caught if he adulterates. First, the fraction of adulterated units within the total output (hence the chance that the manufacturer detects adulteration when testing the aggregated supply) is equal to $\frac{n_{L,i} + \mathbb{E}_{n_{L,-i}} [\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]}{k}$. Conditional on adulteration being detected in the aggregated supply, if farm i 's sample is stored (which occurs with probability t/n), then farm i will be caught with probability $n_{L,i}/m$. Since

the manufacturer tests the aggregated supply with probability q , the ultimate probability for farm i to be caught if he adulterates is equal to

$$\gamma_i(n_{L,i}, a_{-i}(n_{L,-i})) \equiv q \left(\frac{t}{n} \right) \binom{n_{L,i}}{m} \left(\frac{n_{L,i} + \mathbb{E}_{n_{L,-i}} \left[\frac{\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right]}{k} \right). \quad (4.3)$$

Thus, the final expected payoff if farm i chooses to adulterate is equal to $r_H m - \gamma_i(n_{L,i}, a_{-i}(n_{L,-i})) c m$. Conversely, the final payoff if farm i chooses not to adulterate is equal to $r_H(m - n_{L,i}) + r_L n_{L,i}$. Farm i decides whether or not to adulterate depending on which action yields a higher final payoff.

We again focus on analyzing symmetric BNE of this game; i.e., BNE in which the strategy $a_i^*(n_{L,i})$ has the same structure for all i . The next theorem characterizes the unique BNE of this game. We drop the subscript i to simplify notation.

Theorem 4.3.4. For reactive EMA with imperfect testing, there exists a unique symmetric BNE of the game in which a farm's adulteration strategy is a threshold strategy: $a^*(n_L) = 1$ if $n_L \in [0, \beta^{RV})$ and $a^*(n_L) = 0$ if $n_L \in [\beta^{RV}, m]$. The threshold β^{RV} is unique and determined as follows:

- (a) If $n \leq 2(1 - p_L) \left/ \left(\sqrt{p_L^2 + \frac{4(1 - p_L)(r_H - r_L)}{cqt}} - p_L \right) \right.$, then $\beta^{RV} \in (0, m)$ and is the solution to the following equation: $\beta = \frac{nk(r_H - r_L)}{cqt} - (n - 1) \int_0^\beta x f(x, \frac{k}{n}, p_L) dx$.
- (b) If $n > 2(1 - p_L) \left/ \left(\sqrt{p_L^2 + \frac{4(1 - p_L)(r_H - r_L)}{cqt}} - p_L \right) \right.$, then $\beta^{RV} = m$.

In addition, β^{RV} is increasing in the price difference between high- and low-quality output ($r_H - r_L$) and decreasing in testing frequency (q), the number of stored samples (t), and per-unit penalty (c).

Figure 4.2 contrasts the farms' optimal reactive EMA strategies under perfect testing (Theorem 4.3.3) versus imperfect testing (Theorem 4.3.4). Under perfect testing, farms adulterate when the realized number of low-quality units, n_L , is *greater* than the threshold β^{RP} . In contrast, under imperfect testing, farms adulterate when n_L is *smaller* than the threshold β^{RV} . This latter result is because under imperfect testing, the expected penalty from adulterating increases with n_L faster than the revenue gain does. Observe from Equation (4.3) that the chance of farm i being caught adulterating increases with $n_{L,i}$ quadratically, whereas the revenue gain, $(r_H - r_L)n_{L,i}$, increases with $n_{L,i}$ linearly. Therefore, farms find it more beneficial

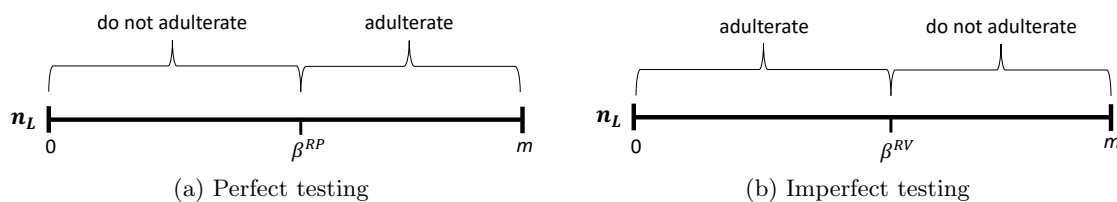


Figure 4.2: Reactive EMA: Farms' Optimal Adulteration Strategies under Perfect and Imperfect Testing

to adulterate when n_L is low. More intuitively, the key driver of this pattern is the farms' "free-riding" behavior; that is, when a farm believes that there is a sufficiently large number of high-quality units in the supply chain that can hide his adulteration (i.e., when n_L and the belief of other farms' n_L are small), the farm is more likely to adulterate. Evidence of such free-riding behavior exists in practice. For example, [Gadzikwa et al. \(2007\)](#) report that organic producers are more likely to adulterate (e.g., using pesticides) and fake organic products when the fake quantity is small relative to the total quantity. In an extension (§C.1.4), we allow the sensitivity of imperfect testing to increase faster than the linear model, and hence, become closer to the sensitivity of perfect testing. We demonstrate that the farms' equilibrium reactive EMA strategy becomes a combination of the two structures in Figure 4.2; i.e., farms adulterate when n_L is either small or large but do not adulterate in between.

Due to the aforementioned contrasting pattern under perfect versus imperfect testing, the effects of the price difference, testing frequency, traceability, and per-unit penalty on β^{RV} are opposite to those for β^{RP} (see Theorem 4.3.3). Comparing these two thresholds, we obtain the following result.

Proposition 4.3.5. (i) If $\beta^{RP} \in (0, m)$, then $\beta^{RV} = m$. (ii) If $\beta^{RV} \in (0, m)$, then $\beta^{RP} = m$.

Proposition 4.3.5 shows that given the same model parameters, whenever farms would possibly adulterate under perfect testing (i.e., $\beta^{RP} \in (0, m)$), then they would always adulterate under imperfect testing. Similarly, whenever farms would possibly not adulterate under imperfect testing (i.e., $\beta^{RV} \in (0, m)$), then they would never adulterate under perfect testing. Consequently, the resulting reactive EMA risk is always higher under imperfect than perfect testing.

Corollary 1. *The reactive EMA risk in the supply chain is always higher under imperfect testing than under perfect testing.*

4.4 The Effect of Supply Chain Dispersion on EMA Risk

We now analyze how supply chain dispersion affects preemptive and reactive EMA risks in the supply chain. We measure preemptive EMA risk by the fraction of adulterating farms, n_a^*/n , under perfect testing, and the farms' adulteration decisions, x^{PV^*} , under imperfect testing. A larger (smaller) value of these terms indicates a larger (smaller) amount of adulterants being used in the total output, and hence, a higher (lower) risk of preemptive EMA. Similarly, we measure reactive EMA risk in two ways: P_n denotes the probability of an individual farm adulterating in a supply chain with n farms, and E_n denotes the expected total amount of adulterated output in the supply chain. Under perfect testing, $P_n = \text{Prob}(n_L > \beta^{RP})$, and $E_n = n \int_{\beta^{RP}}^{k/n} x f(x, k/n, p_L) dx$, where β^{RP} is defined in Theorem 4.3.3. We can similarly define P_n and E_n under imperfect testing. For the analysis in this section, we treat n to be a continuous variable to ensure tractability. Our first result shows how supply chain dispersion (captured by n) impacts preemptive and reactive EMA risk under perfect and imperfect testing.

Proposition 4.4.1. Under perfect testing: When $t < n$ ($t = n$), (i) preemptive EMA risk, measured by n_a^*/n , is increasing (constant) in n ; (ii) reactive EMA risk, measured by P_n and E_n , is increasing in n if $qc \geq \frac{p_L(r_H - r_L)}{3(t/n)}$ ($qc \geq p_L(r_H - r_L)$).

Proposition 4.4.2. Under imperfect testing: (i) Preemptive EMA risk, measured by x^{PV^*} , is increasing in n ; (ii) reactive EMA risk, measured by P_n and E_n , is increasing in n .

Proposition 4.4.1 considers perfect testing and characterizes conditions under which supply chain dispersion increases EMA risk (or not). First, greater dispersion leads to higher preemptive EMA risk under perfect testing only in partially traceable supply chains. That is, this higher risk is solely due to decreased traceability in a more dispersed supply chain (observe that n_a^*/n is constant in n if $t = n$). In Appendix C.1.1, we show that if farms are averse to quality uncertainty, then an additional reason for this higher risk is that a smaller farm faces larger uncertainty in the quality of his output. Second, regardless of traceability, dispersion increases reactive EMA risk under perfect testing if the per-unit expected penalty for adulteration (qc) is high. This result can be explained by the following dynamics. Imagine that the manufacturer diversifies her supply base by procuring one fewer unit from each of the existing farms and procuring the resulting gap from a new farm. When the per-unit expected penalty is *high*, these units originally procured from each existing farm would not be adulterated in expectation. However, when they are instead procured from a new farm, some of them would be adulterated

at least sometimes. Therefore, reactive EMA risk *increases* in a more dispersed supply chain. Conversely, if the per-unit expected penalty is *low*, these units when procured from the original farms would all be adulterated in expectation, while only some of them would be adulterated when procured from the new farm. Hence, reactive EMA risk *decreases* in a more dispersed supply chain.¹⁴

Proposition 4.4.2 considers imperfect testing and shows that both preemptive and reactive EMA risk *always* increase in a more dispersed supply chain, regardless of traceability. With a large number of farms in the supply chain, each farm produces only a small fraction of the total output. As a result, each farm's adulteration has a limited impact on the total amount of adulterants and adulterated output in the aggregated supply, and hence, on the likelihood that the manufacturer would detect adulteration. Therefore, farms feel less risky to adulterate in a more dispersed supply chain, resulting in higher EMA risk. In sharp contrast, if the supply chain consists of only two farms each producing half of the total output, then the adulteration decision of each farm would have a much more prominent impact on the detection probability. Thus, these larger farms would be more cautious in their adulteration decisions.

Table 4.2 summarizes the key insights from Proposition ???. We observe that the effect of supply chain dispersion on EMA risk highly depends on the manufacturer's testing capability. When test is perfect, greater dispersion does not always result in higher EMA risk. However, when test is imperfect, increased dispersion is always harmful. Note that even if the manufacturer tests the aggregated supply with certainty (i.e., $q = 1$), these results remain. The insights from Table 4.2 thus highlight the limitation of relying on product inspection to deter EMA, particularly if the farming supply chain is highly dispersed and testing methods are imperfect (as is the case in most developing countries with a large variety of possibly unknown adulterants). Hence, to more proactively fight EMA, our results call for the need to incorporate a systemic, supply chain perspective to complement an inspection-centered approach adopted in current practices.

4.5 The Effect of Quality Uncertainty on EMA Risk

The following proposition demonstrates how the likelihood of producing low-quality output (p_L) affects both preemptive and reactive EMA risk.

¹⁴We show analytically that when the condition in Proposition 4.4.1 part (ii) is not satisfied, P_n decreases with n (see the proof of Proposition 4.4.1). For E_n , we observe from numerical simulation that it decreases with n when qc is sufficiently low.

	Perfect testing	Imperfect testing
Preemptive EMA	Risk increases (is constant) under partial (full) traceability	Risk increases
Reactive EMA	Risk increases (decreases) if penalty is high (low)	Risk increases

Table 4.2: Effect of Increased Supply Chain Dispersion on EMA Risk

Proposition 4.5.1. (i) For preemptive EMA, both n_a^*/n and x^{PV*} are increasing in p_L^{\max} .

(ii) For reactive EMA under perfect testing, both P_n and E_n are increasing in p_L .

(iii) For reactive EMA under imperfect testing, P_n is decreasing in p_L .

Proposition 4.5.1 parts (i) and (ii) show that as the likelihood of producing low-quality output (p_L^{\max} or p_L) at a farm increases, the supply chain would face a higher risk of EMA for preemptive EMA under either testing sensitivity and reactive EMA under perfect testing. Specifically, a higher p_L^{\max} means there is a greater chance of producing low-quality units, and hence, farms are more motivated to engage in preemptive EMA to reduce that chance. Similarly, a higher p_L means the realized number of low-quality units, n_L , is more likely to be large, as n_L follows a binomial distribution with parameters m and p_L . Therefore, farms are more likely to engage in reactive EMA (recall from Theorem 4.3.3 that under perfect testing, a farm adulterates if $n_L > \beta^{RP}$).

In sharp contrast, Proposition 4.5.1 part (iii) shows that under imperfect testing, the probability that an individual farm engages in reactive EMA *decreases* as p_L increases. This result is related to the farms' "free-riding" behavior discussed in §4.3.2. In particular, when p_L is low, an individual farm expects that the other farms would have many high-quality units. As a result, even if this farm chooses to adulterate, the fraction of adulterated units in the aggregated supply can be sufficiently low that the manufacturer would not be able to detect adulteration. Therefore, farms feel "safe" to adulterate and are more likely to do so. Finally, we observe that under reactive EMA with imperfect testing, the effect of quality uncertainty on E_n , the expected total amount of adulterated output in the supply chain, is nonmonotone. Figure 4.3 illustrates the general pattern of this effect based on extensive numerical simulations. As p_L decreases (i.e., as overall quality becomes better), E_n first increases and then decreases. Furthermore, the decreasing region predominantly corresponds to situations when all farms adulterate all the time in equilibrium (which occurs when p_L is low). In this case, E_n is equal to the expected total number of low-quality units, and hence, it decreases as p_L decreases. These observations

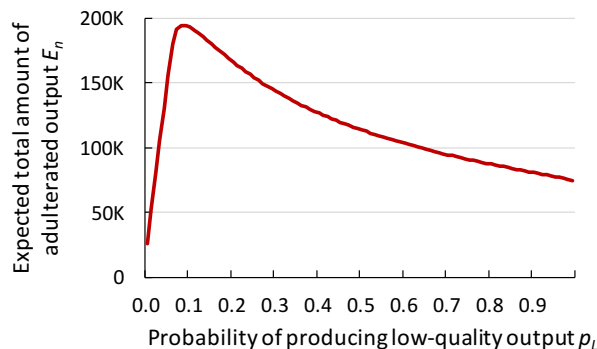


Figure 4.3: Effect of Quality Uncertainty (p_L) on the Expected Total Amount of Adulterated Output (E_n)

	Perfect testing	Imperfect testing
Preemptive EMA	Risk increases	Risk increases
Reactive EMA	Risk increases	Probability of adulteration decreases, while expected total amount of adulterated output first increases then decreases

Table 4.3: Effect of an Increased Probability of Producing Low-Quality Output (p_L^{\max} or p_L) on EMA Risk

highlight that, if the current quality level is bad (i.e., when p_L is high), then a substantial quality investment (to decrease p_L significantly) is necessary to reduce reactive EMA risk with respect to E_n under imperfect testing. Nevertheless, if the focus is to mitigate P_n (e.g., when highly toxic adulterants are concerned), then investing in quality alone is not effective.

Investing in quality to reduce the probability of producing low-quality output is generally believed to be beneficial. Our analysis demonstrates that for reactive EMA with imperfect testing, this strategy may backfire if it is not accompanied by also improving the supply chain’s capability to detect adulteration. This is because a lower chance of producing low-quality output can inadvertently motivate some parties in the supply chain to endogenously adulterate their output and create fake high-quality units. Without the capability to differentiate fake high-quality units from truly high-quality ones, consumers could suffer from consuming adulterated products. An example in China’s dairy industry illustrates this point. After the 2008 melamine-tainted infant formula scandal, many dairy companies made substantial quality investments in their upstream supply chains. However, without similar improvement in testing methods, adulteration in raw milk escalated again in 2011 when a new type of protein-enhancing yet toxic substance was added by milk farms (Handford et al. 2016). We summarize the effect of quality uncertainty on EMA risk for the different modeling scenarios in Table 4.3.

	n	m	r_H	r_L	c	q	t	p_L^{\max}	p_L^{\min}
Guangdong	8	84,000	19.07	0	19.07	{0.1, ..., 0.9}	{1, ..., 8}	{0.2, ..., 0.9}	0.1
Shandong	13	67,000	19.07	0	19.07	{0.1, ..., 0.9}	{1, ..., 13}	{0.2, ..., 0.9}	0.1

Table 4.4: Parameter Values for Guangdong and Shandong Provinces

4.6 Relating Model Predictions to Empirical Observations

In this section, we calibrate our model parameters with field data to the extent possible to examine how well our models' predictions align with empirical evidence in two prominent EMA scenarios.

4.6.1 Preemptive EMA: Misuse of Antibiotics in Poultry Farming in China

As in many developing countries, China's poultry industry is subject to misuse of antibiotics, antivirals, and herbal medicines at poultry farms, especially since multiple outbreaks of avian flu (CCTV 2012, Huang et al. 2017b). We collected farming supply chain data published by China's General Administration of Quality Supervision, Inspection, and Quarantine (AQSIQ), a government agency responsible for entry-exit commodity inspection, certification, accreditation, and import-export food safety. By utilizing this data and various market data, we calibrate our model parameters and predict the risk levels of poultry manufacturers in two leading provinces of poultry production, Shandong and Guangdong, which account for 15% and 8% of the total production (Inouye 2017).

Table 4.4 summarizes the parameter values used in our analysis. The values of n and m are derived by averaging the number of farms and the size of farms (in the number of chickens produced annually) supplying to different poultry manufacturers in each province (18 manufacturers in Guangdong and 34 in Shandong). One immediate observation is that poultry supply chains in Shandong are more dispersed than those in Guangdong, as seen by a larger number of farms supplying to an average manufacturer (n) and the smaller size of an average farm (m). The value of r_H corresponds to the selling price of broiler chickens (in RMB per kilogram) in 2016 (Inouye 2017). Low-quality chickens mean sick chickens which cannot be sold; thus, $r_L = 0$. According to China's food safety law (PRC-NPC 2015), financial penalties on adulterating firms are determined based on the sales value of the products, with a larger marginal increase in penalty at a higher level of product value. We capture this structure by

using $r_H m^2$ as the penalty for a farm of size m .¹⁵ For the manufacturers' testing frequency (q) and the traceability of their supply chains (t), we consider a wide range of values. We use $h(x) = (p_L^{\max} - p_L^{\min})x^2 - 2(p_L^{\max} - p_L^{\min})x + p_L^{\max}$ to ensure that $h(0) = p_L^{\max}$, $h(1) = p_L^{\min}$, and $h(x)$ is convex decreasing. We fix $p_L^{\min} = 0.1$ and vary p_L^{\max} from 0.2 to 0.9 to capture different levels of quality uncertainty in the farms' output. Since poultry farms use a large variety of drugs in their practices, we consider the imperfect testing scenario. Given these parameter values, we calculate the risk levels of an average manufacturer in these two provinces with 936 different parameter combinations.

Our results show that an average poultry manufacturer in Shandong always faces higher preemptive EMA risk than an average manufacturer in Guangdong does. Averaging over all parameter instances, the risk in Shandong is twice of that in Guangdong. This prediction is consistent with empirical evidence; more poultry manufacturers in Shandong have been found to be involved in EMA incidents than those in Guangdong (Huang et al. 2017b). Among the manufacturers in our data, 4 out of 34 (11.8%) Shandong companies and 1 out of 18 (5.6%) Guangdong companies were caught in EMA incidents. Figures 4.4a and 4.4b further illustrate how preemptive EMA risk changes with p_L^{\max} and t in these two provinces. We observe that adulteration increases at a faster rate for Shandong than for Guangdong as p_L^{\max} increases or t decreases. Therefore, manufacturers in Shandong are at a greater risk of increased adulteration due to quality uncertainty, and they can gain more benefit in risk mitigation by improving the traceability of their supply chains.

Lastly in Figure 4.4c, we illustrate how preemptive EMA risk changes with supply chain dispersion (n) for different testing frequencies (q). The values of n are taken from our data. We observe that as the supply chain gets more dispersed, increasing q becomes much less effective in reducing preemptive EMA risk (e.g., risk > 0.6 even if $q = 0.9$ when $n > 20$). Note that we have kept the fraction of traced farms (t/n) constant as we increase n in Figure 4.4c. The reason behind this observation is the following. When the supply chain is composed of a few large farms, each farm's adulteration has a big impact on the total amount of adulterants in the aggregated supply (and hence, the detection probability). Therefore, each farm is cautious about adulterating and sensitive to the testing frequency. However, when the supply chain is composed of many small farms, each farm produces only a small quantity, which, even adulterated, would not increase the detection probability noticeably. Hence, these small farms are not as concerned

¹⁵We verify that all of our results in §4.3 continue to hold with a convex (in m) penalty function (Proposition C.1.12 in Appendix C.1.5).

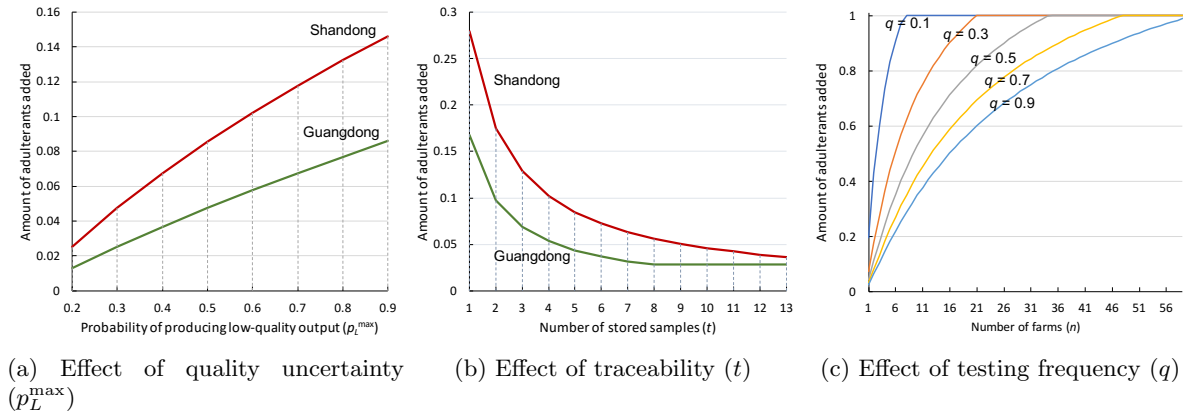


Figure 4.4: Effect of Traceability (t), Quality Uncertainty (p_L^{\max}), and Testing Frequency (q) on Pre-emptive EMA Risk

Note: In Figures 4.4a and 4.4b, we use the average values of n and m presented in Table 4.4 for the two provinces. In Figure 4.4c, we use the whole range of n values in the data and do not distinguish between the two provinces.

about the aggregated supply being tested frequently. As a result, increasing testing frequency has a limited effect on their adulteration behavior, and thus, the resulting preemptive EMA risk in the supply chain.

4.6.2 Reactive EMA: Melamine-Tainted Infant Formula Scandal

In the second empirical case, we examine the risk of two major Chinese dairy companies, Sanlu Group and Bright Dairy, operating at the time of the melamine-tainted infant formula scandal. Regarding Sanlu, the company sourced 2.6 million liters of raw milk from 52,000 small farms primarily through middlemen with little traceability (Chen et al. 2014). These small farms often produced low-quality milk due to poor diets and disease control of the cows (Gale and Hu 2009). In addition, leveraging an exemption of certificate verification from the Chinese government, Sanlu rarely tested the raw milk before production (Flynn and Zhao 2014). All of these data suggest that p_L is high whereas t and q are low in Sanlu's supply chain. In sharp contrast, Bright Dairy sourced 1.625 million liters of raw milk from 2,335 corporate-owned or cooperative farms (Flynn and Zhao 2014). The company stored samples from each cow to ensure full traceability. In addition, it also made significant investments in feed and animal health in the farms and frequently conducted quality tests to uphold its high quality standards. Therefore, in Bright Dairy's supply chain, p_L is low while both t and q are high.

Based on the above discussion, we construct the parameter set as summarized in Table 4.5 for our analysis. In 2008, the average price of raw milk was 6.8 RMB per liter (Gale and Arnade

	n	m	p_L	t	q	r_L	r_H	c
Sanlu Group	52,000	50	{0.5, ..., 0.9}	{520, ..., 2080}	{0.1, ..., 0.5}	6.8	{1.05 r_L , ..., 1.25 r_L }	r_H
Bright Dairy	2,335	700	{0.1, ..., 0.5}	{520, ..., 2080}	{0.5, ..., 0.9}	6.8	{1.05 r_L , ..., 1.25 r_L }	r_H

Table 4.5: Parameter Values for Sanlu Group and Bright Dairy

2015). We use this price as r_L and vary the premium for high-quality milk to be 5% to 25% above r_L . We again use $r_H m^2$ as the penalty function, and consider imperfect testing. Despite qualitative evidence that traceability in Sanlu’s supply chain is almost absent, we allow for some traceability for Sanlu in our analysis. Hence, we likely underestimate its risk level. We analyze a total of 500 parameter combinations for each company and observe a stark contrast in the reactive EMA risk faced by the two companies. On average, the probability of an individual farm adulterating (P_n) is 0.617 for Sanlu versus 0.003 for Bright Dairy. Similarly, the expected fraction of adulterated supply in the total output (E_n/k) is 0.398 for Sanlu versus 0.0005 for Bright Dairy. These predictions match empirical evidence. In particular, Sanlu’s products were the most heavily adulterated in the scandal, whereas Bright Dairy passed all quality inspections by the authorities during the crackdown (Chen et al. 2014).

Figures 4.5a and 4.5b show, for each company, how P_n changes as t and r_H change. While Sanlu’s risk increases quickly as t decreases or r_H increases, Bright Dairy’s risk is almost 0 in all instances. Despite the lack of perfect testing for melamine (at that time), the less dispersed supply chain and better quality assurance established in Bright Dairy’s supply chain have acted as important levers to protect the company from reactive EMA risk. Lastly in Figure 4.5c, we observe a very similar pattern as in Figure 4.4c. That is, increasing testing frequency q has a limited effect on reducing reactive EMA risk when the supply chain is highly dispersed.

4.7 Analyzing Managerial Levers to Mitigate EMA Risk

In this section, we take the manufacturer’s perspective and examine a few managerial levers that could help to mitigate EMA risk in the supply chain. We first consider increasing supply chain traceability t and testing frequency q (Appendix C.2). Since strengthening these levers is costly, we develop an optimization model where the manufacturer’s objective is to minimize total investment costs while satisfying a constraint that the resulting EMA risk in the supply chain cannot exceed a certain level. This modeling approach is common in the risk management literature (e.g., Federgruen and Yang 2008, Meena et al. 2011, Federgruen et al. 2015). We

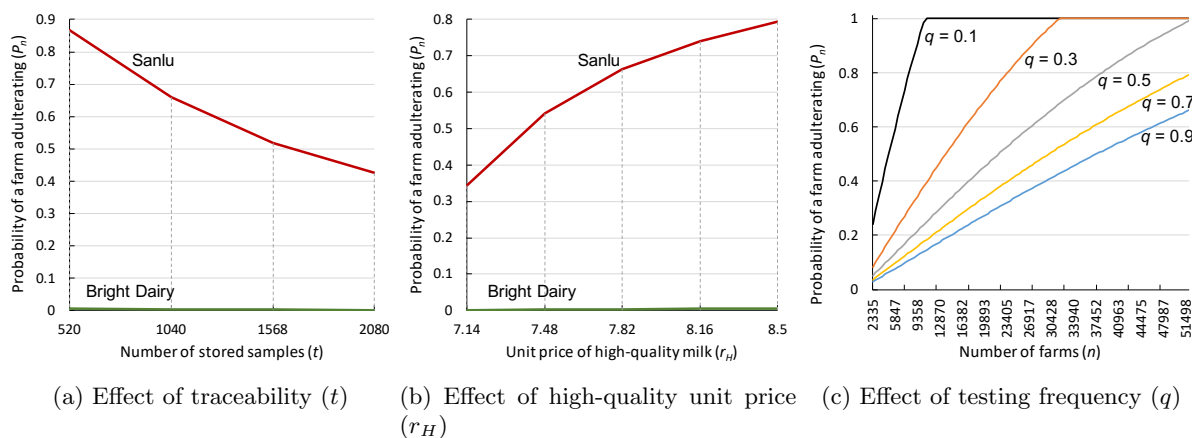


Figure 4.5: Effect of Traceability (t), High-Quality Unit Price (r_H), and Testing Frequency (q) on Reactive EMA Risk

highlight two results. First, whenever a feasible solution exists (i.e., the risk constraint can be satisfied with full traceability and constant testing), the manufacturer optimally utilizes the two levers by prioritizing the more cost-effective one (Theorem C.2.1). Second, given a desirable risk constraint, it is always more difficult for a manufacturer with a more dispersed supply chain to satisfy the constraint. Conditional on being able to satisfy the risk constraint, it is always more costly for a manufacturer with a more dispersed supply chain to do so (Proposition C.2.2). Thus, higher supply chain dispersion results in a greater challenge for a manufacturer to manage and mitigate EMA risk, from both feasibility and financial standpoints. This result aligns with our earlier discussion regarding the adverse impact of supply chain dispersion on EMA risk.

We next consider the lever of testing capability. In particular, the manufacturer can choose to invest in perfect testing for either preemptive or reactive EMA or both. We consider the manufacturer's decision for the two types of EMA separately because in practice, companies contracting with testing labs (for perfect testing) typically can choose what adulterants to test and are priced accordingly. Our focus is to examine how the total EMA risk in the supply chain, accounting for both preemptive and reactive EMA, is affected by the testing capabilities of the manufacturer. To do so, we analyze a setting in which the manufacturer decides whether or not to adopt perfect testing for either type of EMA, and the farms strategically respond by making preemptive and reactive EMA decisions (Appendix C.3). The farms' reactive EMA decisions are conditioned upon their preemptive EMA decisions because the latter impact the likelihood of producing low-quality output, p_L , at the farms. Our main results are twofold. First, by Proposition 4.3.5, perfect testing always decreases reactive EMA risk as compared to imperfect

testing. Thus, if the risk reduction outweighs the investment cost, then the manufacturer should invest in perfect testing to deter reactive EMA. In sharp contrast, perfect testing may lead to a higher preemptive EMA risk than imperfect testing, when the per-unit penalty c is not too low nor high enough.¹⁶ This counterintuitive result can be explained as follows. Recall from Theorems 4.3.1 and 4.3.2 that for preemptive EMA under perfect testing, adulterating farms all adulterate with the maximum dosage, whereas under imperfect testing, they adulterate at a lower level that balances revenue gain and expected penalty. When the penalty is not high enough, many farms choose to adulterate with the maximum dosage under perfect testing. Collectively, they result in a larger amount of adulterants in the total supply chain output than under imperfect testing.

4.8 Conclusions

In this chapter, we develop a new set of analytical models to investigate how exogenous quality uncertainty, supply chain dispersion, traceability, and testing sensitivity (with regard to detecting adulteration) jointly impact farms' strategic adulteration behavior in a farming supply chain consisting of a distributed network of farms. We focus on economically motivated adulteration and consider two distinct scenarios: "preemptive EMA" in which adulteration occurs *before* the uncertain quality of a farm's output is realized with the primary goal of reducing the probability of producing low-quality output; "reactive EMA" in which adulteration occurs *after* the uncertain quality of a farm's output is realized with the primary goal of increasing the perceived quality of the output and creating fake high-quality units. We fully characterize the farms' equilibrium adulteration strategies for both scenarios. We show how these strategies are impacted by quality uncertainty, the level of dispersion and traceability in the supply chain, and the sensitivity of the manufacturer's test for adulteration. We calibrate our model based on real cases and field data and demonstrate that the models' predictions are in line with empirical observations. Furthermore, we examine a number of model extensions to confirm that our main conclusions are robust to a few modeling assumptions. Finally, we analyze several managerial levers available to the manufacturer to mitigate EMA risk in the supply chain.

Our analysis offers important and unique insights for policy makers and commercial entities in food supply chains to more proactively address EMA risk. First, we demonstrate when supply

¹⁶When the penalty is very low, all farms adulterate with the maximum dosage under either testing scenario, and hence, the resulting preemptive EMA risk is independent of the testing capability.

chain dispersion increases EMA risk in farming supply chains and how this result depends on traceability and testing sensitivity. Our results complement the supply chain risk management literature in which multi-sourcing is found to be desirable for guarding against disruption risks (e.g., earthquake, fire, hurricane). We study, instead, quality risks related to both exogenous uncertainty and endogenous decisions within the supply chain (i.e., farms adulterating their output). A recent paper by [Huang et al. \(2017b\)](#) empirically shows the adverse effect of supply chain dispersion on EMA risk in China's farming supply chains, based on supply chain and quality data across five different industries. Second, by explicitly modeling exogenous quality uncertainty (common in agricultural production) and distinguishing it from endogenous adulteration decisions, we caution that quality investments need to be accompanied by improvement in testing capabilities. This is because a lower (exogenous) chance of producing low-quality output could inadvertently induce suppliers to endogenously adulterate the realized low-quality units, if the detection of adulteration relies on having a sufficient amount of adulterants in the total output. In such situations, the ability to differentiate fake high-quality products from truly high-quality ones is essential to combat EMA.

Our results have important practical implications and also open up a fruitful avenue for future research. We recommend that companies address risks resulting from supply chain dispersion by mitigating the potential underlying issues in a dispersed supply chain, for example, by enabling better traceability and risk sharing between farms and manufacturers. Some possible strategies include creating farming cooperatives that can allow for better traceability and knowledge transfer in terms of best practices, or developing fairer contracts and support systems such as protective prices and guaranteed distribution channels for farms (especially smallholder farms who face significant financial pressures). Analytical and empirical research is needed to better understand the effectiveness of these remedies in mitigating EMA risk under different market and socioeconomic environments. In addition, future research can build upon our models to further examine other relevant supply chain settings. For example, it would be valuable to jointly consider quality and disruption risks (or yield and demand uncertainty when dispersion may be beneficial) to develop further insights regarding the role of dispersion in a farming supply chain. Analyzing repeated interactions where the manufacturer adapts her testing strategies over time could also be helpful. For policy makers and regulators, we underscore the importance of collecting data and verifying a food manufacturer's sourcing supply chain to more proactively manage EMA risk in food products. Current practices primarily focus on sampling

products and inspecting facilities, with limited attention to the upstream parts of the supply chains where agricultural production happens. As countries around the world scale up their food defense efforts, with prominent examples such as the recent enactment of the Food Safety Modernization Act in the U.S. and the new food safety law in China, our results offer timely and actionable insights on the relatively overlooked supply chain perspective in the defense of food safety.

Chapter 5

Artificial Shortage in Agricultural Supply Chains

5.1 Introduction

In the fall of 2013, while inflation rate was only 6% in India, onion prices in some regions increased by 500% ([Kapur and Riley 2013](#)). Intuition and canonical economic models suggest that, given relatively stable demand for “essential” commodities such as onions in India, the price surge must be a result of supply shortage ([Pindyck and Rubinfeld 2013](#)). However, data seems to challenge this hypothesis. Figure 5.1 presents onion retail price and supply data in major cities across India over the years of 2010–2019. The lean season for onions in each year is divided into two periods: Period 1 corresponds to July–September, and Period 2 corresponds to October–November.¹ Each point in Figure 5.1 shows the average retail price and average supply in Period 2 relative to those in Period 1 in each city and each year. A normalized price (supply) that is greater than 1 means that the retail price (supply) has increased in the later months of the lean season as compared to the earlier months for that city and that year. The large number of points in the upper right quadrant of Figure 5.1 suggest that simultaneous increase of both price and supply in the later months of the lean season occurred commonly across multiple cities throughout 2010–2019. This type of surge in prices of essential commodities is not unique to India. Similar phenomena have been observed that the prices of other essential commodities such as pulses in India ([Jess 2017](#)), maize in Kenya ([Daily Nation 2017](#)), and onions in Israel

¹Over 70% of the annual total onion production in India is harvested between December and June ([MOAFW 2017](#)). Thus, December–June is considered the peak season for onions, and July–November is considered the lean season.

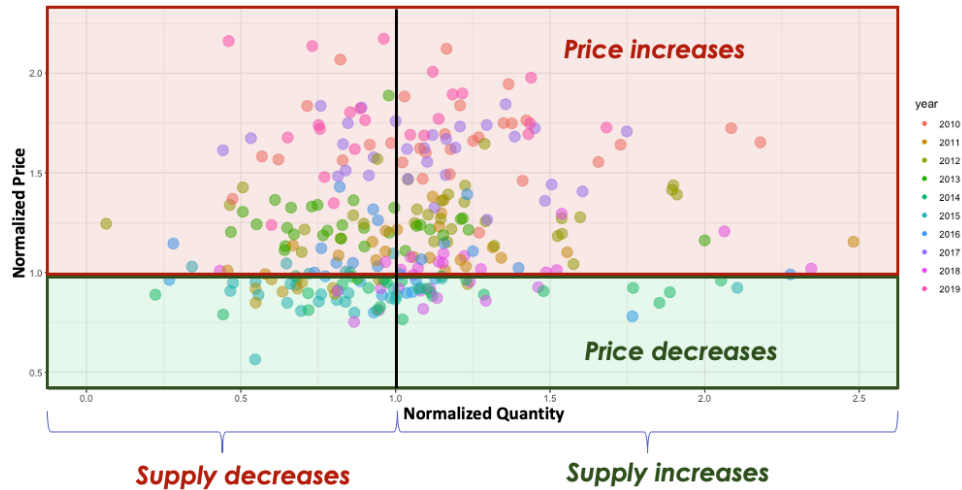


Figure 5.1: Normalized Retail Price and Supply of Onions in India's Agricultural Markets

Note: Each point in Figure 5.1 shows the price and supply of onions in the later months of the lean season relative to those in the earlier months for a city and a year. Simultaneous increase of both price and supply in the later months of the lean season (upper right quadrant) occurred commonly across multiple cities throughout 2010–2019.

and Turkey (Hamodia 2018, Jamieson 2018) soared despite increased supply to the market.

A number of case studies suggest that this phenomenon is caused by traders' strategic behavior of hoarding inventory and gradually releasing supply to benefit from rising prices (Jain 2017, Hindustan Times 2019). As Hindustan Times (2019) notes, "High onion prices may impel some traders to hoard onions and create an artificial shortage in the market, to later sell onions at exorbitant prices." Similarly, 79% of the respondents in a survey blamed deliberate hoarding by traders for price hikes of essential food items and asked for effective government measures to tackle this challenge (Sarkar et al. 2010). Given the essential nature of these commodities, governments across the world spend a significant amount of resources on intervention programs to protect consumers' welfare. The interventions range from direct imports and/or inspections that increase supply in the markets, to cash subsidy aimed at increasing consumers' purchasing power during shortage (Chaki 2013, Business Standard 2015).

This chapter is the first to develop a behavioral game-theoretic model to capture the underlying supply chain and market dynamics that engender, what is termed as, *artificial shortage*. A key aspect of these dynamics is how consumers react to perceived shortage in the market. The model is then used to evaluate the effectiveness of various government interventions in reducing artificial shortage and improving consumer welfare. Empirical analysis using real-world market data on onions in India suggests that the model explains the data well. In addition, the impacts of different government interventions on artificial shortage are estimated to highlight the practical relevance of the model.

Several factors have been hypothesized as drivers of artificial shortage. The first factor is high concentration of market power in the upstream part of the agricultural supply chain. While hoarding supply when there is unmet demand may not be optimal in a perfectly competitive market, agricultural markets in many countries are far from being competitive or perfect. For example, in India and Kenya, smallholder farmers sell to intermediary traders in geographically isolated markets where the traders effectively act as monopolists (Gosh 2017, Bergquist 2019). Similarly, agricultural markets in Israel are concentrated through powerful farmer organizations and councils that actively collude to determine total supply to the markets (Hamodia 2018). In the absence of alternative purchase channels, consumers have no other choice but to rely on these monopolistic sellers for the supply of essential commodities.

Another factor that motivates traders to create artificial shortage is consumers' behavioral response to signals of shortage. This chapter proposes a model of consumer behavior in response to artificial shortage which is inspired by the theory of mental accounting. Specifically, mental accounting posits that consumers allocate their disposable income into various mental accounts, each of which corresponds to fulfilling a certain consumption need (Thaler 1985, Prelec and Loewenstein 1998). Building on this theory, the proposed model assumes that consumers mentally allocate budgets for purchasing essential commodities versus for nonessential items (e.g., entertainment). Recent works have shown that despite fixed income, consumers' budget allocations for different purchases are not fixed, but rather, affected by market signals (e.g., Cheema and Soman 2006, Novemsky and Kahneman 2005). Due to the monopolistic nature of agricultural markets, consumers mainly rely on current supply quantity in the market to form expectations about any potential shortage in the future. Low supply quantity can lead consumers to believe that prices will soon rise. In order to afford future higher prices, consumers are likely to increase their budget allocation for the essential commodity *and* decrease budget for nonessential consumption. An extensive literature has shown that scarcity impacts consumers' willingness-to-pay for the associated products (Lynn 1991, Shi et al. 2020). For example, Lynn and Bogert (1996) find that consumers anticipate price appreciation in future periods when they are exposed to information signals (such as news reports) that indicate product scarcity. Furthermore, qualitative evidence exists that consumers facing price surge are forced to sacrifice nonessential expenditure to maintain necessary consumption of essential items such as food. For example, *Telangana Today* (2017) notes the following: "V Saritha, a housewife, said that it is impossible to prepare curries without onions. 'The dish will lack taste if prepared

without onions. Therefore, people were forced to buy them' ...” A survey by [Sarkar et al. \(2010\)](#) shows that monthly food expenditure increased by as much as 41% among low-income households during times of price hikes. As a result, 72% of the respondents had to decrease their nonfood expenditure (such as entertainment, education, and transportation). An increase in budget allocation for essential products effectively reduces consumers' price sensitivity for these products. Therefore, by creating artificial shortage in the market, the trader essentially exploits consumers' flexible (mental) budget and the resulting lower price sensitivity for economic gains. The model developed in this chapter explicitly captures these important dynamics.

Finally, government interventions also affect the extent of artificial shortage in the market. Government's intervention is often *reactive* in nature, i.e., the government intervenes only *after* there is a concrete evidence of threat to consumer welfare due to artificial shortage ([Hamodia 2018](#)). This reactive action by the government is due to at least two practical reasons ([The Economic Times 2019](#), [The Quint 2019](#)). First, powerful farmer and trader councils often lobby against the government importing and selling essential commodities under normal conditions, citing that such actions would lower prices. Second, the government also has a parallel objective of protecting farmers' income by not increasing supply in the market unless there is a credible threat to consumer welfare. This chapter examines three commonly practiced government interventions for consumers. These interventions are typically designed to increase consumers' purchasing power, or to increase product availability in the market.

Cash Subsidy (CS) is a type of intervention designed to increase consumers' purchasing power by directly transferring cash into the beneficiaries' accounts. In fact, the government of India has been considering to expand the scope of this intervention more broadly. For instance, India's central government asked the state governments to implement CS for food subsidies, and it further increased the subsidy amount due to recent increase in retail prices ([The Economic Times 2016](#), [livemint 2016](#)). The government hopes that CS can help to eliminate corrupt practices in which products dedicated to beneficiaries were siphoned off by middlemen and never reached the intended people. In 2017 alone, the Indian government disbursed 1.8 billion U.S. dollars through various CS programs in the country ([Sharma 2018](#)). The CS scheme currently used by the Indian government prioritizes low-income consumers and transfers the same amount of cash to all beneficiaries that satisfy the income threshold ([Sidhartha 2016](#)). Thus, while not currently used to address artificial shortage specifically, it is conceivable that such a policy will be considered in the near future.

Another type of commonly-used intervention, referred to as supply allocation schemes, focuses on increasing product availability in the market. To do so, the government imports the commodity from foreign markets and sells to consumers at a subsidized price in its own shops (e.g., [Herald 2010](#)). Supply allocation schemes differ based on their eligibility rules. Under *Prioritized Allocation* (PA), the government prioritizes selling the subsidized commodities to low-income consumers. A majority of food distribution programs currently run in developing countries only consider families below the country's poverty line as being eligible ([Wodon and Zaman 2008](#)). In contrast, under *Randomized Allocation* (RA), the government distributes the subsidized commodities to all consumers without income restrictions. In the past, multiple state governments in India have used their public distribution outlets to sell commodities at subsidized prices to either *all* consumers as in RA ([Bhatnagar 2011](#)), or *only* low-income consumers as in PA ([Business Standard 2015](#)) during shortage.

5.1.1 Contributions and Policy Insights

This chapter formally evaluates the effectiveness of different commonly observed government interventions in improving consumer welfare. However, when consumers adjust their budgets for essential consumption due to perceived shortage, how to measure consumer welfare is not obvious. The conventional method evaluates consumer welfare by comparing consumers' valuations of a product with the price at which they purchase the product. This method does not differentiate between consumers' true valuations (as captured by their normal budget allocation) and the seemingly higher "valuations" due to budget adjustment under artificial shortage. Thus, it can lead to the erroneous conclusion that consumer welfare *increases* rather than decreases even though consumers are paying higher prices. This chapter generalizes the conventional model of consumer welfare to capture potential losses for consumers as a result of shortage-induced budget adjustment.

Using this generalized framework to capture welfare, the chapter demonstrates, both analytically and empirically, the disparate effectiveness of different government interventions in curbing artificial shortage of essential commodities. Specifically, randomized allocation always *benefits* consumers by strictly reducing artificial shortage in the market, whereas prioritized allocation is beneficial only when government funds are sufficiently large (otherwise PA leads to no change in the extent of artificial shortage). In sharp contrast, cash subsidy can in fact backfire and create even stronger incentives for more serious artificial shortage if the govern-

ment dedicates too many funds for the intervention. The results show that *reactive* government interventions such as RA can have a *preemptive* effect and reduce traders' incentive to create artificial shortage in the first place. While previous research has focused on the effects of subsidy schemes on consumer welfare in a static setting (Levi et al. 2016), this work highlights the intertemporal effects of government interventions arising from strategic behavior of different stakeholders in the market. In line with these results, Cunha et al. (2018) found empirically through a randomized trial in Mexico that food prices were significantly lower in villages under supply allocation schemes compared to those under cash subsidy schemes.

Utilizing onion price data in India, it is shown that the model developed in this chapter that accounts for shortage-induced budget adjustment can explain the data well. It is estimated that on average, as much as 10% of supply in the lean season is being hoarded by traders to create artificial shortage in the market. This may impact as many as 114 million consumers (9% of India's population) who may suffer from unavailability of onions during this time.² Furthermore, the analysis suggests that cash subsidy (with overspending) could *increase* artificial shortage from 10% to 12%. Conversely, randomized allocation can *reduce* artificial shortage to 6% on average, benefiting over 45 million consumers. These results highlight that, to effectively address artificial shortage of essential commodities, government interventions need to be carefully designed while accounting for the traders' strategic hoarding behavior and consumers' behavioral response to perceived supply shortage.

The remainder of the chapter is structured as follows. §5.2 reviews the related literature. §5.3 describes the base model with no government intervention. §5.4 introduces a new definition of consumer welfare accounting for shortage-induced budget adjustment and describes the modeling of different government interventions. §5.5 analyzes the market equilibrium under each intervention, and §5.6 contrasts the effectiveness of different interventions. §5.7 presents the empirical analysis based on actual onion price data in India. §5.8 concludes the chapter.

²Total onion production in India was estimated to be 21.4 million tons in 2017–2018, and about 30% of this quantity was harvested in the lean season (The Economic Times 2018). Our model predicts that 10% of the supply in the first three months of this season (which accounted for 60% of the 30% total production) was strategically hoarded. Assuming uniform consumption over time, consumers are affected in 3 of the 12 months when artificial shortage occurs. Given yearly per capita onion consumption of 13.5 kg in India in 2013 (Helgi Library 2018), the hoarded quantity affected $21.4 \times 30\% \times 60\% \times 10\% / (0.0135 \times 25\%) = 114.1$ million consumers.

5.2 Related Literature

This work is related to three streams of literature in operations management. The first stream is research that studies the effects of customer behavior on the price-demand curve and the seller's pricing and inventory decisions. The reader is referred to [Özer and Zheng \(2012\)](#) and [Ovchinnikov \(2018\)](#) for comprehensive reviews. Most closely related to this work, [Gallego et al. \(2008\)](#), [Liu and Van Ryzin \(2011\)](#), and [Özer and Zheng \(2015\)](#) examine a seller's optimal pricing and inventory decisions for a seasonal product when consumers' purchase decisions are affected by their perception of inventory availability. Research in this area mainly focuses on the seller's profit maximization and considers consumers' decision of *when* to purchase. The current work differs from these studies in two important aspects. First, prior demand models cannot explain a joint increase in both price and supply under stable demand, which, as discussed earlier, is a key market characteristic resulted from artificial shortage. This work develops a new model to capture the behavioral factor of shortage-induced budget adjustment that is unique in this setting and previously unexplored. Second, this work studies the government's problem of designing interventions to maximize consumer welfare in the presence of the trader's strategic hoarding behavior. Therefore, the model examines strategic interactions between the trader and the government while incorporating consumers' behavior.

The second stream of related literature studies government subsidy programs in various domains, e.g., to incentivize green technology adoption ([Cohen et al. 2015](#), [Hu et al. 2015](#)), or to increase access and consumption of vaccines ([Taylor and Xiao 2014](#)). Economists have also extensively studied the topic of subsidy allocation. Related to food, [Barrett \(2002\)](#) presents a thorough review of food security and assistance programs across different countries. This research stream mainly focuses on consumption effects of different subsidy programs in a single period setting. In contrast, this chapter analyzes effects of subsidy programs on strategic shortage of essential commodities in a two period setting. Within the agricultural domain, recent papers have examined subsidy programs dedicated to improving farmers' welfare (e.g., [Alizamir et al. 2018](#), [Akkaya and Bimpikis 2016](#), [Guda et al. 2016](#)). Instead, this work examines government interventions for protecting consumer welfare in the face of artificial shortage. As such, the current work differs from earlier ones as it analyzes the interaction between the trader's strategic decisions and the government's design of intervention programs. The analysis shows that this interaction critically impacts the effectiveness of various interventions.

This work also relates to the growing literature of socially responsible operations, in which a majority of the works consider settings of developing economies. A main focus in this literature is how to deter socially undesirable behaviors of suppliers and to motivate more responsible practices (see [Atasu 2016](#), [Bouchery et al. 2017](#), for broad reviews). A handful of recent papers specifically focus on the production (farmer) stage of agricultural supply chains. [Hu et al. \(2019\)](#) and [Boyabati et al. \(2019\)](#) examine farmers' herding behavior in their cropping decisions and the benefit of crop rotation. [Lim et al. \(2019\)](#) examine how the design of fair trade certification may impact smallholder farmers' income. [Yayla-Kullu et al. \(2019\)](#) focus on significant volatility in onion prices and analyze when it is optimal to introduce a processed substitute in the market. We add to this body of works by examining a critical, yet understudied, challenge prevalent in many developing countries – how governments should effectively deter strategic hoarding by agricultural traders to ensure access of essential commodities to consumers.

5.3 Base Model: Artificial Shortage under No Government Intervention

Consider a monopolist trader (he) selling an essential commodity over two periods within the same (lean) season. A commodity is considered as “essential” if it is an indispensable ingredient in people’s daily diets and suitable substitutes are difficult to find (e.g., lentils and onions in India; [Pal 2017](#)). The total market size for the commodity in each period is fixed and normalized to 1. This assumption is reasonable because consumption of essential commodities is typically stable. Furthermore, the monopolistic nature of the model is justified by evidence of low competition in agricultural markets in many developing countries ([Gouri 2017](#)). Due to substantial financial and logistical constraints, farmers rely on selling their products to local traders to earn their income. Therefore, the traders purchase the entire production quantity and control the supply to the retail market ([Inter Press Service 2012](#)).

The model dynamics evolve as follows: (i) Each consumer begins with a privately known, heterogeneous budget b_c for purchasing the commodity in each period. The trader only knows that the consumers' budgets in Period 1 are uniformly distributed between $[0, 1]$. In [Appendix D.3.1](#), we show the robustness of our results when consumers' budget distribution is skewed rather than uniform. (ii) The trader sets the market price p_1 by releasing quantity q_1 in Period 1. (iii) Consumers observe p_1 and purchase the commodity if $b_c \geq p_1$. In addition, they adjust

their budgets for Period 2 based on the perceived shortage signaled by q_1 . In Appendix D.3.2, we show the robustness of our results when consumers may purchase fractional units when their budget is below the price (as opposed to a binary purchase decision). (iv) The trader sets the market price p_2 by releasing quantity q_2 in Period 2, and all consumers whose updated budget exceeds or equals p_2 purchase the commodity. Appendix D.1 summarizes the model notations. The trader's objective is to maximize his total expected revenue, $p_1q_1 + p_2q_2$, from selling in both periods. In our main model, we assume that the trader has no inventory constraint to focus on the phenomenon of artificial shortage. In Appendix D.3.3, it is shown that our results continue to hold when the trader is actually constrained in inventory and incurs an inventory holding cost for unsold units in Period 1. With unconstrained, known inventory and a monopolist, supply and demand exactly clear in each period. Therefore, the quantity released in Period 1 (q_1) is exactly equal to the quantity sold at price p_1 given consumers' budgets; i.e., $q_1 = 1 - p_1$. We now formally define the phenomenon of artificial shortage given the model setup.

Definition 1. *The trader is considered to have created "artificial shortage" if both of the following conditions are satisfied: (i) the quantity sold in Period 1 is smaller than that in Period 2; and (ii) despite increased quantity in Period 2, the price in Period 2 is higher than that in Period 1.*

A key feature in our model is that consumers adjust their budgets for Period 2 based on perceived shortage in the market. This feature is built upon the theory of mental accounting (Thaler 1985). Specifically, consumers allocate their disposable income to a set of mental budgets for different purchase needs (e.g., food, entertainment, education, and so on). As discussed in §5.1, if consumers learned that the supply of an essential commodity in the market was low (from reading news, talking to friends, or directly observing it in the market), then they would likely anticipate that prices were on the rise. As a result, consumers are likely to increase their budgets for purchasing the essential commodity and decrease budgets for nonessential consumption (Sarkar et al. 2010, Telangana Today 2017).

The Period 1 quantity, q_1 , released in the market by the trader serves as a signal that could affect consumers' perception of shortage. In particular, with a lower q_1 , consumers would perceive a more severe shortage in the market, and as a result, they would increase their budgets more substantially. Hence, the updated budget of a consumer with initial budget b_c is modeled as $b_c(1 + \delta(q_1))$, where $\delta(q_1)$ denotes the relative increase in the consumer's budget. We model $\delta(q_1)$ to be linearly decreasing: $\delta(q_1) = g(1 - q_1)$. The parameter g is called the consumers'

budget adjustment factor. A larger value of g implies that consumers increase their budgets to a larger extent given q_1 . In the main model, all consumers increase their budgets by the same relative amount regardless of their initial budgets. The assumption that *all* consumers update their budgets is driven by the focus on “essential” products in this work. In Appendix D.3.4, we show the robustness of our results if consumers with higher initial budgets increase their budgets by a larger relative amount.

Note that $\delta(1) = 0$ and $\delta(q_1) > 0$ for all $q_1 \in [0, 1)$. That is, if q_1 satisfies the maximum market demand in Period 1 ($q_1 = 1$), then consumers should not perceive any shortage and would not increase their budgets in Period 2. Conversely, if some consumers cannot obtain the commodity in Period 1 ($q_1 < 1$), then consumers perceive shortage and increase budget. (Alternatively, one could also consider the case that consumers perceive shortage and increase their budgets only when $q_1 \leq q_{\min}$ for some $q_{\min} < 1$. In this case, it is assumed that $\delta(q_1) = 0$ for all $q > q_{\min}$, and all of our insights continue to hold.) Finally, since consumption of essential commodities such as onions is required in *every* period, it is unlikely that consumers would wait and make a strategic decision about when to purchase.³

Given the updated budget in Period 2, a consumer purchases the commodity at price p_2 if and only if $b_c(1+g(1-q_1)) \geq p_2$. Since b_c is uniformly distributed on $[0, 1]$, the demand in Period 2 (and hence the quantity released, q_2) is characterized as $q_2 = 1 - p_2/(1+g(1-q_1))$. Therefore, the increase in consumers’ budget due to perceived shortage effectively reduces consumers’ price sensitivity (because $1 + g(1 - q_1) > 1$). As a result, the trader can charge a higher price by creating artificial shortage in the market. This is formalized in the following result.

Proposition 5.3.1. Under no government intervention, the optimal quantity released and the optimal price charged by the trader in each period are as follows: $(q_1^*, p_1^*) = (1/2 - g/8, 1/2 + g/8)$, and $(q_2^*, p_2^*) = (1/2, 1/2 + g/4 + g^2/16)$, where g is consumers’ budget adjustment factor.

Observe from Proposition 5.3.1 that, as long as $g > 0$, we have both $q_1^* < q_2^*$ and $p_1^* < p_2^*$. This result violates the typical decreasing relationship between supply and price. Note that if consumers’ budgets for the commodity stayed constant (and hence, the demand curve did not change) between the two periods, then it would be optimal for the monopolist trader to sell the same quantity at the same price in both periods; i.e., artificial shortage as defined

³Yayla-Kullu et al. (2019) make a similar assumption in their analysis of the onion markets in India. In addition, stockpiling by consumers is difficult because the majority of consumers do not have proper storage facilities for perishable products, and they strongly prefer buying fresh every day especially in tropical countries such as India.

in Definition 1 would not arise. Therefore, Proposition 5.3.1 highlights a key characteristic of artificial shortage: By leveraging consumers' budget adjustment in response to artificial shortage, the trader benefits by selling a *larger* quantity at a *higher* price in Period 2. In Appendix D.3.3, we show that this result extends to the case when the trader incurs a moderate inventory holding cost for unsold units in Period 1.

5.4 Modeling Government Interventions

Due to artificial shortage created by the trader under no intervention, consumers suffer from having to pay a higher price for an essential commodity. Governments have the responsibility to intervene and protect consumers' welfare in such situations. As discussed in §5.1, governments in practice often intervene *reactively* in Period 2, i.e., after artificial shortage has already occurred in the market. Thus, we focus on this setting in the main model. In Appendix D.3.5, we analyze an extension where the government may intervene preemptively (i.e., before artificial shortage occurs).

With government intervention, we model the interactions among the trader, the government, and the consumers as a three-stage dynamic game. The game proceeds similarly as the model dynamics described at the start of §5.3, with the following additional actions by the government. First, the government commits to a certain intervention scheme at the beginning of the two periods. Second, after observing the trader's quantity decision in Period 1 and before the trader makes the Period 2 decision, the government decides the detailed parameters of the scheme it has committed to (see step (iv) in Figure 5.2). We analyze three commonly-used government interventions: (i) cash subsidy (CS), (ii) prioritized allocation (PA), and (iii) randomized allocation (RA). Under CS, the government directly transfers cash to a fraction (f) of consumers with the lowest income levels (and correspondingly, the lowest budgets). Each eligible consumer receives the same amount of cash (c) in Period 2. For both PA and RA, the government imports a quantity q_g of the commodity from foreign markets and sells to consumers at a subsidized rate (p_g) in Period 2. Under PA, the government prioritizes selling the commodity to the lowest-income consumers until it exhausts all q_g units. PA is implemented in practice by only allowing consumers below a certain income level to buy at subsidized rates (Business Standard 2015). Under RA, all consumers in the market have an equal chance to purchase the commodity from the government. RA is implemented in practice by declaring that all consumers are eligible to

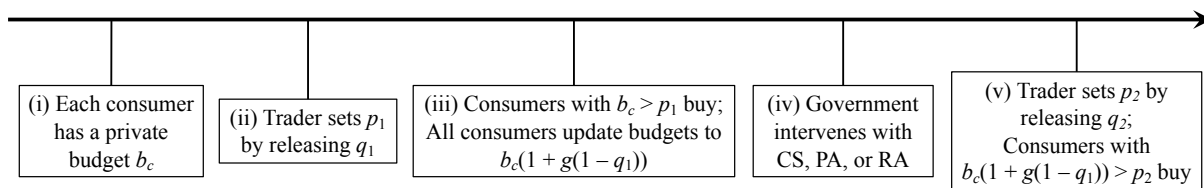


Figure 5.2: Dynamics of the Game

buy at subsidized rates (Business Standard 2015). For each intervention, the government has a total budget of b to spend. We treat this budget as predetermined (and hence exogenous to our model), as governments typically make these budgeting decisions at a time much earlier than when the money is actually used for intervention programs (Ministry of Agriculture 2015). In the main model, consumers' budget adjustment factor (g) is unaffected by government intervention. In Appendix D.3.6, we show the robustness of our results when the government's intervention reduces the extent that consumers increase their budgets due to artificial shortage (i.e., reducing g).

5.4.1 Modeling Consumer Welfare under Shortage-induced Budget Adjustment

Since the government intervenes in Period 2, we consider its objective to be maximizing consumer welfare in Period 2. Recall from §5.3 that consumers update their budgets for Period 2 in response to artificial shortage in the market. A key question thus arises as to how we should define and evaluate consumer welfare given shortage-induced budget adjustment. The conventional definition evaluates consumer welfare by quantifying the total monetary savings gained by consumers due to paying a price lower than their valuations of the product (Mankiw 2014). Figure 5.3a illustrates this definition in our setup. The solid black line corresponds to the price-demand curve given consumers' updated budgets in Period 2: $p_2(q) = (1 - q)(1 + g(1 - q_1))$. The solid gray line corresponds to the price-demand curve given consumers' original budgets: $p_1(q) = 1 - q$. Due to the essential nature of the commodity, we also capture a consumption utility (λ) for all consumers who are able to purchase the commodity, and a disutility of nonconsumption ($-\lambda$) for those who cannot purchase it.⁴ The conventional definition treats the

⁴Including the λ parameter in defining consumer welfare is inspired by recent development in consumer utility models that consider both dimensions of monetary savings/overspending and functional utility from consumption/disutility from nonconsumption (Kőszegi and Rabin 2006). This approach is appropriate in our setting because for essential commodities, monetary savings are not interchangeable with functional consumption utility. The functional utility of consuming an essential product would far exceed the monetary savings of not buying the product, prompting consumers to be willing to overspend to purchase the product. Our results continue to hold as long as the disutility of nonconsumption is larger than the utility of consumption.

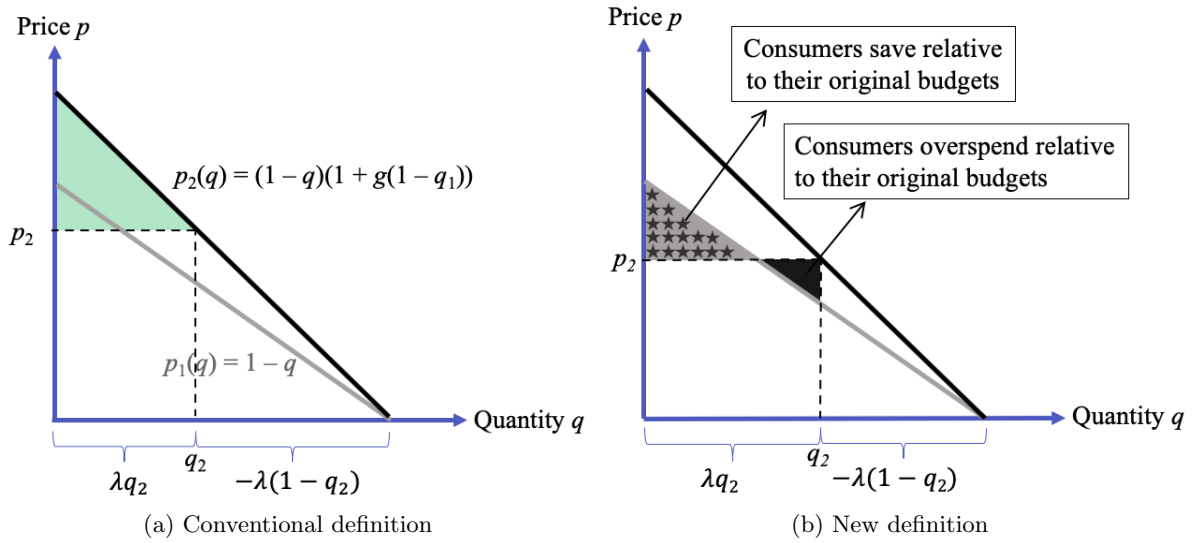


Figure 5.3: Evaluating Consumer Surplus with Shortage-Induced Budget Adjustment

Period 2 price-demand curve as representing consumers' new valuations of the product. As such, it would lead to the erroneous conclusion that consumer surplus *increases* due to artificial shortage (because $p_2(q) > p_1(q)$ for all q when $g > 0$). Put it differently, the conventional definition suggests that consumers *save* more money because they have allocated a larger budget for the purchase. However, this definition ignores the fact that the budget increase comes at the expense of reducing budgets for other spending categories; consumers do so because the commodity is essential (Sarkar et al. 2010). The increase, in fact, represents anticipated overspending due to rising prices. Consumers are therefore likely to feel a loss, not a gain. Hence, a new definition of consumer surplus is needed to properly capture the impact of shortage-induced budget adjustment on consumers' welfare.

In particular, we propose that the consumer surplus should be evaluated based on their original budgets for the commodity. Since the original budgets are allocated based on normal spending without being affected by market conditions, they should more accurately capture consumers' true valuations. Compared to their original budgets, consumers who are able to purchase the commodity in Period 2 given their updated budgets may fall into one of two situations. First, as captured by the star gray region in Figure 5.3b, these consumers would have been able to purchase the commodity even with their original budgets. Therefore, they earn a positive surplus. Second, as captured by the solid black region in Figure 5.3b, these consumers are able to purchase only because they have increased their budgets; i.e., they overspend relative to their normal budgets. Hence, they suffer a negative surplus because the budget increase is at the expense of reducing budgets for other categories. This new definition of consumer welfare,

denoted as CW^{upd} , can be characterized as follows, which continues to capture the utility of consumption/disutility of nonconsumption as discussed earlier.

$$CW^{upd}(q_1, q_2) = \int_0^{q_2} (1 - q) dq - (1 - q_2)(1 + g(1 - q_1))q_2 + 2\lambda q_2 - \lambda. \quad (5.1)$$

To rule out the unrealistic scenario that consumers increase their budgets to exceed the value of consuming the commodity, we assume that the maximum budget that a consumer allocates for the commodity is smaller than its consumption utility (i.e., $1 + g < \lambda$). Given the new definition, the following result characterizes the consumer welfare under artificial shortage without government intervention.

Proposition 5.4.1. Given the trader's optimal decisions characterized in Proposition 5.3.1, the consumer welfare in Periods 1 and 2 are $CW_\phi^1 = (g - 4)^2/128 - \lambda g/4$ and $CW_\phi^2 = 1/8 - g/8 - g^2/32$.

In what follows, we first analyze the trader's and the government's strategic decisions in the game under each intervention. We then examine which intervention is the most effective in maximizing consumer welfare in the presence of artificial shortage.

5.5 Analysis of Government Interventions

We solve for the trader's and the government's equilibrium decisions in the three-stage game using backward induction. The analysis starts by first characterizing the price-demand curve in Period 2 accounting for consumers' budget adjustment and government interventions. Using this characterization, we solve for the trader's decision in Period 2. Next, we characterize the government's decisions when it anticipates the trader's best response in Period 2. Finally, we analyze the trader's decision in Period 1 when he anticipates the government's best response as well as his own decision in Period 2. Since we characterize the trader's optimal decision in Period 2 for *all* possible actions by the government, the key insights can be easily extended to scenarios where one or both of the government's decision variables are exogenous and cannot be optimized over.

5.5.1 Cash Subsidy (CS)

Recall from §5.4 that in this setting, the government makes the following decisions: (i) what fraction of consumers (f) to subsidize, starting from those with the lowest income; and (ii)

how much cash (c) to subsidize for each eligible consumer. The government's total budget for the intervention is b . The trader decides quantities released in Periods 1 (q_1) and 2 (q_2) to maximize his total revenue. Let $p_1(q_1)$ denote the realized price in Period 1 when the quantity released is q_1 . Since the trader decides q_2 after the government has implemented CS, the price-demand curve in Period 2 is affected by the government's decisions. Let $p_2(q_2, q_1, c, f)$ denote the realized price in Period 2 when the quantities released in the two periods are q_1 and q_2 , and the government has distributed a per-consumer subsidy of c to a fraction f of consumers. Let $CW_2(q_1, c, f, q_2)$ denote the consumer surplus in Period 2 given the trader's and the government's decisions. Solving the game is equivalent to solving the following tri-level optimization problem:

$$\max_{q_1} \{q_1 p_1(q_1) + q_2^*(c^*, f^*, q_1) p_2^*(q_2^*, q_1, c^*, f^*)\} \quad (5.2)$$

$$s.t. \quad (c^*(q_1), f^*(q_1)) = \arg \max_{c, f} \{CW_2(q_1, c, f, q_2^*) \mid cf \leq b\}, \quad (5.3)$$

$$q_2^*(c, f, q_1) = \arg \max_{q_2} q_2 p_2(q_2, q_1, c, f) \quad (5.4)$$

Equation (5.4) represents the trader's problem in Period 2 after the government has already declared its CS policy. Equation (5.3) represents the government's problem of selecting the optimal CS policy, $(c^*(q_1), f^*(q_1))$, to maximize consumer surplus in Period 2 anticipating the trader's best response in Period 2. Finally, the objective function in Equation (5.2) represents the trader's problem of maximizing total revenue in both periods.

We start by characterizing the price-demand curve in Period 2 given the government's CS policy (c, f) . If a consumer's original budget (b_c) is lower (greater) than f , then she receives (does not receive) a cash subsidy of c from the government. Recall from §5.3 that all consumers update their budget allocations in Period 2 to $b_c(1 + g(1 - q_1))$. Consumers buy as long as the price in period 2, p_2 , is not greater than their updated budget in period 2 (i.e., $c + b_c(1 + g(1 - q_1))$ if they are subsidized and $b_c(1 + g(1 - q_1))$ otherwise). Because we consider essential commodities, we assume that consumers who receive cash subsidy will spend *all* of it on the commodity and do not decrease their own budget allocation in anticipation of the government's policy. That is, subsidized consumers buy if and only if $b_c \geq (p_2 - c)/(1 + g(1 - q_1))$, and non-subsidized consumers buy if and only if $b_c \geq p_2/(1 + g(1 - q_1))$. Thus, cash subsidy enables purchase by lower-budget consumers. The total demand is the sum of demand from both subsidized and non-subsidized consumers.

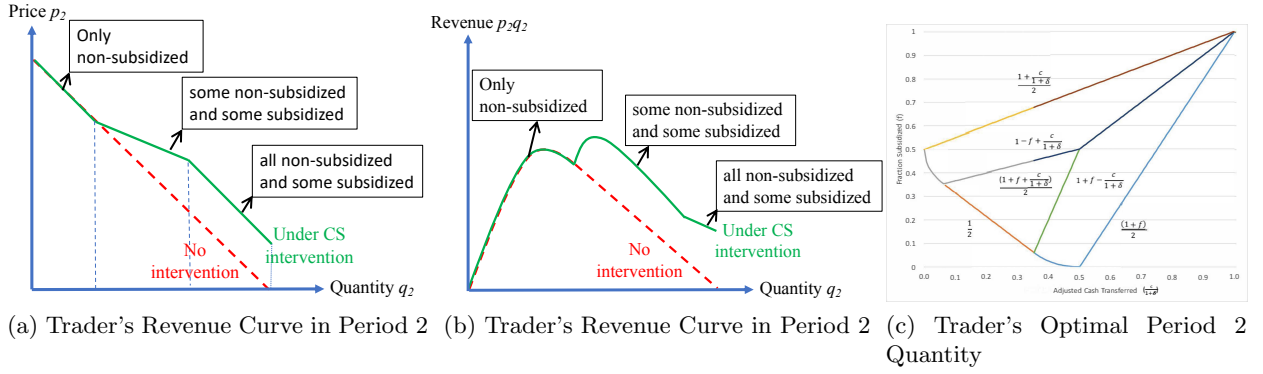


Figure 5.4: Revenue Curve and Optimal Quantity for the Trader in Period 2 Under CS

Proposition D.2.1 characterizes the price-demand curve in Period 2, $p_2(q_2, q_1, c, f)$, under CS. Figure 5.4a illustrates Proposition D.2.1. We observe that the price-demand curve moves upward due to the availability of additional cash. In addition, Proposition D.2.1 and Figure 5.4a describe three cases of demand compositions in Period 2: (i) only a portion of non-subsidized consumers buy (when q_2 is low); (ii) a portion of both non-subsidized and subsidized consumers buy (when q_2 is medium); and (iii) all non-subsidized consumers and a portion of subsidized consumers buy (when q_2 is large).

Trader's optimal decision in Period 2: Figure 5.4c depicts the optimal quantity released by the trader for different values of cash subsidy (c) and subsidized fraction (f ; characterized formally in Proposition D.2.2). When the subsidy amount is very small and the number of subsidized consumers are not large enough (lower left region in Figure 5.4c), it is optimal for the trader to not release any additional quantity in the market. Government spending in this region effectively remains un-utilized as subsidized consumers are unable to buy in equilibrium even given the subsidy. As the subsidy amount increases, the revenue gain from serving subsidized consumers increases and the trader starts releasing a larger quantity.

Government's optimal cash subsidy policy: Given the best response function of the trader in period 2 ($q_2^*(c, f, q_1)$), we solve the government's problem of maximizing consumer welfare under a budget constraint (Equation (5.3)). Let $s \equiv q_2 - p_2/(1 + g(1 - q_1))$ be the number of subsidized consumers who buy in equilibrium. Following the approach introduced in §5.4, the total consumer welfare under CS in Period 2 can be quantified as follows.

$$CW_2(q_1, c, f, q_2) = \int_0^{q_2-s} p_1(q) dq + \int_{1-f}^{1-f+s} p_1(q) dq - p_2(q_2, q_1, f, c) q_2 + sc + \lambda q_2 - \lambda(1 - q_2) \quad (5.5)$$

The first and the second terms in Equation (5.5) capture the original budget allocations of

non-subsidized and subsidized consumers who buy in equilibrium. The next two terms capture the net amount that is paid by consumers for buying the commodity. While $p_2(q_2, q_1, f, c)q_2$ is the total amount paid by consumers to the trader, sc captures the part that comes from the government's subsidy. Finally, the last two terms capture consumers' consumption utility and non consumption disutility.

Identifying the optimal solution for the government under the budget constraint is not trivial. This is because the trader's revenue is bimodal and there exist regions where it is optimal for the trader to only sell to non-subsidized consumers (see lower left region in Figure 5.3). Although increasing the subsidy amount can mean more savings for subsidized consumers who can buy, it also means fewer subsidized consumers because of the budget constraint. The government must carefully balance the subsidized fraction (f) with the subsidy amount (c) in the optimal solution. Theorem 5.5.1 characterizes the government's optimal decisions under different budget levels.

Theorem 5.5.1. Let $f^*(q_1)$ and $c^*(q_1)$ be the optimal subsidized fraction and per-consumer subsidy amount under the optimal CS policy. Then $f^*(q_1) = \min(1, b/c^*(q_1))$, and $c^*(q_1)$ is given as follows.

- (i) If $b \leq \frac{(3 - 2\sqrt{2})}{8}(1 + g(1 - q_1))$, then $c^*(q_1) = \frac{(1 + g(1 - q_1) - 2\sqrt{b(1 + g(1 - q_1))})}{2}$.⁵
- (ii) If $b \in (\frac{(3 - 2\sqrt{2})}{8}(1 + g(1 - q_1)), \frac{(1 + g(1 - q_1))}{4}]$, then $c^*(q_1) = (1 + g(1 - q_1) + \sqrt{(1 + g(1 - q_1))(1 + 12b + g(1 - q_1))})/6$.
- (iii) If $b \in (\frac{(1 + g(1 - q_1))}{4}, 1 + g(1 - q_1)]$, then $c^*(q_1) = \sqrt{b(1 + g(1 - q_1))}$.
- (iv) If $b > 1 + g(1 - q_1)$, then $c^*(q_1) = 1 + g(1 - q_1)$.

Theorem 5.5.1(i) shows that when the government's budget is very small, the optimal subsidy amount (c^*) is decreasing in the government's budget. The government's budget in this region is so small that it cannot increase both c^* and f^* . Thus, by decreasing c^* (and thereby increasing f^*), the government ensures that the trader's revenue from selling to subsidized consumers (the second mode in Figure 5.4b) becomes equal to the trader's revenue when he only sells to non-subsidized consumers (the first mode in Figure 5.4b). As the government's budget increases, the subsidy amount and subsidized fraction both increase until the budget b becomes equal to $1 + g(1 - q_1)$.

⁵There are multiple optimal q_2^* for the trader in this case. We assume that the government can adjust $c^*(q_1)$ by an infinitesimal amount to induce the trader to choose the larger q_2^* .

Trader's optimal decision in Period 1: We finally analyze q_1^* , the optimal quantity that the trader releases in Period 1 (Equation (5.2)). The trader's objective function is again a bimodal curve. The first mode represents the point at which the trader's revenue from Period 2 is equal to that under no intervention (Theorem 5.5.1(i)) and is unaffected by the government's budget. As a result, the trader continues to create the same amount of shortage as in the no intervention case at this mode. Conversely, at the second mode, the trader's revenue in Period 2 is affected by the government's subsidy. Define t_1 as the smallest budget b such that the trader's revenue at the second mode becomes equal to that at the first mode (see Lemma D.2.3). Theorem 5.5.2 characterizes the trader's optimal decision in Period 1.

Theorem 5.5.2. Given t_1 defined in Lemma D.2.3, the optimal quantity released by the trader in Period 1, q_1^* , is as follows. (i) If $b < t_1$, then $q_1^* = 1/2 - g/8$.

(ii) If $b \in [t_1, (8 + 4g + g^2)/32]$, then $q_1^* \in [0, 1]$ and is the solution to the following fixed point equation:

$$1 - 2q_1 + \frac{g^2(-1 + q_1) - g(1 + 6b + \sqrt{1 + g(1 - q_1)(1 + g + 12b - gq_1)})}{9\sqrt{(-1 + g(-1 + q_1))(-1 - 12b + g(-1 + q_1))}} = 0.$$

(iii) If $b \in [(8 + 4g + g^2)/32, (4 + 2g + g^2)/4]$, then $q_1^* \in [0, 1]$ and is the solution to the following fixed point equation: $1 - 2q_1 - bg/(2\sqrt{b(1 + g(1 - q_1))}) = 0$.

(iv) If $b > (4 + 2g + g^2)/4$, then $q_1^* = (1 - g)/2$.

If the government's budget is very large (Theorem 5.5.2(iv)), the trader releases the smallest q_1^* . In contrast, if the government budget is very small (Theorem 5.5.2(i)), q_1^* is the same as in the no intervention case. This result can be explained by the following dynamics. When the government budget is very small, any further increase in consumer budget, compared to the no intervention case, is not beneficial for the trader because the loss from lower sales in Period 1 is larger than the gain from increased consumer budget in Period 2. However, as the government's budget increases, consumers' spending ability in Period 2 increases, and the gain from increasing consumer budget starts to outweigh the loss from decreased sales in Period 1. Thus, it becomes optimal for the trader to withhold more inventory in Period 1 to increase consumer budget and benefit from the government's CS intervention.

5.5.2 Supply Allocation Schemes

We next analyze the case in which the government intervenes through a supply allocation scheme. Recall from §5.3 that in this case the government imports q_g units (at price p_b) and sells them at a subsidized price (p_g) in Period 2. Similar to §5.5.1, the government has a total budget of b . We analyze two allocation schemes in particular: (i) Under prioritized allocation (PA), the government prioritizes distribution of the commodity to the lowest-income consumers; (ii) under randomized allocation (RA), all consumers in the market have an equal chance of purchasing the commodity from the government.

We highlight two additional features that are modeled for supply allocation schemes (Equation (5.7)). First, we capture in the government's objective function the revenue it earns from selling the commodity. The parameter μ reflects the relative importance that the government places on earning revenue in comparison to maximizing consumer welfare. A high value of μ implies that the government is sensitive to spending. The government's preference for a higher revenue ($\mu p_g q_g$) in the objective function can be equivalently viewed as its preference for a lower regulatory cost, defined as $\mu(b - q_g p_g)$. Second, we model an additional constraint that the subsidized price (p_g) charged by the government must be lower than the price charged by the trader (p_2) in equilibrium. Similar as in §5.5.1, Equations (5.8) and (5.6) represent the trader's problem in Period 2 and Period 1 respectively.

$$\max_{q_1} \{q_1 p_1(q_1) + p_2^*(q_2^*, q_g^*, p_g^*, q_1) q_2^*(q_g^*, p_g^*, q_1)\} \quad (5.6)$$

$$s.t. \quad (q_g^*(q_1), p_g^*(q_1)) = \arg \max_{q_g, p_g} \{CW_2(q_1, q_g, p_g, q_2^*) + \mu q_g p_g \mid p_g \leq p_2(q_2^*, q_g, p_g, q_1), p_b q_g \leq b\}$$

$$q_2^*(q_g, p_g, q_1) = \arg \max_{q_2} q_2 p_2(q_2, q_g, p_g, q_1) \quad (5.8)$$

We again solve the above problem using backward induction starting with the trader's problem in Period 2. Under PA, consumers may not be able to buy from the government for two reasons. First, if the government's subsidized price (p_g) is higher than the consumer's updated budget ($b_c(1+g(1-q_1))$), then the consumer cannot afford it. Second, if the government has limited inventory, consumers who can afford may still be unable to buy because they are not prioritized under PA. In contrast, under RA, all consumers who can afford the subsidized price would be able to buy from the government with a positive probability. This probability depends on the government's inventory relative to the total number of consumers who can afford the subsidized price. If the government has q_g units and the total number of consumers who can

afford the subsidized price, referred to as the potential demand, is q_d , then each consumer will be able to buy from the government with probability $\min(1, q_g/q_d)$.

Trader's optimal decision in Period 2: Proposition D.2.4 characterizes the price-demand curve for the trader under both PA and RA in Period 2. Figure 5.5 illustrates Proposition D.2.4. If the number of units sold by the trader is larger (smaller) than a threshold under PA, then the market price in Period 2 will be smaller than (the same as) that under no intervention (Figure 5.5a). This result is due to the following dynamics. Under PA, a total of q_g consumers will buy from the government, and hence, the market size for the trader is reduced. If the trader releases a small quantity, then the consumers who buy from the trader in equilibrium are those with high Period 2 budgets. These consumers could not buy from the government anyways due to low priority under PA. Therefore, the trader charges the same Period 2 price as if the government did not intervene. In contrast, if the quantity released by the trader is larger than a threshold ($q_g - q_d$), then the trader effectively engages in quantity competition with the government. As a result, the trader's Period 2 price drops relative to the no intervention case. Different from the case of PA, we observe from Figure 5.5b that the price-demand curve under RA always shifts downward regardless of the trader's quantity decision. Since all consumers have equal chance of buying from the government under RA, there is a positive probability that those consumers who would have otherwise bought from the trader actually buy from the government. Thus, the market price in Period 2 is always smaller than that in the no intervention case (Proposition D.2.4(ii)). The trader's revenue in Period 2 is again bimodal with the modes depending on the budget level of those consumers served by the government. Proposition D.2.5 characterizes the optimal decisions for the trader in Period 2 under both PA and RA.

Government's optimal allocation policy: To solve for the government's optimal decisions for PA and RA, we first characterize the resulting consumer welfare CW^{PA} and CW^{RA} as follows.

$$\begin{aligned}
 CW^{PA}(q_1, q_g, p_g, q_2) \equiv & \int_0^{\min(q_2, q_d - q_g)} p_1(q) dq + \int_{q_d}^{\max(q_d, q_2 + q_g)} p_1(q) dq - p_2(q_2, q_g, p_g, q_1) q_2 \\
 & + \int_{q_d - q_g}^{q_d} p_1(q) dq - p_g q_g + \lambda(q_2 + q_g) - \lambda(1 - (q_2 + q_g))
 \end{aligned} \tag{5.9}$$

Recall that $q_d \equiv 1 - p_g/(1 + \delta)$ is the potential demand for the government when it charges a subsidized price of p_g . The first three terms in Equation (5.9) account for the saving and overspending for consumers who buy from the trader in equilibrium. The next two terms

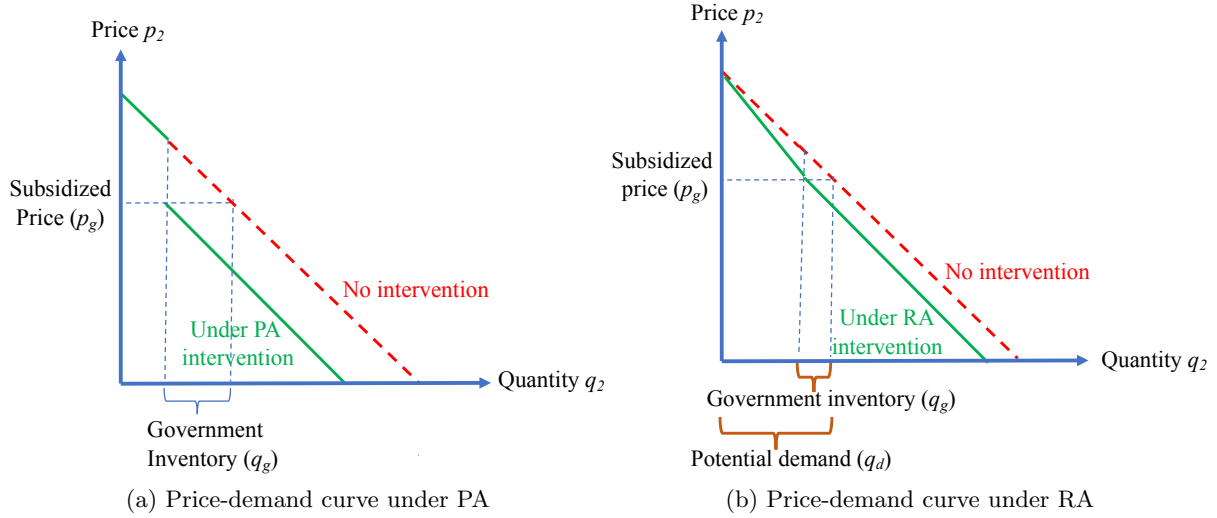


Figure 5.5: Effect of Supply Allocation Schemes on Period 2 Price-Demand Curve account for the saving and overspending for consumers who buy from the government. The last two terms capture the consumption utility and nonconsumption disutility. Similarly, consumer welfare under RA can be quantified as follows.

$$\begin{aligned}
 CW^{RA}(q_2, q_g, q_d) \equiv & \int_0^{q_2/(1-q_g/q_d)} p_1(q) dq + \frac{q_g}{q_d} \int_{q_2/(1-q_g/q_d)}^{q_d} p_1(q) dq - p_g q_g \\
 & - p_2(q_2, q_g, p_g, q_1) q_2 + \lambda(q_2 + q_g) - \lambda(1 - (q_2 + q_g))
 \end{aligned} \quad (5.10)$$

Under RA, each consumer whose Period 2 budget is greater than the government's subsidized price (p_g) has a positive probability of buying the product from the government. Consumers with high budgets will buy from the trader if they can not buy from the government. However, those with lower budgets can only afford to buy from the government (with probability q_g/q_d). The first two terms in Equation (5.10) capture these scenarios. The remaining terms follow from Equation (5.9).

The government's objective is to maximize a weighted sum between consumer welfare and revenue earned from selling, with $\mu < 1$ being the weight the government places on revenue. Theorem 5.5.3 characterizes the government's optimal PA and RA policies (q_g^*, p_g^*).

Theorem 5.5.3. Under both PA and RA, the optimal amount of inventory to procure is $q_g^* = b/p_b$. The optimal subsidized price to charge, p_g^* , is specified as follows.

(i) Under PA, if $\mu \leq 2\lambda/(1+g)$, then $p_g^* = 0$.

(ii) Under RA: (a) If $\mu \leq (5 + 6g(1 - q_1) + 8\lambda)/(8(1 + g(1 - q_1)))$, then $p_g^* = 0$.

(b) If $\mu \in ((5 + 6g(1 - q_1) + 8\lambda)/(8(1 + g(1 - q_1))), (8 + 8\lambda(2 - q_g^*)^2 + (-4 + 8g(1 - q_1))q_g^* +$

$(1 - 2g(1 - q_1))q_g^{*2}/(8(1 + g(1 - q_1)))$], then

$$p_g^* = (1 + g(1 - q_1))(1 - (\sqrt{1 - 2g(1 - q_1) + 8\lambda}) / (2\sqrt{-1 - 2g(1 - q_1) + 2\mu + 2g(1 - q_1)\mu})).$$

(c) If $\mu > (8 + 8\lambda(2 - q_g^*)^2 + (-4 + 8g(1 - q_1))q_g^* + (1 - 2g(1 - q_1))q_g^{*2}) / (8(1 + g(1 - q_1)))$, then $p_g^* = (1 + g(1 - q_1))(1 - 1/(2 - q_g^*))$.

The condition on μ in Theorem 5.5.3(i) implies that the government prioritizes consumer welfare over revenue earnings. This condition is reasonable given the social goals of these interventions. Under this condition, it is optimal for the government to distribute the commodity for free. The reason behind this result is the following. If the government sells at any positive price, then some of the lowest-budget consumers would overspend when buying from the government, which hurts overall consumer welfare. In contrast, if the government distributes its inventory for free, then no one overspends and consumer welfare is maximized (to the extent that the government exhausts its inventory). Conversely, under RA, the government could be selling at a positive price in equilibrium. By allowing any consumer to buy at the subsidized price, the government and the trader are effectively competitors in the market. When μ is large (i.e., the government values revenue earned to a large extent; Theorem 5.5.3(ii)-(c)), the constraint that the government's selling price cannot be larger than the trader's is binding, and the two parties are selling at the same price in equilibrium. As μ decreases, the government puts more emphasis on consumer welfare compared to its revenue earnings. Ultimately, when μ is sufficiently small, the government again distributes its inventory for free as in the case of PA (Theorem 5.5.3(ii)-(a)).

Trader's optimal decision in Period 1: Anticipating the government's optimal PA and RA policies, the trader's optimal quantity decision in Period 1 is characterized in Theorem 5.5.4.

Theorem 5.5.4. Let $q_g^* = b/p_b$. (i) Under PA: If $\mu \leq 2\lambda/(1 + g)$, then (a) $q_1^* = 1/2 - g/8$ if $q_g^* \leq 1/2$; (b) $q_1^* = \frac{1 - gq_g^*(1 - q_g^*)}{2}$ if $q_g^* > 1/2$. (ii) Under RA: The trader's total revenue is piecewise concave and unimodal in q_1 . The optimal decision q_1^* is always greater than $1/2 - g/8$.

Theorem 5.5.4(i) highlights two observations for PA. First, the trader's optimal quantity released in Period 1 remains the same as in the no intervention case if the government has insufficient inventory (Theorem 5.5.4(i)-(a)). If the government only has enough inventory to satisfy the lowest-budget consumers who would anyways not buy from the trader in Period 2, then the trader has no incentive to change his quantity decision compared to that under no

intervention. Second, if the government procures a large amount of inventory in Period 2, then it is indeed beneficial for the trader to release a larger quantity in Period 1. This is because the trader anticipates losing revenue to the government in Period 2. In order to make up for the loss, the trader finds it profitable to release a larger quantity and earn additional revenue in Period 1. In sharp contrast, under RA, the trader always has to compete with the government for demand in Period 2, and hence, suffers from revenue loss. Therefore, it is always beneficial for the trader to release a larger quantity in Period 1 as compared to the no intervention case to compensate for the revenue loss in Period 2.

5.6 Policy Insights: Which Intervention is Most Effective?

Given the trader's and the government's optimal decisions under CS, PA, and RA characterized in §5.5, we next examine the effectiveness of these three interventions in reducing artificial shortage. Specifically, we answer the following questions: (i) How does the optimal quantity released by the trader in Period 1 compare to the no intervention case? (ii) How does this optimal quantity change as the government's budget for each intervention increases? We focus on the trader's Period 1 quantity decision because this decision captures the extent of artificial shortage that the trader creates in the market. The same insights continue to hold if we instead focus on the consumer welfare in Period 1. A smaller quantity released in Period 1 creates stronger artificial shortage and is undesirable. In what follows, we use the notation q_I^1 with $I \in \{\phi, CS, PA, RA\}$ to denote the trader's optimal Period 1 quantity under no intervention, CS, PA, and RA, respectively.

Effectiveness of cash subsidy: Proposition 5.6.1 answers the above two questions for CS.

Proposition 5.6.1. Recall t_1 defined in Lemma D.2.3. (i) Comparing q_{CS}^1 to q_ϕ^1 , we have: (a) $q_{CS}^1 = q_\phi^1$ if $b \leq t_1$; (b) $q_{CS}^1 \geq q_\phi^1$ if $b \in (t_1, (8 + 4g + g^2)/32]$; (c) $q_{CS}^1 < q_\phi^1$ if $b > (8 + 4g + g^2)/32$. (ii) As the government's budget, b , increases: (a) q_{CS}^1 is constant if $b < t_1$; (b) q_{CS}^1 decreases if $b > t_1$.

Figure 5.6a illustrates Proposition 5.6.1. Proposition 5.6.1(i)-(a) shows that the trader in equilibrium creates the same extent of shortage as under no intervention when the government's budget for CS is small (when $b < t_1$). In this case, the trader's revenue in Period 2 under CS is equal to that under no intervention. Thus, it is optimal for the trader to release the same

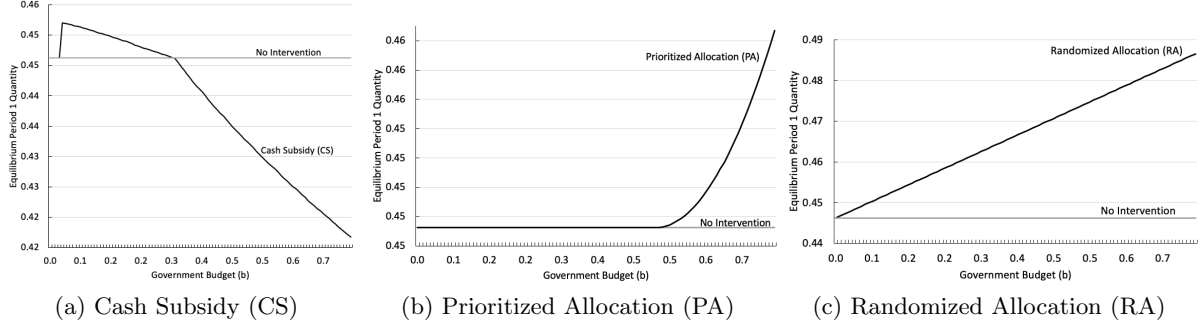


Figure 5.6: Effect of Government Budget (b) on Equilibrium Period 1 Quantity (q_I^1) under Different Government Interventions

Note: We use the following parameters in this example: $\lambda = 1.43$, $g = 0.43$, $b \in \{0, \dots, 0.75\}$.

quantity in Period 1. In other words, with a small budget, CS is ineffective in mitigating artificial shortage in the market and “improves” consumer welfare only by giving out free money. At the other extreme, when the government’s budget for CS is too large (when $b > (8 + 4g + g^2)/32$; Proposition 5.6.1(i)-(c)), the government’s intervention actually *aggravates* artificial shortage in the market. This can be seen from Proposition 5.6.1(i)-(c) that the CS policy motivates the trader to release an even smaller quantity in Period 1 as compared to no intervention. The trader does so to take advantage of consumers’ access to free government money and benefit from their increased spending power by charging a high market price in Period 2. Only when the government’s budget is in the intermediate range (Proposition 5.6.1(i)-(b)) do we observe that the CS policy effectively mitigates artificial shortage in the market by inducing a larger quantity released by the trader in Period 1. Following the same reasoning, Proposition 5.6.1(ii)-(b) highlights that an increased budget for CS is often counterproductive as “oversubsidizing” by the government inadvertently encourages the trader’s strategic hoarding behavior.

Effectiveness of supply allocation schemes: Proposition 5.6.2 describes the effectiveness of PA and RA with respect to the same two questions outlined earlier.

Proposition 5.6.2. (i) Under PA and $\mu \leq 2\lambda/(1 + g)$, we have: (a) if $b \leq p_b/2$, then

$q_{PA}^1 = q_\phi^1$ and q_{PA}^1 is constant in b ; (b) if $b > p_b/2$, then $q_{PA}^1 \geq q_\phi^1$ and q_{PA}^1 increases with b .

(ii) Under RA, $q_{RA}^1 > q_\phi^1$ and q_{RA}^1 increases with b for all $b > 0$.

Figure 5.6b illustrates Proposition 5.6.2(i). Proposition 5.6.2(i)-(a) shows that under PA, when the government’s budget is small relative to the import price (p_b), then the intervention is not effective in reducing artificial shortage in the market. This is because the consumers who buy from the government in this case would have not bought from the trader in equilibrium

anyways. Since there is no revenue loss for the trader due to the government’s intervention, it is optimal for the trader to create the same level of shortage in Period 1 as he would do under no intervention. Conversely, when the government’s budget is large (Proposition 5.6.2(i)-(b)), anticipation of the intervention would motivate the trader to sell a larger quantity in Period 1. The larger the government’s budget, the more the trader would sell in Period 1. This is because at least some of the consumers that are served by the government in Period 2 would have been served by the trader had the government not intervened. This anticipated revenue loss in Period 2 incentivizes the trader to increase sales in Period 1, thereby reducing shortage in the market. In sharp contrast to PA, Proposition 5.6.2(ii) demonstrates that the RA policy *always* decreases artificial shortage in the market, and a larger budget is always beneficial (Figure 5.6c). This is because the RA policy always creates competition between the trader and the government in Period 2. As such, the trader has to increase sales in Period 1 to compensate for the reduced revenue due to this competition.

Table 5.1 summarizes the key insights from our analysis. From the perspective of addressing the root cause of the problem – discouraging the trader’s strategic hoarding behavior – we observe that government intervention through randomized allocation is the most effective, followed by prioritized allocation, and lastly cash subsidy.⁶ For both RA and PA, a larger budget for intervention is always useful, although the increased budget for PA would materially reduce artificial shortage only after the budget exceeds a certain threshold. For CS, however, increasing budget is not always desirable. If the budget starts out being very limited, then increasing the budget to a critical threshold (t_1 characterized in Lemma D.2.3) can be helpful as it motivates the trader to substantially increase quantity in Period 1. Further increasing budget beyond that point would again encourage strategic hoarding by the trader and could eventually lead to even more hoarding than under no intervention. Thus, our analysis suggests that the effectiveness of the CS policy is more sensitive to the government’s spending as compared to either supply allocation scheme. Finally, since PA by design prioritizes selling inventory to low-income consumers, whereas low-income consumers have to compete with high-income ones for limited supply under both CS and RA, PA is most beneficial for low-income consumers. These results highlight the disparate effects of different government interventions on addressing artificial shortage of essential commodities and underscore the importance of properly designing

⁶Lemma D.2.6 confirms that in all three interventions: (a) the aggregate consumer welfare across the two periods is always larger than that under no intervention; and (b) as the government budget increases, the aggregate consumer welfare also increases.

Budget b	Cash subsidy	Prioritized allocation	Randomized allocation
Low	Stays constant	Stays constant	Decreases
High	Increases	Decreases	Decreases

Table 5.1: Effects of Increased Budget for Government Intervention on Artificial Shortage

these interventions.⁷

Preemptive intervention: While governments in practice mainly respond to artificial shortage reactively, it will be instructive to examine when preemptive intervention (i.e., before artificial shortage occurs) may be desirable. To this end, we analyze an extension in which the government procures inventory *at the beginning of* the two periods and decides how to optimally release the inventory across the two periods (Appendix D.3.5). Here, the government’s objective is to maximize total consumer welfare in both periods. We show that the government’s optimal strategy depends on whether it adopts PA or RA. Under PA, it is indeed optimal for the government to intervene *preemptively* by releasing more quantity in Period 1 than in Period 2. However, under RA, it is optimal for the government to intervene *reactively* by releasing more quantity in Period 2 than in Period 1. These contrasting results are due to the following reasons. Under PA, the government’s intervention effectively increases total supply in the market by serving lower-budget consumers who could not afford to purchase from the trader anyways. Therefore, intervening preemptively has the dual benefit of both mitigating consumers’ potential budget increase and serving more consumers, whereas intervening reactively benefits consumers primarily through serving more consumers. In contrast, under RA, the government’s intervention creates supply competition for the trader over *all* consumers. As a result, the trader responds to the government’s quantity decision in either period to ensure that the total supply in the market across both periods remains the same regardless of whether the government intervenes preemptively or reactively. In this case, if the government were to intervene preemptively, then the trader would further reduce first-period quantity to induce consumers to increase their budgets and then benefit by selling at a higher price in Period 2 (as the government would have limited supply in Period 2). Conversely, if the government

⁷Lemma D.2.7 summarizes the impact of different government interventions on the trader’s revenue as compared to the case of no intervention. In particular, the trader’s revenue always *increases* under CS; however, it either *decreases* or remains the same under PA and RA. This result is due to the following reasons. First, under both PA and RA, the trader competes with the government to sell to consumers. Therefore, *none* of the government’s budget is transferred to the trader under PA or RA. In contrast, under CS, the trader is able to extract the government’s budget from the consumers by appropriately adjusting the Period 2 price in response to the government’s CS policy.

intervenes reactively, then it creates the maximum extent of supply competition in Period 2. Such competition would force the trader to sell more quantity in Period 1, which in turn helps to mitigate consumers' budget increase in the first place, thus bringing more benefit to the consumers.

5.7 Relating Model Predictions to Empirical Observations

To further validate the empirical relevance of our model, we calibrate the model parameters using field data from agricultural markets in India and perform a counterfactual analysis to measure the effectiveness of the three government interventions studied. Using our model, we answer the following questions empirically: (i) What is the estimated level of artificial shortage in the markets? (ii) How would this shortage be affected had the government intervened with the three policies we study?

India is the second largest consumer as well as producer of onions after China ([Gummagol-math 2013](#)). We obtain monthly retail price data for onions from 33 large cities in India between 2011 and 2015 published by India's National Horticulture Board. Anecdotal evidence suggests that traders deliberately create shortage to increase prices in the downstream supply chain during the lean season ([Dash 2013](#)). Since we do not have access to retail quantity, we take a structural estimation approach ([Musalem et al. 2010](#)) for our analysis.

In particular, we assume that the observed retail prices result from the trader's optimal decisions in our model. Given this assumption, we estimate the relevant model parameters, including consumers' budget adjustment factor g , and the extent of artificial shortage in the market without government intervention. We then use the estimated model parameters to perform a counterfactual analysis to quantify the effects of government interventions on the extent of artificial shortage. In addition, we also structurally estimate several alternative models and demonstrate that the proposed model fit the data best. We next elaborate on the procedure of our analysis and the corresponding results.

Data preparation: Over 70% of the total production of onions in India is harvested between December and June ([MOAFW 2017](#)). During these months of high supply, there is little incentive for traders to create artificial shortage. Therefore, we do not consider these months in our analysis and focus on the price data from July to November every year. We further divide these 5 months into two periods: July–September corresponds to Period 1 of

our model when traders may create artificial shortage, and October–November corresponds to Period 2 of our model. That is, we take the average retail price of July–September in a year as p_1^{obs} in our model for that year, and take the average retail price of October–November as p_2^{obs} in our model. Our results are robust to an alternative splits of these 5 months where period 1 corresponds to July–August and period 2 corresponds to September–November. We adjust the retail prices by the monthly Consumer Price Index (CPI) to account for inflation and take the log of the CPI-adjusted prices to stabilize the series. We perform our analysis at the city–year level; i.e., we observe for each city every year, the price pair (p_1^{obs}, p_2^{obs}) as defined above.

Model specification: Because government interventions for onions have been sporadic and irregular, it is reasonable to consider the observed retail prices as resulted from the trader’s optimal decisions under no intervention. We thus have the following price-demand relations from §5.3: $p_1(q_1) = \alpha - \beta q_1$, and $p_2(q_1, q_2) = (\alpha - \beta q_2)(1 + g(1 - \beta q_1/\alpha))$. (In §5.3, we normalize the demand parameters α and β to be 1.) The trader chooses (q_1, q_2) , equivalently (p_1, p_2) , to maximize his total revenue from both periods. The resulting optimal retail prices are $p_1^* = \alpha(4 + g)/8$ and $p_2^* = \alpha(8 + 4g + g^2)/16$. In the estimation, we assume that log of the observed retail price pair $(\log(p_1^{obs}), \log(p_2^{obs}))$ follow a bivariate normal distribution with mean $(\log(p_1^*), \log(p_2^*))$, variance (σ_1, σ_2) , and covariance $\sigma_{1,2}$. The covariance captures potential temporal correlation in prices across the two periods. We estimate the parameters $(\alpha/\beta, g, \sigma_1, \sigma_2, \sigma_{1,2})$ by maximum likelihood estimation, taking each retail price pair in a city–year as an observation. To further strengthen the empirical relevance of our model, we contrast its goodness-of-fit to a nested model with $g = 0$; i.e., when consumers do not update their budgets in Period 2. In addition, we analyze in Appendix D.3.7 a reference price model that captures consumers’ psychological loss (gain) when prices are higher (lower) than their reference price (Tversky and Kahneman 1991). We show that $p_2^* \leq p_1^*$ in equilibrium under this model, contradicting the observed data.

Results: Table 5.2 summarizes the estimation results. We first observe that the consumer model with budget adjustment explains the data better than the nested model, based on both the likelihood ratio test (rejecting the nested model at $p < 0.0001$) and the Akaike Information Criterion (AIC).⁸ Under the consumer model with budget adjustment, the estimated budget adjustment factor (g) is significantly positive with a value of 0.43. Comparing the trader’s optimal decision q_1^* under $g = 0.43$ versus $g = 0$, we predict that traders on average withhold

⁸AIC is defined as $AIC = 2P - LL$, where P is the number of model parameters, and LL is the log-likelihood value of the estimation. Thus, AIC trades off model fit and complexity. A model with a lower AIC value is preferred (Mishra et al. 2014).

Parameters	Budget Adjustment Model	No Budget Adjustment Model ($g = 0$)
α	3328 (0.000)***	4073 (0.000)***
g	0.43 (0.049)***	—
β	—	—
σ_1	0.19 (0.020)***	0.20 (0.020)***
σ_2	0.17 (0.018)***	0.17 (0.018)***
$\sigma_{1,2}$	0.15 (0.018)***	0.15 (0.017)***
$-2\log\text{-likelihood}$	185	230
AIC	195	266

Table 5.2: Maximum Likelihood Estimation Results of Candidate Models

Note. The budget adjustment model is our model developed in §5.3. Standard errors are in parentheses. “—” means the variable is not estimated in the corresponding model. ***: $p < 0.01$, p values are derived from Wald test.

10% of inventory to create artificial shortage in the market.⁹ Taken together, the analysis suggests that consumers indeed increase their budgets for essential commodities (onions in this case) in response to shortage in the markets.

Counterfactual analysis: We finally use the estimate of g to quantify the effects of different government interventions on curbing artificial shortage if the government were to engage in the three types of interventions studied earlier. To strengthen the robustness of our results, we separately estimate the consumer budget adjustment factor g , on different subsets of our data to account for potential variation in budget adjustment factor across cities. (We create multiple subsets of our dataset by filtering cities that are in the top 25%, 50%, and 75% respectively based on their population. For each subset, we estimate g separately.) The resulting estimates range from 0.34 to 0.47. We also use public data sources to calibrate the government’s total budget (b) and the price at which it can purchase onions (p_b). We find that for most of the subsidy policies run by the government in the agricultural sector, the subsidy amount per consumer is at least 50% of the market price (Sharma 2012). Thus we allow government’s average budget per consumer to vary from 50% to 75% of the equilibrium price in Period 2 under no government intervention. Because importing from foreign markets is costly, we fix the import price (p_b) to be 1. Given these estimated and calibrated model parameters, we can predict the extent of artificial shortage in the market for each of the four scenarios: no government intervention, cash subsidy, prioritized allocation, and randomized allocation. Specifically, we calculate the amount of shortage by comparing the trader’s optimal decision q_1^* under the estimated value of

⁹From Proposition 5.3.1, $q_1^*(g) = \frac{\alpha}{2\beta}(1 - \frac{g}{4})$. Using the estimate of $g = 0.43$, we calculate the % difference between $q_1^*(0)$ and $q_1^*(0.43)$.

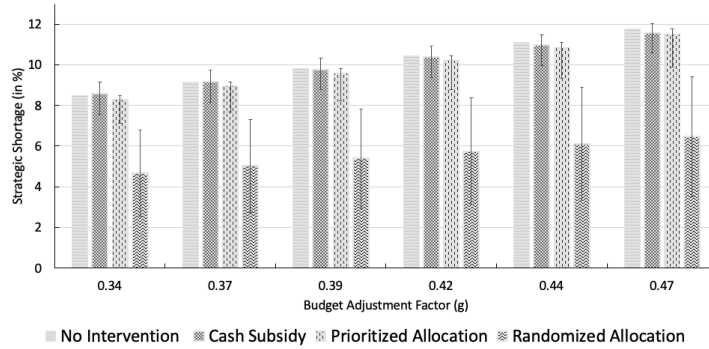


Figure 5.7: Counterfactual Analysis

g versus that under $g = 0$, using the model solutions derived in §5.3 and §5.5.

Figure ?? summarizes the (predicted) average shortage in the markets for the four scenarios under different values of consumers’ budget adjustment factor g . The “error bars” in each column of government intervention present the minimum and maximum amount of shortage predicted by our model. We highlight three observations. First, we observe from the no intervention columns that about 8–12% of the onion supply is being strategically hoarded by the traders when the government does not intervene. Second, comparing the cash subsidy and the prioritized allocation columns with the no intervention columns, we observe very similar amount of average shortage in the market. This result indicates that neither intervention is effective within the context of our data. In sharp contrast, randomized allocation always leads to less shortage or larger supply in the market. Across all parameter values, randomized allocation on average reduces the amount of shortage in the market by about 40% (from 10% under no intervention to 6% under RA). Finally, consistent with Theorem 5.6.1, we observe that cash subsidy can inadvertently *increase* artificial shortage when the government’s total budget is too large.

5.8 Conclusions

This chapter develops a behavioral game-theoretic model to investigate the phenomenon of artificial shortage in agricultural supply chains of essential commodities. Building upon the theory of mental accounting, we propose the first model to capture consumer’s behavioral factor of shortage-induced budget adjustment and its effect on consumer welfare. The analysis suggests that consumers’ budget adjustment plays a significant role in motivating traders to create artificial shortage in monopolistic agricultural markets. Next, we analyze the effectiveness of cash subsidy and supply allocation schemes, both of which are government interventions

commonly observed in practice. Leveraging price data of onions in India, it is estimated that on average, as much as 10% of the supply in the lean season is being hoarded by traders to create artificial shortage in the market across different states.

This chapter demonstrates, both analytically and empirically, that government interventions can have very different effects on artificial shortage in the market. Randomized allocation always *benefits* consumers by strictly reducing artificial shortage in the market, whereas prioritized allocation is beneficial only when government funds are sufficient. In sharp contrast, cash subsidy can in fact backfire and *aggravate* artificial shortage if the government dedicates too many funds for the intervention. These results highlight that, to effectively address artificial shortage of essential commodities, government interventions need to be carefully designed while accounting for the traders' strategic hoarding behavior and consumers' budget adjustment in response to perceived supply shortage.

The results in this chapter have important practical implications and answer critical questions pertaining to optimal interventions for food security in monopolistic markets. Recent research has highlighted the positive externality of supply allocation schemes as they increase the beneficiaries' intake of calories from all food groups (Kaul 2018). This chapter shows an additional positive externality of supply allocation policies; that is, such policies may allow governments to effectively mitigate market distortions without taking direct actions against traders (e.g., forcing the release of inventory through inspections). While cash subsidy schemes have recently gained wide popularity in many countries, we caution governments to carefully consider the strategic response of different stakeholders in the supply chain when implementing cash subsidy schemes.

Appendix A

Appendix of Chapter 2

A.1 Data Processing

We start with the group of commodities that constitute 90% of all transactions on the platform (11 commodities in total). We download the price and supply quantity data for these commodities for all regulated mandis (both in and outside Karnataka) in India from AgMarknet. We define the primary markets for a given commodity as the largest markets in which 90% of the total quantity in the data was traded. Table A.1 shows the total number of primary markets within and outside Karnataka for this group of commodities. To ensure that there are a sufficient number of both control and treatment markets for the DID analysis, we excluded all commodities with less than 10 markets either in or outside Karnataka, i.e., areca nut, copra, and dry chillies. Of the remaining commodities, a common linear pre-trend of prices between the treatment and control markets prior to UMP’s implementation is rejected in the parallel trend test for bengal gram and jowar. These two commodities are thus excluded from the analysis. Next, for each market-commodity-year, we remove daily observations for which either the supply quantity or the modal price is outside the [1%, 95%] interval of all data from that market in that year. Doing so prevents our results from being affected by potential outliers. We then aggregate the daily data to the weekly level by taking a quantity-weighted average of the daily modal prices. This aggregation is necessary because different markets are open on different days of the week. Finally, our control covariates are constructed from district-level or state-level data. Hence, we first identify the state and district of each market and then apply the corresponding covariate values.

We provide more details regarding the data sources for our analysis below.

Table A.1: Total Number of Primary Markets Within and Outside Karnataka

Commodity	Within Karnataka	Outside Karnataka
Cotton	12	162
Green Gram	17	139
Groundnut	13	58
Maize	61	177
Paddy	21	378
Tur	13	140
Areca nut	16	7
Bengal Gram	13	209
Copra	9	5
Dry Chillies	2	3
Jowar	23	132

- (i) **Wholesale market data:** We obtain data on daily modal prices and supply quantity of the commodities in our analysis for all regulated mandis in India from <http://agmarknet.gov.in> for the period of 2012–2017.
- (ii) **UMP data:** We obtain the dates of UMP implementation for different Karnataka markets from the state government to code the implementation dummy $I_{m,t}$ in our DID model. The government also provided us with the lot-level data on UMP (from April 2016 to December 2018) for the analysis of systemic differences between the high-impact and low-impact groups. The UMP data contains the following information for each lot: quantity, winning price, winning trader ID, commission agent ID, and date of transaction. Table A.2 presents summary statistics of the auction data.
- (iii) **Data for control covariates:** The data to construct the control covariates are collected from four government sources: <http://www.imd.gov.in>, <http://www.aps.dac.gov.in>, <http://www.vdsa.icrisat.ac.in>, and <http://www.niti.gov.in>. All district-level covariates come from the first three data sources, and all state-level covariates come from the last data source.

Table A.2: Summary Statistics of Auction Data from UMP

	Total lots	Total traders	Total agents	Avg. price	Avg. lot size
Cotton	1,241,400	1,374	2,935	5184	17.7
Green gram	320,536	1,377	2,497	4,800	6.88
Groundnut	737,984	1,324	2,834	4,049	3.70
Maize	560,760	4,090	6,108	1,342	111
Paddy	892,139	4,470	4,934	1,946	115
Tur	401,383	1,679	3,464	4,603	9.89

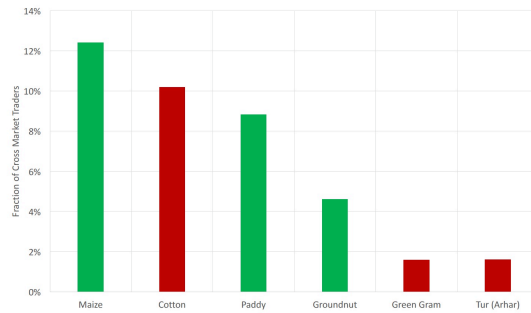


Figure A.1: The fraction of traders who have won lots in more than one market for different commodities. Green (red) indicates the high-impact (low-impact) group. The unit of observation is commodity-trader pair and the total number of observations is 14,314.

A.2 Cross-market Trading

We use the lot-level data from UMP to identify cross-market traders. All traders who have ever transacted in more than one market are considered as cross-market traders for the analysis. Figure A.1 summarizes the fraction of cross-market traders for different commodities, separating the high-impact group (green) from the low-impact group (red). We observe that cross-market trading indeed occurs to a greater extent for the high-impact group than for the low-impact group.

A.3 Dates of Integration

Figure A.2 plots the number of Karnataka markets that were integrated on UMP starting in 2014.

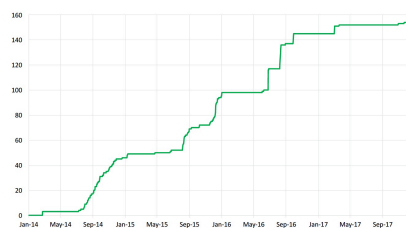


Figure A.2: Number of Karnataka markets that were integrated into UMP since 2014.

A.4 Parallel Trend Test

Table A.3 summarizes the statistics from the parallel trend test described in *Materials and Methods* in the main paper. Note that a non-statistically significant γ_2 means that a common linear pre-trend of prices between the treatment and control markets cannot be rejected.

Table A.3: Results from the Parallel Trend Test

Model	Paddy	Groundnut	Maize	Green Gram	Cotton	Tur
	2.53×10^{-5}	-2.43×10^{-5}	-3.04×10^{-5}	-1.41×10^{-5}	-4.63×10^{-6}	-8.86×10^{-5}
	(2.85×10^{-5})	(5.88×10^{-5})	(2.32×10^{-5})	(8.44×10^{-5})	(2.48×10^{-5})	(7.01×10^{-5})
Obs.	41,127	11,960	26,975	15,440	14,459	14,617

Notes. We report the coefficient estimate, γ_2 , for the parallel trend test described in the main paper. Standard errors (reported in parentheses) are clustered at the market level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. Other controls include the following district-level features: annual production, annual yield, monthly rainfall, and state-level per capita GDP. The unit of observation is week-market pair.

A.5 Robustness Tests

In our robustness analysis, we consider (i) five alternative model specifications; (ii) p -value adjustments to account for multiple hypothesis testing; (iii) alternative specifications with different control covariates. We consider our main result for a commodity to be robust if the direction and statistical significance of the coefficient for the implementation dummy are consistent between the main model and these robustness tests. We confirm that this is the case for all six commodities.

A.5.1 Alternative Specifications

- (i) To account for potential learning by market participants, we reestimate the main DID model (model [1]) with $I_{m,t} = 1$ if market m has been integrated to UMP for at least 3 months by week t and 0 otherwise (Table A.4 column 3).
- (ii) We reestimate model [1] by adding state-specific trends (Table A.4 column 4). This approach is advocated by ref. 1 as a robustness analysis for the DID method.
- (iii) To ensure that the results are not driven by the presence of small states in the control set, we reestimate model [1] with a subset of the data that only contains the largest states (in terms of the total number of markets) which cover 95% of all the markets in our data (Table A.4 column 5). These states include Maharashtra, Madhya Pradesh, Uttar Pradesh, Punjab, Tamil Nadu, Gujarat, Andhra Pradesh, Karnataka, Rajasthan, Telangana, Chhattisgarh, Haryana, Odisha, West Bengal, Kerala, and Jharkhand.
- (iv) To account for potential spillovers across markets within the same district, we aggregate our data at the district level and reestimate model [1] (Table A.4 column 6). The implementation dummy $I_{d,t}$ is equal to 1 if at least one of the markets in district d has been

Table A.4: Estimated Impact of UMP on the Prices of Different Commodities

Commodity	Main model	3-Month lag	State-specific trend	Large states only	District-level	Joint regression
Paddy	0.051*** (0.013)	0.058*** (0.013)	0.035*** (0.012)	0.051*** (0.013)	0.064*** (0.013)	0.041*** (0.016)
Observations	131,903	131,903	131,903	124765	56,135	–
Groundnut	0.036*** (0.014)	0.040*** (0.014)	0.047** (0.019)	0.036*** (0.014)	0.037*** (0.015)	0.062*** (0.021)
Observations	34,944	34,944	34,944	34,788	18,088	–
Maize	0.035*** (0.007)	0.043*** (0.007)	0.019* (0.011)	0.035*** (0.007)	0.066*** (0.024)	0.037*** (0.008)
Observations	89,478	89,478	89,478	88,796	42,355	–
Cotton	0.015 (0.012)	0.015 (0.011)	-0.023 (0.010)	0.015 (0.012)	0.027 (0.017)	0.027 (0.01)
Observations	46,055	46,055	46,055	44,690	18,001	–
Tur	0.008 (0.016)	0.008 (0.015)	0.011 (0.026)	0.009 (0.016)	-0.014 (0.020)	-0.040 (0.031)
Observations	52,953	52,953	52,953	51,295	28,312	–
Green Gram	-0.009 (0.013)	-0.007 (0.014)	-0.038 (0.030)	-0.008 (0.013)	0.017 (0.020)	0.004 (0.01)
Observations	51,742	51,742	51,742	49,581	28,766	–

Notes. The dependent variable is the logarithm of modal prices in all models. Standard errors (in parentheses) are clustered at the market level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. In all models, we control for district-level yearly production, yearly yield, monthly rainfall, and state-level per capita GDP, as well as market and week fixed effects. The analysis is at the week-market level and the total number of observations in the last column is 407,705.

integrated to UMP at week t and 0 otherwise.

- (v) We estimate a joint regression with all six commodities together in which each commodity dummy interacts with the implementation dummy. We use linear hypothesis test to perform pairwise comparisons among these coefficient estimates. We confirm that the coefficient estimate for each of groundnut, paddy, and maize is significantly greater than that for each of cotton, green gram, and tur (Table A.4 column 7).

The results from all these specifications are included in Table A.4 and confirm the robustness of our results.

A.5.2 Multiple Hypothesis Testing

We adjust the p -values using the Bonferroni method [McDonald \(2009\)](#) to confirm that our conclusions continue to hold after accounting for multiple comparisons across commodities and

outcomes in the analysis. In particular, we perform two adjustments: (i) we adjust the p -values for the same outcome across commodities (Table A.5), (ii) we adjust the p -values for the same commodity across outcomes (Table A.6).

Table A.5: Unadjusted and Adjusted p -Values (Accounting for Multiple Comparisons across Commodities for the Same Outcome)

Commodity	Modal price		Minimum price		Maximum price	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Paddy	0.000	0.000	0.000	0.000	0.000	0.000
Groundnut	0.010	0.080	0.370	1.000	0.000	0.000
Maize	0.000	0.000	0.000	0.000	0.000	0.000
Cotton	0.200	1.000	0.290	1.000	0.000	0.000
Green Gram	0.490	1.000	0.700	1.000	0.620	1.000
Tur	0.620	1.000	0.400	1.000	0.800	1.000
Bengal Gram	0.940	1.000	0.990	1.000	0.200	1.000
Jowar	0.000	0.000	0.000	0.000	0.000	0.000

Notes. We report the unadjusted and adjusted p -values of γ estimates for the implementation dummy in the DID Model (Eq. [1]) of the main paper.

Table A.6: Unadjusted and Adjusted p -Values (Accounting for Multiple Comparisons across Outcomes for the Same Commodity)

Commodity	Modal price		Minimum price		Maximum price	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Paddy	0.000	0.000	0.000	0.000	0.000	0.000
Groundnut	0.010	0.030	0.370	1.000	0.000	0.000
Maize	0.000	0.000	0.000	0.000	0.000	0.000
Cotton	0.200	0.600	0.290	0.870	0.000	0.002
Green Gram	0.490	1.000	0.700	1.000	0.620	1.000
Tur	0.620	1.000	0.400	1.000	0.800	1.000
Bengal Gram	0.940	1.000	0.990	1.000	0.200	1.000
Jowar	0.000	0.000	0.000	0.000	0.000	0.000

Notes. We report the unadjusted and adjusted p -values of γ estimates for the implementation dummy in the DID Model (Eq. [1]) of the main paper.

A.5.3 Alternative Control Covariates

We reestimate the DID model (model [1]) in three additional specifications with different sets of control covariates. In particular, (i) we only control for production but do not control for quantity in a market, (ii) we do not control for production or quantity, (iii) we do not control for any covariates. Table A.7 summarizes the estimates of γ (for the implementation dummy) in our main model and the above three specifications.

Table A.7: Estimated Impact of UMP: Specifications with Different Control Covariates

Commodity	γ estimates			
Groundnut	0.036*** (0.014)	0.036*** (0.014)	0.037*** (0.014)	0.028* (0.015)
Paddy	0.051*** (0.013)	0.053*** (0.012)	0.059*** (0.012)	0.060*** (0.013)
Maize	0.035*** (0.007)	0.034*** (0.007)	0.034*** (0.007)	0.038*** (0.007)
Cotton	0.015 (0.012)	0.015 (0.012)	0.015 (0.012)	0.014 (0.012)
Green Gram	-0.009 (0.013)	-0.008 (0.013)	-0.009 (0.012)	0.001 (0.012)
Tur	0.008 (0.016)	0.009 (0.017)	0.009 (0.016)	0.015 (0.016)
Other Controls	Y	Y	Y	-
Rainfall	Y	Y	Y	-
Production	Y	Y	-	-
Quantity	Y	-	-	-

Notes. Standard errors (in parentheses) are clustered at the market level. γ is the coefficient for the implementation dummy in Model [1]. “-” means the corresponding variable is not present in the model. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$. Other controls include district-level annual yield and state-level per capita GDP. The unit of observation is week-market pair. Total observations for each commodity are the following: groundnut 34,944; paddy 131,903; maize 89,478; cotton 46,055; green gram 51,742; and tur 52,953.

A.6 Falsification Test

As another robustness analysis, we reestimate model [1] with the data prior to UMP’s implementation and assuming randomly selected placebo dates were the implementation dates. We repeat this test 1,000 times with a randomly drawn placebo implementation date each time. Table A.8 presents summary statistics on the coefficient estimates of the implementation dummy across these 1,000 runs and compares to the estimated impact with the true implementation date (the last column). We observe that for the three commodities that UMP has yielded a significant positive impact, the estimated effects from the falsification tests are significantly smaller. For the remaining three commodities, we do not find any statistically significant treatment effects in the falsification tests.

A.7 Farmer Heterogeneity

To further explore potentially heterogeneous impacts of UMP on market prices for farmers with different characteristics, we obtain data on the following characteristics at the district

Table A.8: Estimated Impact in 1,000 Replications of Falsification Tests

Commodity	Mean	Median	Minimum	Maximum	95% confidence interval	Estimated impact with true implementation date
Groundnut	0.005	0.005	-0.065	0.008	[-0.053, -0.005]	0.036***
Paddy	-0.014	-0.014	-0.035	0.028	[-0.012, 0.021]	0.051***
Maize	-0.029	-0.029	-0.036	0.013	[-0.029, -0.002]	0.035***
Cotton	-0.005	-0.005	-0.029	0.017	[-0.020, 0.009]	0.014
Green Gram	-0.016	-0.018	-0.081	0.048	[-0.057, 0.024]	-0.009
Tur	-0.033	-0.034	-0.080	0.026	[-0.063, 0.002]	0.008

Note: ***: $p < 0.01$.

level from the Village Dynamics in South Asia (VDSA) database provided by the International Crop Research Institute for the Semiarid Tropics (ICRISAT): annual fertilizer usage per unit of farmland, the fraction of small farmers (defined as those having less than 2 hectares of farmland) among all farmers in the district, literacy rate among rural population, and total length of road in the district (normalized by the total area of the district). To examine whether the launch of UMP has differential impact along the above four farmer characteristics, we estimate the following model for each characteristic and each commodity separately.

$$\log(P_{m,t}) = \gamma_0 I_{m,t} + \gamma_H I_{m,t} \times C_m^H + \delta \log(Q_{m,t}) + \Theta X_{m,t} + \alpha_m + \beta_t + \epsilon_{m,t}. \quad (\text{A.1})$$

The dependent variable is the logarithm of the modal price in market m at week t . C_m^H is the key variable that captures farmer heterogeneity. For each characteristic, $C_m^H = 1$ if the value of the characteristic in the district where market m is located is above the median of the distribution of that characteristic across all districts in the state, and $C_m^H = 0$ otherwise. In other words, a statistically significant value of γ_H would indicate that farmers in districts at the top 50% of a certain characteristic experience a significantly different impact from UMP compared to the rest of the farmers. All other variables follow from Eq. [1] in the paper.

Table A.9 summarizes the estimates of γ_0 and γ_H for each farmer characteristic and each commodity. We make three observations. First, we observe a statistically significant differential

Table A.9: Estimated Differential Impacts of UMP by Farmer Characteristics

Commodity	Fertilizer usage		Fraction of small farms		Length of road		Literacy rate	
	γ_0	γ_H	γ_0	γ_H	γ_0	γ_H	γ_0	γ_H
Groundnut	0.013 (0.016)	0.054** (0.021)	0.050** (0.021)	-0.027 (0.023)	0.040** (0.016)	-0.009 (0.025)	0.068*** (0.017)	0.057*** (0.021)
Observations	33542		34173		30925		34522	
Paddy	0.015 (0.012)	0.063*** (0.022)	0.056** (0.026)	-0.008 (0.028)	0.048*** (0.017)	0.006 (0.025)	0.038*** (0.014)	0.023 (0.023)
Observations	124098		124910		101531		125323	
Maize	0.039*** (0.012)	-0.009 (0.013)	0.026*** (0.007)	0.018 (0.013)	0.025*** (0.007)	0.019 (0.014)	0.032*** (0.010)	0.005 (0.013)
Observations	87685		87694		69205		87948	
Cotton	0.005 (0.015)	0.020 (0.023)	-0.020 (0.011)	0.073*** (0.021)	0.006 (0.011)	0.023 (0.027)	0.004 (0.016)	0.022 (0.024)
Observations	40813		41808		37114		42039	
Green Gram	-0.013 (0.014)	0.011 (0.026)	-0.006 (0.019)	-0.006 (0.024)	-0.017 (0.013)	0.020 (0.028)	-0.010 (0.019)	0.001 (0.024)
Observations	47243		48272		40530		48899	
Tur	0.012 (0.021)	-0.007 (0.031)	0.024 (0.016)	-0.047 (0.037)	0.019 (0.016)	-0.030 (0.038)	-0.011 (0.016)	0.067 (0.037)
Observations	51283		51599		38377		51772	

Notes. Standard errors (clustered at the market level) are reported in parentheses. ***: $p < 0.01$; **: $p < 0.05$. Other controls include the following district level features: annual production, annual yield, monthly rainfall and state level per capita GDP. The unit of observation is week-market pair.

impact along the characteristic of fertilizer usage per unit of farmland for groundnut and paddy. In particular, groundnut and paddy farmers whose fertilizer usage is above the median benefit from a 5.4% and 6.3% price increase due to UMP, whereas those with below-median fertilizer usage do not benefit from a significant price increase. These results support the hypothesis that farmers producing higher-quality commodities (groundnut and paddy specifically) benefit more from the implementation of UMP.

Note however that for maize, we do not find differential impact along the characteristic of fertilizer usage, although we observe a significant overall price increase. This can be because compared to groundnut and paddy, price for maize is less sensitive to its quality as maize is predominantly used as cattle feed in the country (Singh 2014). We find some evidence in line with this hypothesis in the UMP data. Specifically, we should expect to see more price variation across lots within the same day for commodities whose prices are more sensitive to quality. Thus, we calculate the coefficient of variation (CV) in the market prices for each of

these three commodities, i.e., we divide the standard deviation of prices across all lots in a given day and a given market by the average price across these lots. The average CVs in the market prices (across all days and markets) for groundnut, paddy, and maize are 0.135, 0.141, and 0.043. Furthermore, we run a regression with the dependent variable being the CV in market prices at the market-day level, the independent variables being an indicator for groundnut and an indicator for paddy (i.e., maize is the baseline group), and clustering standard errors at the market level. We observe that the coefficient estimates for the groundnut and paddy dummies are both positive and statistically significant ($\beta = 0.092$ and 0.098 , $p < 0.001$). These results suggest that quality matters more in determining market prices for groundnut and paddy than for maize. As a result, UMP results in a significantly higher impact for groundnut and paddy farmers who produce higher-quality commodities, whereas there is no differential impact among maize farmers.

Second, we observe in the case of cotton that the modal price in markets with a larger fraction of small farmers significantly increases due to UMP. This result suggests that UMP may benefit farmers with weaker bargaining power more. Third, in the case of groundnut, we find a significant positive (negative) impact on the modal price in markets with a smaller (larger) fraction of literate farmers. A linear hypothesis test shows that the sum of γ_0 and γ_H in this case is not statistically significant ($p = 0.52$). In other words, UMP yields a significant price increase for less literate farmers but does not yield a significant impact for more literate farmers. This result again suggests that UMP may benefit weaker farmers more (e.g., illiterate farmers are less likely to seek price information and thus more likely to be exploited prior to UMP). Due to the sporadic nature of these results, we consider them as mainly exploratory findings in this research.

A.8 Comparing High-impact and Low-impact Groups

In this section, we do additional analysis using the UMP data to confirm H2-H4. In particular, we examine how the level of supply concentration among agents (H2), the level of market concentration among traders (H3), and the fraction of quantity traded through auctions (H4) differ between commodities in the high-impact versus low-impact group. Table A.10 presents the correlation matrix among these three features at the commodity-market-week level. We

Table A.10: Correlation Matrix among Market Features for H2–H4

	Market concentration among traders (H3)	Fraction of quantity traded through auction (H4)
Supply concentration among agents (H2)	0.37	−0.24
Market concentration among traders (H3)	–	0.12

Note. The unit of analysis is at the commodity-market-week level.

Table A.11: Systemic Differences between High-impact and Low-impact Groups

Commodity	Supply concentration among agents (H2)	Market concentration among traders (H3)	Fraction of quantity traded through auction (H4)
γ_0	0.35***	0.44***	0.34***
γ_H	-0.05**	-0.12***	-0.13***
Observations	23,308	23,308	23,358

Notes. Values reported are the estimates of γ_0 and γ_H in Eq. [A.2]. Standard errors (in parentheses) are clustered at the market level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. The unit of observation is week-market pair.

estimate the following regression for each of these features separately:

$$\text{Market Factor}_{m,w,c} = \gamma_H I_c + \gamma_0 + \epsilon_{m,w,c}. \tag{A.2}$$

The dependent variable is one of the market factors in market m at week w for commodity group c . I_c is the high-impact group dummy: $I_c = 1$ if commodity c is paddy, groundnut, or maize, and 0 otherwise. A statistically significant positive (negative) value of γ_H would indicate that the market factor for commodities in the high-impact group is significantly larger (smaller) compared to those in the low-impact group. Table A.11 summarizes the estimates of γ_H and γ_0 for different market factors. In line with H2, H3b, and H4, we find that γ_H is negative and statistically significant for each of the three factors.

A.9 Structural Changes in the Markets

In this section, we examine how various factors of market structure have evolved over time post UMP’s implementation. Using the UMP data from April 2016 to December 2018, we examine whether any of the following factors has changed significantly since April 2016: the number of traders in a market, the number of commission agents in a market, the level of supply concentration among agents, the level of market concentration among traders, the fraction of quantity traded through auctions, the fraction of quantity won by “home” traders, and the average number of bids per lot. Since we do not know the location of a trader, we define the “home market” for a trader as the market from which the trader has purchased the most quantity. Hence, for a given market, its home traders are those traders whose home market is the current market. We estimate the following regression for each of these factors and each commodity separately:

$$\text{Market Factor}_{m,w} = \gamma w + \beta_{\text{month}} + \alpha_m + \epsilon_{m,w}. \quad (\text{A.3})$$

The dependent variable is one of the market factors in market m at week w . The key independent variable is w , the numeric week. We control for month fixed effects (β_{month}) and market fixed effects (α_m). The coefficient of interest is γ . A positive (negative) and significant value of γ would indicate that the corresponding market factor has increased (decreased) over time.

Table A.12 summarizes the estimates of γ for different commodities and market factors. The major significant change we find is regarding the average number of bids per lot (the last column in Table A.12). We observe a significant increase in this factor for the four commodities whose modal price or maximum price has significantly increased due to UMP (groundnut, paddy, maize, and cotton). In contrast, this factor has significantly decreased for green gram and tur. These results provide additional evidence that UMP has resulted in increased competition among traders for the high-impact group. Besides this factor, the other significant results are: (i) market concentration among traders for paddy has decreased over time, and (ii) the fraction of quantity traded through auctions for maize and cotton has also decreased. We expect that these changes are likely to be gradual as opposed to a sudden change immediately following the launch of UMP. Hence, using the UMP data to approximate market structure differences across commodities pre-implementation still has merits.

Table A.12: Change in Market Structure over Time

Commodity	No. of traders per market	No. of agents per market	Supply conc. agents	Market conc. traders	Fraction of quantity auctioned	Quant. won by home traders	Avg. no. of bids per lot
Groundnut	0.001 (0.001)	0.001 (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0009* (0.0004)
Paddy	-0.002 (0.002)	-0.002 (0.004)	-0.001 (0.001)	-0.001*** (0.000)	0.000 (0.000)	0.0001 (0.0001)	0.0014*** (0.0000)
Maize	0.001 (0.001)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	-0.0002*** (0.000)	0.0002 (0.0001)	0.0033*** (0.0006)
Cotton	-0.0004 (0.0011)	-0.0008 (0.0013)	0.0001 (0.0003)	-0.0001 (0.0003)	-0.0006* (0.0003)	-0.0001 (0.0001)	0.0018*** (0.0006)
Green Gram	0.000 (0.001)	0.002 (0.002)	-0.001 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0002 (0.0001)	-0.0022*** (0.0008)
Tur	0.000 (0.001)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0001 (0.0001)	-0.0011** (0.0005)

Notes. Values reported are the estimates of γ in Eq. [A.3]. Standard errors (in parentheses) are clustered at the market level. ***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.1$. The analysis is at the week-market level. Total observations for each commodity is the following. Groundnut: 3631, Paddy: 8629, Maize: 10867, Cotton: 5299, Green Gram: 3081, Tur: 3637.

Table A.13: Change in the Fraction of Quantity Won by Home Traders over Time

	Groundnut	Paddy	Maize	Cotton	Green gram	Tur
Estimate of γ in Eq. [A.3]	0.000	0.0001	0.0002	-0.0001	0.0002	0.0001
(Standard error)	(0.000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations	3,631	8,629	10,867	5,299	3,081	3,637

Note. Standard errors are clustered at the market level. The unit of observation is week-market pair.

A.9.1 Spillovers

We use the UMP data (only available for Karnataka markets on UMP from April 2016 to December 2018) to perform three analyses to examine whether traders might have shifted among different markets. First, we investigate whether the fraction of quantity won by home traders in a market has changed over time. The definition of a home trader follows as before. We estimate Eq. [A.3] with the fraction of quantity won by home traders as the dependent variable. Table A.13 summarizes the estimates of γ for all commodities. We observe no statistically significant change in this fraction over time. Importantly, across the six commodities, 88%–98% of the traders have traded in only one market in the entire data.

Second, we consider the subset of traders who have purchased in multiple markets, which

Table A.14: Change in the Fraction of Quantity Purchased at Home Market by Cross-Market Traders over Time

	Groundnut	Paddy	Maize	Cotton	Green gram	Tur
Estimate of γ in Eq. [A.4]	-0.0002	0.0004	0.0005**	-0.0001	-0.0002	0.0012
(Standard error)	(0.0004)	(0.0003)	(0.0002)	(0.0003)	(0.0005)	(0.0009)
Observations	2,178	17,404	20,998	6,531	657	670

Notes. Standard errors are clustered at the market level. The unit of observation is trader-week pair. **: $p < 0.05$.

we call “cross-market” traders, and investigate whether they might have shifted among different markets. In total, there are 61, 395, 508, 140, 22, 27 cross-market traders for groundnut, paddy, maize, cotton, green gram, and tur respectively, and they collectively account for 792,469 out of 4,154,202 transactions (19%) in the data. We calculate for each cross-market trader in each week, within the total quantity purchased by the trader, what is the fraction that he purchased from his home market. We then analyze whether this fraction has changed significantly over time for each commodity separately by estimating the following model:

$$\text{Fraction of Home Market Quantity}_{b,w} = \gamma w + \beta_{\text{month}} + \alpha_m + \epsilon_{b,w}, \quad (\text{A.4})$$

where b is the index for trader, w is the numeric week, and we control for month fixed effects (β_{month}) and market fixed effects (α_m). Table A.14 presents the estimates of γ in Eq. [A.4] for all commodities. The only statistically significant result we observe is that for maize, the fraction of quantity that cross-market traders purchase at their home market has increased over time.

Third, we analyze whether the traders’ home markets have changed. To do so, we redefine a trader’s home market on a monthly basis. That is, a trader’s home market in a month is defined as the market in which the trader purchased the most quantity among all markets in that month. We then count the number of unique monthly home markets for each trader of each commodity. Figure A.3 presents the distribution of the number of unique monthly home markets across traders for each commodity. We observe that across the six commodities, 93%–99% of traders have only one main market from which they consistently purchase the most quantity. Taking all three analyses together, we do not observe a strong spillover or frequent market shifting on the trader side in our data.

We next turn to the farmer side. In the absence of farmer IDs in the UMP data, we cannot

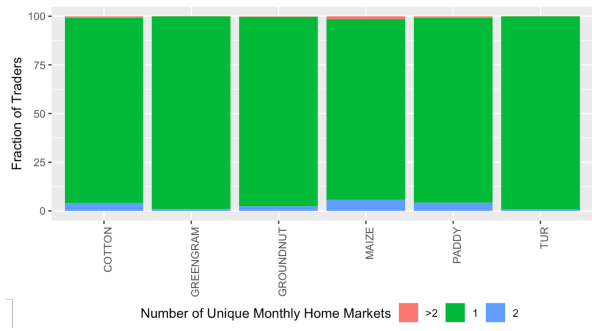


Figure A.3: Distribution of the number of unique monthly home markets across traders for each commodity. The unit of observation is trader-commodity pair and the total number of observations is 14,314.

Table A.15: Change in Quantity Distribution among Karnataka Markets over Time

	Groundnut	Paddy	Maize	Cotton	Green gram	Tur
Estimate of γ_0 in Eq. [A.5]	0.001	-0.002	-0.001	0.004	0.011	-0.009
(Standard error)	(0.004)	(0.002)	(0.002)	(0.009)	(0.018)	(0.013)
Observations	6,200	11,863	14,397	5,575	3,092	3,643

Note. Standard errors are clustered at the market level. The unit of observation is week-market pair.

perform a similar analysis at the farmer level to analyze spillovers on the supply side. Instead, we investigate whether the quantity distribution among Karnataka markets has changed over time. In particular, we analyze whether there exists a significant change in the fraction of the total quantity in the state being traded in a given market when the market becomes integrated to UMP. We estimate the following model considering all markets in Karnataka:

$$f_{m,w}^q = \gamma_0 I_{m,w} + \alpha_m + \beta_w + \epsilon_{m,w}. \tag{A.5}$$

The variable $f_{m,w}^q$ is the fraction of total quantity in the state that is traded in market m at week w . $I_{m,w}$ is the implementation dummy that is equal to 1 if market m has been integrated on UMP at week w and 0 otherwise. We control for market fixed effects (α_m) and week fixed effects (β_w). Table A.15 summarizes the estimates of γ_0 in Eq. [A.5] for all commodities. We do not observe any statistically significant change in the quantity distribution across markets for any of the commodities. Therefore, we do not find evidence of substantial market shifting on the farmer side due to UMP.

A.10 Normalized Modal Prices of High-Impact and Low-Impact Group

Figure A.4 presents the normalized average weekly modal prices of the high-impact group (Fig. A.4a) and low-impact group (Fig. A.4b) in Karnataka markets (dark green) and in markets outside Karnataka (yellow) from our data. The normalization is performed for each commodity by subtracting the mean price from each price observation and then dividing the difference by the standard deviation of the prices. We then average these normalized prices across commodities within each group. The first, second, and third vertical dashed lines indicate the time when the first market, 50% of all markets in Karnataka, and the last market were integrated into UMP.

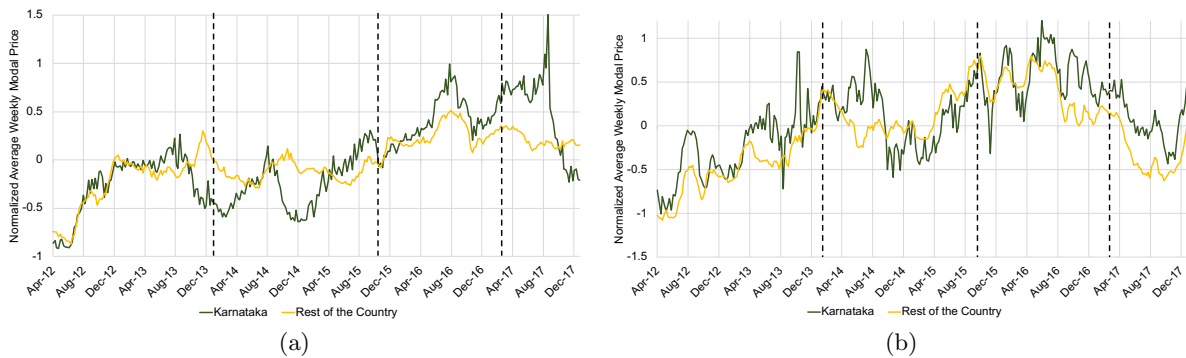


Figure A.4: Normalized average weekly modal prices of the high-impact group (Fig. A.4a) and the low-impact group (Fig. A.4b) in Karnataka markets (dark green) and in markets outside Karnataka (yellow). The total number of observations in each figure is 600.

Appendix B

Appendix of Chapter 3

B.1 Operational Feasibility of Dutch Auction

This section focuses on evaluating the operational feasibility of Dutch auction that can also potentially intensify anticipated loser's regret ([Filiz-Ozbay and Ozbay 2007](#)). In a Dutch auction, the auctioneer begins with a high asking price, and lowers it until some participant accepts the price. One particularly relevant analysis to understand the operational feasibility of Dutch auction for UMP is that of eNAM launched by the central government of India. eNAM and UMP share close similarities in market operations and cultural characteristics but differ in the auction mechanism used on the platform. eNAM uses English auctions to determine market prices. In an English auction, traders sequentially increase their bids on a lot until only one trader remains. This last trader wins the lot and pays the current highest bid. However, traders initially participating in eNAM complained that having to constantly monitor and potentially adjust bids for hundreds of lots involved too much effort, seriously disrupted other aspects of their operations (e.g., arranging post-auction logistics), and became operationally infeasible during peak seasons. As a result, trader participation on the eNAM was not sustained, and the platform has not realized its intended impact to date ([Business Line 2019](#)). Similar to English auction on eNAM, implementing Dutch auction on UMP would require substantial trader efforts to monitor bids for many lots. In light of such participation constraints on the traders' side, and due to the requirements of significant updates to the UMP platform, the Karnataka government rejected the idea of adopting a Dutch auction.

B.2 Additional Details of the Behavioral Auction Model

We provide additional details here regarding how we model the traders' bidding behavior in a two-stage auction. Specifically, let $q + H$ be the highest bid in the first stage. We assume that the top-rank trader bids $q + H$ again, and the second-rank trader bids $q + H$ if and only if his valuation $q + v$ is greater than $q + H$ (i.e., he makes a positive profit when bidding). Given this assumption, we characterize the expected payoff of a trader whose valuation is $q + v$ and who bids $q + b$ in the first stage, taking anticipated non-qualification regret into account. Let $B^{(1)}$, $B^{(2)}$, and $B^{(3)}$ denote the top-, second-, and third-rank bids in the first stage. Similarly, let $V^{(1)}$, $V^{(2)}$, and $V^{(3)}$ denote the top-, second-, and third-rank valuations.

There are three possibilities. If the trader is ranked first, then he wins with probability $1/2$ (1) if the second-rank trader bids (does not bid) again in the second stage (the first two terms in Equation (B.1)). In both the cases, his revenue is $(v - b)$. If the trader is ranked second, then he wins with probability $1/2$ if he bids again in the second round (the third term in Equation (B.1)). His revenue in this case is $(v - H)$. If the trader does not qualify, then he feels non-qualification regret (see §3.3.2), with magnitude of the regret proportional to the lost revenue from non-qualification, i.e., $(v - H)^+$ (Filiz-Ozbay and Ozbay 2007, Engelbrecht-Wiggans and Katok 2008). Equation (B.1) summarizes the trader's expected payoff given these three possibilities. The parameter $\lambda \geq 0$ captures the intensity of non-qualification regret.

$$\begin{aligned} \pi(q, v, b) = & \left((v - b)E[\mathbb{1}\{q + b > V^{(2)}\} | B^{(1)} = q + b] + \frac{(v - b)}{2}E[\mathbb{1}\{V^{(2)} > q + b\} | B^{(1)} = q + b] \right) P(B^{(1)} = q + b) \\ & + E\left[\frac{q + v - B^{(1)}}{2} \mathbb{1}\{q + v > B^{(1)}\} | B^{(2)} = q + b\right] P(B^{(2)} = q + b) \\ & - \lambda E[(q + v - B^{(1)})^+ | B^{(3)} = q + b] P(B^{(3)} = q + b) \end{aligned} \tag{B.1}$$

We follow Filiz-Ozbay and Ozbay (2007) to prove Theorem 3.4.2 and Proposition 3.4.3. To do so, we first show the following lemmas.

Lemma B.2.1. Incentive compatible bidding strategy in the two-stage auction with anchoring and anticipated regret is a strictly increasing function when $\lambda > 1/2$.

Lemma B.2.2. The local and global incentive constraints in a two-stage auction with anchoring and anticipated regret are equivalent.

B.3 Proofs

Proof of Lemma 3.4.1 (Filiz-Ozbay and Ozbay 2007)

The expected profit of a trader who bids b when his valuation is $q + v$ and the other two traders bid B_1 and B_2 is $\pi(v, q, b) = E[(v + q - b)\mathbb{1}\{b > B_1 \cap b > B_2\}] - \lambda_L(v - \max B_1, B_2)^+$. The f.o.c. is $2(1 + \lambda)(v - b) = vb'$. This is because the trader wins only if his bid is higher than the other two bids and feels loser's regret if he does not win. It is easy to check that under a first price sealed bid auction, the symmetric equilibrium bidding function, $B^*(v, q)$, that maximizes $\pi(v, q, b)$ is given by $B^*(v, q) = q + \frac{2 + 2\lambda_L}{3 + 2\lambda}v$.

Proof of Lemma B.2.1

Let $b(\cdot) : [0, 1] \rightarrow \mathbb{R}$ be the equilibrium bidding function for a two-stage auction. Let two valuations v_1 , and v_2 be such that $v_2 > v_1$, and b^1 and b^2 be the corresponding bids. Finally, let b^w be the maximum bid. Let us define \mathcal{E}_i^1 to be the event that bidder is ranked first with bid b_i and no one else bids again. Similarly, let \mathcal{E}_i^2 be the event that bidder is ranked first with bid b_i but the second ranked bidder bids again. Let \mathcal{E}_i^3 be the event that bidder is ranked second with bid b_i and he bids again. Let \mathcal{E}_i^4 to be the event that bidder is ranked third and he feels regret from non qualification. Let $\mathcal{E}'_{j,i}$ be the event that the trader is ranked j when he bids b_i . Let $\bar{\mathcal{E}}_i$ be the event that the trader qualifies when he bids b_i .

Since we consider incentive compatible bids, following Filiz-Ozbay and Ozbay (2007), we should have,

$$\begin{aligned} P(\mathcal{E}_1^1)(v_1 - b_1) + P(\mathcal{E}_1^2)\frac{(v_1 - b_1)}{2} + P(\mathcal{E}_1^3|b^w)\frac{(v_1 - b^w)^+}{2} - \lambda P(\mathcal{E}_1^4|b^w)(v_1 - b^w) \geq \\ P(\mathcal{E}_2^1)(v_1 - b_2) + P(\mathcal{E}_2^2)\frac{(v_1 - b_2)}{2} + P(\mathcal{E}_2^3|b^w)\frac{(v_1 - b^w)^+}{2} - \lambda P(\mathcal{E}_2^4|b^w)(v_1 - b^w) \end{aligned} \quad (\text{B.2})$$

and,

$$\begin{aligned} P(\mathcal{E}_2^2)(v_2 - b_2) + P(\mathcal{E}_2^2)\frac{(v_2 - b_2)}{2} + P(\mathcal{E}_2^3|b^w)\frac{(v_2 - b^w)^+}{2} - \lambda P(\mathcal{E}_2^4|b^w)(v_2 - b^w) \geq \\ P(\mathcal{E}_1^2)(v_2 - b_1) + P(\mathcal{E}_1^2)\frac{(v_2 - b_1)}{2} + P(\mathcal{E}_1^3|b^w)\frac{(v_2 - b^w)^+}{2} - \lambda P(\mathcal{E}_1^4|b^w)(v_2 - b^w) \end{aligned} \quad (\text{B.3})$$

Adding the two inequalities above and rearranging the terms, we get:

$$\begin{aligned}
 & (P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1))(v_2 - v_1) + (P(\mathcal{E}_2^2) - P(\mathcal{E}_1^2))\frac{(v_2 - v_1)}{2} + (P(\mathcal{E}_2^3|b^w) - P(\mathcal{E}_1^3|b^w))\frac{(v_2 - v_1)}{2} \\
 & + \lambda(P(\mathcal{E}_1^4|b^w) - P(\mathcal{E}_2^4|b^w))(v_2 - v_1) \geq 0
 \end{aligned} \tag{B.4}$$

Note that since $P(\mathcal{E}'_{1,i}) = P(\mathcal{E}_i^1) + P(\mathcal{E}_i^2)$, we have the following:

$$\begin{aligned}
 & \frac{(v_2 - v_1)}{2}(P(\mathcal{E}'_{1,2}) - P(\mathcal{E}'_{1,1})) + \frac{(v_2 - v_1)}{2}(P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1)) + \frac{(v_2 - v_1)}{2}(P(\mathcal{E}_2^3|b^w) - P(\mathcal{E}_1^3|b^w))\frac{(v_2 - v_1)}{2} \\
 & + \lambda(v_2 - v_1)(P(\mathcal{E}'_1|b^w) - P(\mathcal{E}'_2|b^w)) \geq 0 \\
 \implies & \frac{(v_2 - v_1)}{2}(P(\mathcal{E}'_{1,2}) - P(\mathcal{E}'_{1,1})) + \frac{(v_2 - v_1)}{2}(P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1)) + \frac{(v_2 - v_1)}{2}(P(\mathcal{E}'_{2,2}) - P(b^w > v_2) - (P(\mathcal{E}'_{2,1}) - \\
 & + \lambda(v_2 - v_1)(1 - P(\mathcal{E}'_{2,1}) - P(\mathcal{E}'_{1,1}) - P(b^w > v_1) - (1 - P(\mathcal{E}'_{2,2}) - P(\mathcal{E}'_{1,2}) - P(b^w > v_2))) \geq 0 \\
 \implies & \frac{(v_2 - v_1)}{2}(P(\mathcal{E}'_{1,2}) - P(\mathcal{E}'_{1,1}) + P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1) + P(\mathcal{E}'_{2,2}) - P(\mathcal{E}'_{2,1})) \\
 & + \lambda(v_2 - v_1)(+P(\mathcal{E}'_{2,2}) - P(\mathcal{E}'_{2,1}) + P(\mathcal{E}'_{1,2}) - P(\mathcal{E}'_{1,1})) + (v_2 - v_1)(P(b^w > v_1) - P(b^w > v_2))(1/2 - \lambda) \geq 0 \\
 \implies & \frac{(v_2 - v_1)}{2}(P(\bar{\mathcal{E}}_2) - P(\bar{\mathcal{E}}_1) + P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1)) + \lambda(v_2 - v_1)(P(\bar{\mathcal{E}}_2) - P(\bar{\mathcal{E}}_1)) + (v_2 - v_1)(P(b^w > v_1) - P(b^w > v_2)) \geq 0
 \end{aligned} \tag{B.5}$$

Since $\lambda > 1/2$ and $v_2 > v_1$, the last term is negative and a necessary condition for the inequality to hold is

$$\frac{(v_2 - v_1)}{2}(P(\bar{\mathcal{E}}_2) - P(\bar{\mathcal{E}}_1) + P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1)) + \lambda(v_2 - v_1)(P(\bar{\mathcal{E}}_2) - P(\bar{\mathcal{E}}_1)) \geq 0 \tag{B.6}$$

It is easy to check that either $P(\bar{\mathcal{E}}_2) - P(\bar{\mathcal{E}}_1)$ and $P(\mathcal{E}_2^1) - P(\mathcal{E}_1^1)$ are both positive or they are both negative. Since $v_2 > v_1$, we should have $P(\bar{\mathcal{E}}_2) - P(\bar{\mathcal{E}}_1) \geq 0$. This gives $b_2 \geq b_1$. Finally, $b_2 > b_1$ since otherwise, there exists an interval $[v_1, v_2]$ such that $b_1 = b_2 = b(v)$ for any $v \in [v_1, v_2]$ but $\bar{b}_v = b(v) + \epsilon$ is a profitable deviation given that all opponents are bidding $b(v)$.

Proof of Lemma B.2.2

Global IC constraints trivially satisfy local IC constraints. To prove the converse, consider the case when local IC constraint holds, i.e. at $z = v_1$, $\frac{\partial \pi(v_1, z)}{\partial z} = 0$. Consider any $y < v_1$

$$\begin{aligned}
 \pi(v_1, v_1) - \pi(v_1, y) &= \int_y^{v_1} \frac{\partial \pi(v_1, z)}{\partial z} dz \\
 &= \int_y^{v_1} (\pi_z(v_1, z) - \pi_z(z, z)) dz \\
 &= \int_y^{v_1} \int_z^{v_1} (\pi_{z,k}(k, z)) dk dz
 \end{aligned} \tag{B.7}$$

Note that the cross derivative of $\pi(k, z)$, is $(1 + 2\lambda)(1 - z) + b(z)b'(z)$ when $b^{-1}(z) > 1$ and $(b^{-1}(k) - z) + b(z)b'(z)$ otherwise. If $v_1 > z > b(1)$, then $\pi_{k,z}(k, z)$ is positive and if $y < z < b(1)$, then since $k > z > b(z)$, we again have that $\pi_{k,z}(k, z)$ is positive.

Now consider, $b^{-1}(v_1) > y > v_1$

$$\begin{aligned}
 \pi(v_1, v_1) - \pi(v_1, y) &= - \int_{v_1}^y \frac{\partial \pi(v_1, z)}{\partial z} dz \\
 &= \int_{v_1}^y (\pi_z(v_1, z) - \pi_z(z, z)) dz \\
 &= \int_{v_1}^y \int_{v_1}^z (\pi_{z,v}(k, z)) dk dz
 \end{aligned} \tag{B.8}$$

Note that since $k > v_1$, $b^{-1}(v_1) > y > z$, we again have that $(b^{-1}(k) - z) > 0$ in this range as well. Therefore, we have shown for every y , $\pi(v_1, v_1) > \pi(v_1, y)$ and the global IC constraints hold.

Proof of Theorem 3.4.2

We will assume $q = 0$ w.l.o.g. to simplify exposition throughout the section. Let $b(\cdot) : [0, 1] \rightarrow \mathbb{R}$ be the equilibrium bidding function for a two-stage auction. Let us define \mathcal{E}^1 to be the event that bidder is ranked first and no one else bids again. Similarly, let \mathcal{E}^2 be the event that bidder is ranked first but the second ranked bidder bids again. Let \mathcal{E}^3 be the event that bidder is ranked second with but he bids again. Let \mathcal{E}^4 to be the event that bidder is ranked third and he feels regret from non qualification. Any representative bidder in the two stage auction maximizes the expected revenue maximization problem to decide the optimal incentive compatible bidding strategy:

$$\begin{aligned}
 \max_s \pi(v, b(s)) &= (v - b(s))P(\mathcal{E}^1) + \frac{(v - b(s))}{2}P(\mathcal{E}^2) + P(\mathcal{E}^3)E[v - b^w | \mathcal{E}^3] - \lambda P(\mathcal{E}^4)E[v - b^w | \mathcal{E}^4] \\
 &= (v - b(s))P(\max\{V_i\} \leq b(s)) + \frac{(v - b(s))}{2}P(b(s) \leq \max\{V_i\} \leq s) + 2P(V_1 < s) \int_s^1 \frac{v - b(v_2)^+}{2} dv_2 \\
 &\quad - 2\lambda \int_s^1 \int_s^{v_2} (v - b(v_2))^+ dv_1 dv_2 \\
 &= (v - b(s))(b(s))^2 + \frac{(v - b(s))}{2}(s^2 - b(s)^2) + s \int_s^1 (v - b(x))^+ dx - 2\lambda \int_s^1 (v - b(x))^+(x - s) dx
 \end{aligned} \tag{B.9}$$

The second inequality follows since the equilibrium bidding function is strictly increasing. Since the local and global IC constraints are equivalent (by Lemma B.2.2), the corresponding first order condition is $\frac{\partial \pi(v, b(s))}{\partial s} = 0$ when $s = v$. Taking the derivative and setting it to 0 gives us the condition in the theorem.

$$\frac{\partial b(v)}{\partial v} = 2(1 + \lambda) \frac{\int_v^1 (v - b(x))^+ dx}{(v - b(v))^2 + 2b(v)^2}$$

The last part is to show that a solution to this equation always exists. Similar to asymmetric first-price auctions, our ODE is not well defined at the boundary because it is of the $\frac{0}{0}$ form. Nevertheless, existence of a solution can still be shown (Lebrun 1999, Krishna 2009, Plum 1992) by showing that $\frac{\partial b(v)}{\partial v}$ is well defined and continuous for all values of $0 < v \leq 1$ and the limit at 0 exists. Continuity follows trivially since $b(z)$ is continuous and strictly less than z from Lemma B.2.1.

$$\begin{aligned}
 \left. \frac{\partial b(v)}{\partial v} \right|_{v=0} &= \lim_{\epsilon \rightarrow 0} 2(1 + \lambda) \frac{\int_{\epsilon}^1 (\epsilon - b(x))^+ dx}{(\epsilon - b(\epsilon))^2 + 2b(\epsilon)^2} \\
 &= 2(1 + \lambda) \lim_{\epsilon \rightarrow 0} \frac{\int_{\epsilon}^{b^{-1}(\epsilon)} (\epsilon - b(x)) dx}{(\epsilon - b(\epsilon))^2 + 2b(\epsilon)^2} \\
 &= \frac{-2\epsilon + b(\epsilon) + b^{-1}(\epsilon)}{2(\epsilon(1 - b'(\epsilon)) + b(\epsilon)(-1 + 3b'(\epsilon)))} \\
 &= \frac{(-2 + b'(\epsilon))b'(b^{-1}(\epsilon)) + 1}{2b'(b^{-1}(\epsilon))((1 - b'(\epsilon))^2 + 2b'(\epsilon)^2)}
 \end{aligned} \tag{B.10}$$

The second inequality follows from the fact that $\epsilon \leq b(1)$. The third and fourth inequality follow using L'Hopital's Rule. Since b is strictly increasing from Lemma B.2.1, we have that the last term is finite. Further, for all $0 < v \leq 1$, $\frac{\partial b(v)}{\partial v}$ is always well defined.

Proof of Proposition 3.4.2

Let $b^f(\cdot)$ and $b^t(\cdot)$ be the equilibrium bidding functions in the first-price sealed bid and

two-stage auctions respectively. For brevity, we consider the case when $\lambda_L = 0$. Our results continue to hold when for any $\lambda_L \geq 0$. Note that in the first-price sealed bid auction, $b^f(v) = (\int_0^v 2z^2 dz)/(v^2)$. Consider a $\delta < 1$ and $\epsilon > 0$ and let $\gamma = \frac{1}{(1-\delta)^4 - \epsilon^4}$. In order to show the result, we will show that $\forall v \in [\epsilon, 1-\delta]$, there exists a λ s.t. $\gamma b^f(v) \leq b(v)$. For all $0 < v \leq 1-\delta$, note that

$$b^t(v) = \int_0^v \frac{2(1+2\lambda) \int_z^1 (z - b^t(x))^+ dx}{(z - b^t(z))^2 + 2b^{t2}(z)} dz$$

Consider $\lambda^* = \max_{v \in [\epsilon, 1-\delta]} \max_{z \in [0, v]} \frac{2\gamma z^2(z - b(z))^2 + 2b^2(z)}{4v^2(\int_z^1 (z - b(x))^+ dx)}$. Since the denominator is never 0 in the range, note that λ^* is well defined. Further, it is easy to check that this choice of λ^* ensures that $\gamma b^f(v) \leq b^t(v) \forall v \in [\epsilon, 1-\delta]$. Therefore, when $\lambda = \lambda^*$, $E[\pi^f] = \int_0^1 3b^f(z)z^2 dz \leq \gamma \int_\epsilon^{(1-\delta)} 3b^f(z)z^2 dz \leq \int_\epsilon^{(1-\delta)} 3b^t(z)z^2 dz \leq \int_0^1 3b^t(z)z^2 dz = E[\pi^t]$.

The last part is to show that for any v , if $\lambda_2 \geq \lambda_1$, the equilibrium bidding b_2 is greater than b_1 . By optimality, we have

$$(v - b_1)P(\mathcal{E}_1^1) + \frac{(v - b_1)}{2}P(\mathcal{E}_1^2) + P(\mathcal{E}_1^3)E[v - b^w | \mathcal{E}_1^3] - \lambda_1 P(\mathcal{E}_1^4)E[v - b^w | \mathcal{E}_1^4] \geq (v - b_2)P(\mathcal{E}_2^1) + \frac{(v - b_2)}{2}P(\mathcal{E}_2^2) + P(\mathcal{E}_2^3)E[v - b^w | \mathcal{E}_2^3] - \lambda_1 P(\mathcal{E}_2^4)E[v - b^w | \mathcal{E}_2^4] \quad (\text{B.11})$$

$$(v - b_2)P(\mathcal{E}_2^1) + \frac{(v - b_2)}{2}P(\mathcal{E}_2^2) + P(\mathcal{E}_2^3)E[v - b^w | \mathcal{E}_2^3] - \lambda_2 P(\mathcal{E}_2^4)E[v - b^w | \mathcal{E}_2^4] \geq (v - b_1)P(\mathcal{E}_1^1) + \frac{(v - b_1)}{2}P(\mathcal{E}_1^2) + P(\mathcal{E}_1^3)E[v - b^w | \mathcal{E}_1^3] - \lambda_2 P(\mathcal{E}_1^4)E[v - b^w | \mathcal{E}_1^4] \quad (\text{B.12})$$

Adding the two equations, we get,

$$(\lambda_2 - \lambda_1)(P(\mathcal{E}_1^4)E[v - b^w | \mathcal{E}_1^4] - P(\mathcal{E}_2^4)E[v - b^w | \mathcal{E}_2^4]) + (\lambda_2 - \lambda_1)(2 \int_{b_1}^1 \int_{b_1}^x (v - b(x))^+ dy dx - 2 \int_{b_2}^1 \int_{b_2}^x (v - b(x))^+ dy dx) \quad (\text{B.13})$$

Since $\lambda_2 > \lambda_1$, the last inequality will hold only if $b_1 < b_2$.

B.4 Interview Questions

B.4.1 General Questions

- (i) Name:
- (ii) Number of years of experience in the business:
- (iii) Number of commodities you trade in:

B.4.2 Two-Stage Auction

- (i) What are some factors that affect your bid prices?
(If no response, suggest the following:)
 - (a) Total lots in the market
 - (b) Number of traders
 - (c) Quality of the lot
 - (d) Yesterday's closing bid
- (ii) Are there public benchmarks when quoting a price?
(If no response, suggest the following:)
 - (a) Stock Exchange
 - (b) Prices in major market
 - (c) Prices in futures market

B.4.3 Anticipated Regret

Assume that there are 15 traders in the market today and you participate in the two-stage auction. In particular, only top 3 bidders in the first round will qualify for the second-round bidding.

- (i) The maximum that the buyer would be willing to pay you is 5100-5300 (Average: 5200)
What is your first-round bid for lots of different quality?

	Your bid in the first-stage
High-Quality Lot	
Medium Quality	
Low Quality	

(ii) Suppose at the end of the first stage, you find that you did not qualify for the second round. Please rate the intensity of the emotions listed below you would experience in that situation:

	1(not at all)	2	3	4	5	6	7	8	9 (very much)
Anger									
Envy									
Irritation									
Regret									
Sadness									

B.4.4 Competitive Arousal versus Anchoring

(i) Suppose at the end of the first round, you find that you qualify for the second round and the buyer's maximum quote was 5300-5500. Fill the table (Rivalry)

First-Stage Top Bid	Number of Qualified Bidders	Lots in the Market	Your Bid in the Second-Stage
5200	3	100 (Low Supply)	
5200	3	600 (High Supply)	
5200	3	100 (Low Supply)	
5200	3	600 (High Supply)	

(ii) Suppose at the end of the first stage, you find that you qualify for the second round and the buyer's maximum price your buyer is willing to pay is 5400. Fill the table (Time Pressure)

First-Stage Top Bid	Time for the Second-Stage	Your Bid in the Second-Stage
5200	4 minutes	
5200	15 minutes	

B.4.5 Budget Constraints

- (i) Do you have a total budget everyday?
- (ii) Do you exceed your everyday budget target?
- (iii) How concerned are you if you exceed your budget target?
 - (a) Extremely concerned because I have to borrow.
 - (b) Less concerned because I always have enough cash
 - (c) Not concerned because I can sell it at higher price to my buyers
 - (d) Not concerned because I can sell it at higher price to other traders
- (iv) Do you have a total inventory target everyday?
- (v) Do you ever exceed your inventory target?
- (vi) How concerned are you if you exceed your inventory target?
 - (a) Extremely concerned because I have to store.
 - (b) Less concerned because I always have enough storage capacity.
 - (c) Not concerned because I can sell more to my buyers.
 - (d) Not concerned because I can sell to other traders.

Appendix C

Appendix of Chapter 4

C.1 Analytical Details on Model Extensions

C.1.1 Preemptive EMA When Farms are Averse to Quality Uncertainty

To model this extension, we subtract a penalty, $\lambda \mathbb{E}[\mathbb{I}\{n_L > \gamma m\}]$, from the farm's expected payoff function formulated in §4.3.1. This penalty captures the farm's aversion to producing too many low-quality units, which can threaten the livelihood of his family. The next two theorems characterize farms' equilibrium adulteration behavior under perfect and imperfect testing when they are averse to quality uncertainty.

Theorem C.1.1. For preemptive EMA with perfect testing and farms averse to quality uncertainty, the total number of adulterating farms in any NE of the game is characterized by $n_a^{RA*} = \min \{n, \lceil T^{RA} - 1 \rceil\}$, where

$$T^{RA} \equiv n^2 \frac{(r_H - r_L)(p_L^{\max} - p_L^{\min}) + (\lambda/m) (\bar{F}(\gamma, m, p_L^{\max}) - \bar{F}(\gamma, m, p_L^{\min}))}{cqt},$$

and $\bar{F}(\gamma, m, h(x)) \equiv 1 - \int_0^{\gamma m} f(x, m, h(x)) dx$. We have the following three cases.

- (i) If $T^{RA} \leq 1$, then no farm adulterates.
- (ii) If $T^{RA} > n$, then all farms adulterate.
- (iii) If $T^{RA} \in (1, n]$, then any subset of n_a^{RA*} farms adulterating while the rest $(n - n_a^{RA*})$ farms not adulterating constitutes a NE of the game.

Theorem C.1.2. Define $T(x) \equiv -h'(x) \frac{(r_H - r_L) + \lambda f(\gamma m, m, h(x)) \left(\frac{\gamma(1-2h(x))+h(x)}{2h(x)(1-h(x))} \right)}{cq(t/n)((n+1)/n)}$. For

preemptive EMA with imperfect testing and farms averse to quality uncertainty, if $\gamma \leq \frac{p_L^{\min}}{1 - 2p_L^{\min}(1 - p_L^{\min})}$, then there exists a unique symmetric NE of the game in which x_{RA}^{PV*} is determined as follows.

- (a) If $c < T(1)$, then all farms adulterate to the maximum level; i.e., $x_{RA}^{PV*} = 1$.
- (b) If $c \geq T(1)$, then the farms adulterate to some extent; i.e., $x_{RA}^{PV*} \in (0, 1]$ and is the solution to the equation: $x = T(x)$.

We observe that the farms' equilibrium adulteration behavior under either testing scenario follows a very similar pattern as in Theorems 4.3.1 and 4.3.2 in §4.3.1. The condition on γ in Theorem C.1.2 means that a farm should begin to exhibit aversion when the number of his low-quality units is not too large. This condition is reasonable given our focus on smallholder farms. Our next result shows that the risk of preemptive EMA in the supply chain is higher when farms are averse to quality uncertainty than when they are expected-profit maximizers, under both perfect and imperfect testing.

Proposition C.1.3. $n_a^{RA*} \geq n_a^*$ and $x_{RA}^{PV*} \geq x^{PV*}$.

With respect to the effect of supply chain dispersion on preemptive EMA risk in the supply chain, we show as in Proposition ?? that greater dispersion leads to higher risk under imperfect testing.

Proposition C.1.4. If $\gamma \leq \frac{p_L^{\min}}{1 - 2p_L^{\min}(1 - p_L^{\min})}$, then $\frac{\partial x_{RA}^{PV*}}{\partial n} \geq 0$.¹

For the case of perfect testing, we cannot characterize the effect of dispersion on risk analytically. Therefore, we perform extensive numerical simulation and observe that in a total of 32,000,000 numerical instances we run, greater dispersion always leads to a higher risk.²

C.1.2 EMA Risk When the Quality of All Units from the Same Farm is Perfectly Correlated

In this extension we analyze a setting where *all* units of the same farm are of low (high) quality with probability $p_L^{max}(1 - p_L^{max})$. This setup captures scenarios where quality of units of the same farm are highly correlated (in contrast to being independent as in our base model). We

¹We treat n as a continuous variable in this analysis for tractability.

²We use the following parameter values in the numerical simulation: $k = 100000$, 100 m values in $\{100, \dots, 1000\}$, 20 p_L^{max} values in $\{0.1, \dots, 0.9\}$, $p_L^{\min} = 0.1$, $c = 19.07$, 20 q values in $\{0.1, \dots, 0.9\}$, $r_H = 19.07$, $r_L = 0$, $t = n$, $\lambda = \{1, 2, \dots, 40\}$, and 20 γ values in $\{0.1, \dots, 0.9\}$.

again analyze a setting where farms decide whether or not to adulterate low quality units. We first note that our results for preemptive EMA in §4.3.1 will not change in this setup. This is because farms' expected payoff does not change in either perfect or imperfect testing with this updated distribution of low quality units. Next, we characterize farms' equilibrium adulteration behavior under both perfect and imperfect testing for reactive EMA.

Theorem C.1.5. (i) For reactive EMA with perfect testing,

- (a) if $r_H - r_L \leq q(t/n)c$, then none of the farms adulterate low quality units.
- (b) If $r_H - r_L > q(t/n)c$, then all the farms always adulterate low quality units.

(ii) For reactive EMA with imperfect testing,

- (a) if $q(t/n)c \leq \frac{(r_H - r_L)n}{p_L^{max}(n-1) + 1}$, then all the farms always adulterate low quality units.
- (b) If $cq(t/n) \in (\frac{(r_H - r_L)n}{p_L^{max}(n-1) + 1}, (r_H - r_L)n)$, then a mixed strategy Nash equilibrium exists where all farms adulterate low quality units with probability $p_{ad} \in (0, 1)$ and
$$p_{ad} = \frac{(r_H - r_L)n^2}{cqt(n-1)p_L^{max}} - \frac{1}{(n-1)p_L^{max}}.$$
- (c) if $(r_H - r_L)n \leq q(t/n)c$, then farms never adulterate low quality units.

Theorem C.1.5 shows that under perfect testing, farms always (never) adulterate low quality units if the expected per unit penalty is smaller (larger) than the expected per unit revenue gain from adulteration. In contrast, since expected penalty increases with total amount of adulterants under imperfect testing, a symmetric mixed strategy Nash equilibrium that randomizes between adulteration and non adulteration balances the payoffs. Our next proposition shows that similar to Proposition 4.4.2, the risk of reactive EMA as measured by the total expected number of adulterated output in the supply chain is again increasing in supply chain dispersion.

Proposition C.1.6. For reactive EMA with imperfect testing where all units of the same farm are of low quality with probability p_L^{max} , $\frac{\partial E_n}{\partial n} \geq 0$.

C.1.3 Reactive EMA with Decision on How Much to Adulterate

In §4.3.2, we focus on the setup where farms adulterate either all or none of the realized low-quality units. An alternative setup is for farms to decide how many of the realized low-quality units to adulterate. This setup can capture scenarios in which a farm adulterates to fake the overall quality of his output to a desirable level. We first note that our results for perfect

testing will not change under this setup. This is because under perfect testing, any amount of adulteration induces the same level of expected penalty, whereas the revenue gain increases in the number of units being adulterated. Hence, a farm would always adulterate all of his low-quality units if he decides to adulterate. Theorem C.1.7 below characterizes the farms' equilibrium adulteration strategy under imperfect testing, where $a^*(n_{L,i})$ denotes the number of realized low-quality units that farm i adulterates in equilibrium.

Theorem C.1.7. For reactive EMA with imperfect testing where a farm can choose how many of the realized low-quality units to adulterate, there exists a unique symmetric BNE of the game in which a farm's adulteration strategy is a threshold strategy: $a^*(n_{L,i}) = n_{L,i}$ if $n_{L,i} \in [0, \beta_f^{RV}]$ and $a^*(n_{L,i}) = \beta_f^{RV}$ if $n_{L,i} \in (\beta_f^{RV}, m]$, for all i . The threshold β_f^{RV} is unique and determined as follows:

(a) If $cq \left(\frac{t}{n} \right) \geq \frac{n(r_H - r_L)}{2 + (n-1)p_L}$, then $\beta_f^{RV} \in (0, m)$ and is the solution to the equation:

$$2\beta = \frac{nk(r_H - r_L)}{cqt} - (n-1) \left(\int_0^\beta xf(x, \frac{k}{n}, p_L) dx + \int_\beta^m \beta f(x, \frac{k}{n}, p_L) dx \right).$$

(b) If $cq \left(\frac{t}{n} \right) < \frac{n(r_H - r_L)}{2 + (n-1)p_L}$, then $\beta_f^{RV} = m$.

Theorem C.1.7 shows that when farms can choose to adulterate a fraction of his low-quality units, then they adulterate all low-quality units up to a threshold, after which they adulterate a constant number of low-quality units. This structure is very similar to that in Theorem 4.3.4. The only difference is that in the current setup, when a farm has many low-quality units, the farm would adulterate just enough to make the marginal revenue gain from adulteration equal to the marginal penalty, i.e., adulterating β_f^{RV} units (as opposed to not adulterating at all in Theorem 4.3.4). Similar to Proposition 4.4.2, we show that the risk of reactive EMA as measured by the total expected number of adulterated output in the supply chain is increasing in supply chain dispersion.

Proposition C.1.8. For reactive EMA with imperfect testing where a farm can choose how many of the realized low-quality units to adulterate, $\frac{\partial E_n}{\partial n} \geq 0$.

C.1.4 Alternative Models of Imperfect Testing Sensitivity

In this extension, we analyze settings of imperfect testing to model scenarios where detection probability is not linearly increasing in the relative amount of adulterated output in the total

supply chain output (as in §4.3.1 and §4.3.2). In particular, we examine three alternative models where the detection probability is (i) convex increasing in the relative amount of adulterated output, (ii) convex increasing in the relative amount of adulterated output and reaches 1 in the interior of the $(0, 1)$ interval, and (iii) linearly increasing in the relative amount of adulterated output and reaches 1 in the interior of the $(0, 1)$ interval. The last two alternatives capture scenarios in which detection probability increases quickly as a small amount of adulterants are added.

Formally, let the detection probability $S_1 : [0, 1] \rightarrow [0, 1]$ be a convex increasing function such that $S_1(0) = 0$ and $S_1(1) = 1$ under scenario (i). That is, the detection probability should be 0 (1) if none (all) of the output is adulterated. Under imperfect testing, if farm i adulterates with x_i , then the chance that the manufacturer detects adulteration when testing the aggregated supply is equal to $\frac{S_1(x_i) + \sum_{-i} S_1(x_{-i})}{n}$. If the manufacturer further tests the individual sample of farm i , then she detects adulteration in the sample with probability $S_1(x_i)$. Since the manufacturer tests the aggregated supply with probability q , the ultimate probability for farm i to be caught if he adulterates is equal to

$$\gamma_i(x_i, x_{-i}) \equiv q \left(\frac{t}{n} \right) \left(\frac{S_1(x_i) + \sum_{-i} S_1(x_{-i})}{n} \right) S_1(x_i).$$

Similarly, the probability for farm i to be caught if he adulterates under imperfect testing is equal to

$$\gamma_i(n_{L,i}, a_{-i}(n_{L,-i})) \equiv q \left(\frac{t}{n} \right) S_1 \left(\frac{n_{L,i}}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_1 \left(\frac{n_{L,i} + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right) \right].$$

Under model (ii), let the detection probability $S_2 : [0, 1] \rightarrow [0, 1]$ be such that $S_2(0) = 0$, $S_2(a)$ is convex increasing in a for $a \in [0, \tau)$, and $S_2(a) = 1$ for $a \in [\tau, 1]$, for some $\tau \in (0, 1)$. Lastly, let the detection probability $S_3 : [0, 1] \rightarrow [0, 1]$ under model (iii) be a piecewise increasing function such that $S_3(x) = \min(\alpha x, 1)$ for some $\alpha \geq 1$. Note that α captures the detection level of testing. Higher the α , higher is the range in which detection of adulterants happens with perfect accuracy. Under preemptive case, if farm i adulterates with x_i , then the chance that the manufacturer detects the adulteration when testing the aggregated supply is equal to $\min\left(\frac{\alpha(x_i + \sum_{-i} x_{-i})}{n}, 1\right)$. If the manufacturer further tests the individual sample of farm i , then she detects the adulterants in the sample with probability $\min(\alpha x_i, 1)$.

We characterize farms equilibrium adulteration behavior under both preemptive and reactive

EMA for the three models. Under preemptive EMA, the equilibrium adulteration behavior (Theorem C.1.9) again follows the same structure as in Theorem 4.3.2 with updated thresholds. In particular, if penalty is smaller than a threshold, then all farms adulterate upto the maximum level and if it is large enough then farms adulterate to some extent but not to the maximum level.

Theorem C.1.9. (i) For preemptive EMA with imperfect testing and convex increasing testing sensitivity modeled by $S_1(\cdot)$, there exists a unique symmetric NE in which x^{PV^*} is determined as follows.

- (a) If $c < -h'(1)(r_H - r_L)/[S_1'(1)q(t/n)((n+1)/n)]$, then all farms adulterate to the maximum level; i.e., $x^{PV^*} = 1$.
- (b) If $c \geq [-h'(1)(r_H - r_L)/[S_1'(1)q(t/n)((n+1)/n)]]$, then the farms adulterate to some extent; i.e., $x^{PV^*} \in (0, 1)$ and is the solution to the following equation: $-h'(x) = S_1(x)S_1'(x)q(t/n)((n+1)/n)c/(r_H - r_L)$.

(ii) For preemptive EMA with imperfect testing and convex increasing testing sensitivity modeled by $S_2(\cdot)$, there exists a unique symmetric NE in which x^{PV^*} is determined as follows.

- (a) If $c < -h'(\tau)(r_H - r_L)/[S_2(\tau)S_2'(\tau)q(t/n)((n+1)/n)]$, then all farms adulterate to the maximum level; i.e., $x^{PV^*} = 1$.
- (b) If $c \geq -h'(\tau)(r_H - r_L)/[S_2(\tau)S_2'(\tau)q(t/n)((n+1)/n)]$, let $x^* \in (0, \tau)$ be the solution to the equation: $-h'(x) = S_2(x)S_2'(x)q(t/n)((n+1)/n)c/(r_H - r_L)$. There are two cases:
 - i. If $c < h(x^*)/[q(t/n)(1 - S_2(x^*)^2)]$, then all farms adulterate to the maximum level; i.e., $x^{PV^*} = 1$.
 - ii. If $c \geq h(x^*)/[q(t/n)(1 - S_2(x^*)^2)]$, then all farms adulterate to some extent; i.e., $x^{PV^*} = x^*$.

(iii) For preemptive EMA with imperfect testing and testing sensitivity modeled by S_3 , there exists a unique symmetric NE of the game in which x^{PV^*} is determined as follows.

- (a) If $c < -h'(1/\alpha)(r_H - r_L)/[\alpha q(t/n)((n+1)/n)]$, then all farms adulterate to the maximum level; i.e. $x^{PV^*} = 1$.

(b) If $c \leq -h'(1/\alpha)(r_H - r_L)/[\alpha q(t/n)((n+1)/n)]$, then let $x^* \in (0, 1/\alpha)$ be the solution to the following equation: $-h'(x)/x = q(t/n)((n+1)/n)c\alpha^2/(r_H - r_L)$.

i. If $c < (h(x^*) - h(1))(r_H - r_L)/[\alpha^2 q(t/n)(1 - x^{*2})]$ then $x^{PV^*} = 1$

ii. If $c \geq (h(x^*) - h(1))(r_H - r_L)/[\alpha^2 q(t/n)(1 - x^{*2})]$ then $x^{PV^*} = x^*$

Under reactive EMA, we again analyze a setting where a farm adulterates either all or none of his realized low-quality units; i.e., the adulteration strategy can be characterized by the mapping $a_i(n_{L,i}) : \{1, \dots, m\} \rightarrow \{0, 1\}$.

Theorem C.1.10. (i) For reactive EMA with imperfect testing and convex increasing testing sensitivity modeled by $S_1(\cdot)$, there exists a unique symmetric BNE of the game in which a farm's adulteration strategy is a threshold strategy: $a^*(n_{L,i}) = 1$ if $n_{L,i} \in [0, \beta^S)$ and $a^*(n_{L,i}) = 0$ if $n_{L,i} \in [\beta^S, m]$, for all i . The threshold β^S is unique and determined as follows:

(a) If $cq \left(\frac{t}{n} \right) \geq \frac{(r_H - r_L)}{\mathbb{E}_{n_{L,-i}} \left[S_1 \left(\frac{m + \sum_{-i} n_{L,-i}}{k} \right) \right]}$, then $\beta^S \in (0, m)$ and is the solution to the equation:

$$\beta^S(r_H - r_L) = q \left(\frac{t}{n} \right) S_1 \left(\frac{\beta^S}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_1 \left(\frac{\beta^S + \sum_{-i} n_{L,-i} \mathbb{I}\{n_{L,-i} < \beta^S\}}{k} \right) \right] cm,$$

where $\mathbb{I}\{\cdot\}$ is an indicator function whose value is 1 if the argument is true and 0 otherwise.

(b) If $cq \left(\frac{t}{n} \right) < \frac{(r_H - r_L)}{\mathbb{E}_{n_{L,-i}} \left[S_1 \left(\frac{m + \sum_{-i} n_{L,-i}}{k} \right) \right]}$, then $\beta^S = m$.

(ii) For reactive EMA with imperfect testing and convex increasing testing sensitivity modeled by $S_2(\cdot)$, in any BNE of the game, there exist thresholds β^l and β^u such that $\beta^l < \beta^u$ and the following must hold:

(a) If $cq \left(\frac{t}{n} \right) \geq \frac{(r_H - r_L)}{\mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{m + \sum_{-i} n_{L,-i}}{k} \right) \right]}$, then in equilibrium $a^*(n_{L,i}) = 1$ if $n_{L,i} < \beta^l$ or if $n_{L,i} > \beta^u$; $a^*(n_{L,i}) = 0$ if $n_{L,i} \in [\beta^l, \beta^u]$. Furthermore, we must have $\beta^u \geq \tau m$.

(b) If $cq \left(\frac{t}{n} \right) < \frac{(r_H - r_L)}{\mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{m + \sum_{-i} n_{L,-i}}{k} \right) \right]}$, then all farms always adulterate.

(iii) For reactive EMA with imperfect testing and testing sensitivity modeled by S_3 , there exists a symmetric BNE of the game in which a farm's adulteration strategy is a threshold strategy: $a^*(n_{L,i}) = 1$ if $n_{L,i} \in [0, \beta^S]$ or if $n_{L,i} \in [\beta^U, m]$.

$$\begin{aligned}
 & \text{(a) If } cq(t/n) \leq \max\left(\frac{r_H - r_L}{\alpha}, \frac{k(r_H - r_L)}{\alpha(p\alpha(k - m) + m)}\right) \text{ then } \beta^S = \beta^U = m \\
 & \text{(b) If } cq(t/n) > \max\left(\frac{r_H - r_L}{\alpha}, \frac{k(r_H - r_L)}{\alpha(p\alpha(k - m) + m)}\right) \text{ then} \\
 & \quad \beta^S = \max\left(0, \frac{k(r_H - r_L)}{cq\alpha^2} - \left(\frac{k}{m} - 1\right)\left(\int_0^{\beta^S} xf(x, m, p)dx + \int_{\beta^U}^m xf(x, m, p)dx\right)\right) \text{ and} \\
 & \quad \beta^U = \min\left(m, \frac{cqm(t/n)}{r_H - r_L}, \frac{\left(\frac{k}{m} - 1\right)\left(\int_0^{\beta^S} xf(x, m, p)dx + \int_{\beta^U}^m xf(x, m, p)dx\right)}{\frac{k(r_H - r_L)}{cqm\alpha} - 1}\right)
 \end{aligned}$$

Theorems C.1.9 and C.1.10 below show that the farms' equilibrium adulteration behavior under scenario (i) follows a very similar structure as in Theorems 4.3.4 in and §4.3.2. Under scenario (ii) and (iii), the structure of the equilibrium strategy in Theorem C.1.10 can be viewed as a combination of the equilibrium strategies described in Theorems 4.3.3 and 4.3.4 in §4.3.2. In particular, the equilibrium strategy here combines the adulteration strategies identified in perfect and imperfect testing. If the number of low quality units are very low, then the testing sensitivity is similar to the imperfect testing case and farms adulterate when their low quality units are less than a threshold. In contrast, if low quality units are greater than a threshold, then testing sensitivity is similar to perfect testing and farms adulterate when their low quality units are greater than a threshold.

C.1.5 Alternative Penalty Structures

In this extension, we first analyze cases in which per unit penalty is linearly increasing in the amount of adulterants. This section is valuable in capturing scenarios where harmful effects of adulterants are increasing in the amount of adulterants. We model such case by assuming that penalty is linearly increasing in the amount of adulterants. Thus for a farm that adds x_i adulterants per unit under preemptive testing and is caught, the penalty is $cm(mx)$. We again find that the equilibrium structure under all the scenarios is similar to the ones identified in Section §4.3.2 and §4.3.1 except in one case: For penalty alternative (i) and preemptive EMA with perfect testing, adulterating farms use an amount of adulterants that balances the revenue gain with the penalty from adulterating. The additional tradeoff arise because the penalty on being caught adulterating is no longer constant but increasing in the amount of adulterants.

Theorem C.1.11. (i) For preemptive EMA with perfect testing, there exists a unique symmetric NE in which x^{PP^*} is determined as follows.

(a) If $c \geq -h'(0)(r_H - r_L)n/[mq(t/n)(n + 1)]$, then

$$n_a^* = \left\lceil \frac{-h'(0)(r_H - r_L)n^2}{mcqt} - 1 \right\rceil \quad (\text{C.1})$$

(b) If $c < -h'(0)(r_H - r_L)n/[mq(t/n)(n + 1)]$, then $n_a^* = n$

Let x^* be the solution to the following equation: $-h'(x)(r_H - r_L) = mcq(t/n)(n_a^*/n)$. Then $x_a^{PP^*} = \min(x^*, 1)$ and $x_a^{PP^*} \in (0, 1)$.

(ii) For preemptive EMA with imperfect testing, there exists a unique symmetric NE in which x^{PV^*} is determined as follows.

(a) If $c < -h'(1)(r_H - r_L)/[mq(t/n)((2n + 1)/n)]$, then all farms adulterate to the maximum level; i.e., $x^{PV^*} = 1$.

(b) If $c \geq [-h'(1)(r_H - r_L)/[mq(t/n)((2n + 1)/n)]$, then the farms adulterate to some extent; i.e., $x^{PV^*} \in (0, 1)$ and is the solution to the following equation: $-h'(x)/x^2 = q(t/n)((2n + 1)/n)c/(r_H - r_L)$.

(iii) For reactive EMA with perfect testing,

(a) if $r_H - r_L \leq q(t/n)cm$, then none of the farms adulterate; i.e. $\beta^{RP^*} = 0$.

(b) If $r_H - r_L > q(t/n)cm$, then all the farms adulterate to the maximum level; i.e., $\beta^{RP^*} = m$

(iv) For reactive EMA with imperfect testing, there exists a unique symmetric BNE of the game in which a farm's adulteration strategy is a threshold strategy: $a^*(n_{L,i}) = 1$ if $n_{L,i} \in [0, \beta^{RV})$ and $a^*(n_{L,i}) = 0$ if $n_{L,i} \in [\beta^{RV}, m]$, for all i . The threshold β^{RV} is unique and determined as follows:

(a) If $cq \left(\frac{t}{n} \right) \geq \frac{k(r_H - r_L)}{m(m + (n - 1)mp)}$ then $\beta^{RV} \in (0, m)$ and is the solution to the equation: $\beta^2 + \beta(n - 1) \int_0^\beta xf(x, \frac{k}{n}, p_L)dx = \frac{nk(r_H - r_L)}{cqt}$.

(b) If $cq \left(\frac{t}{n} \right) < \frac{k(r_H - r_L)}{m(m + (n - 1)mp)}$, then $\beta^{RV} = m$.

Next, we find qualitative evidence that larger companies who are caught adulterating face more severe penalty than smaller ones. For example, they face longer jail terms and are fined more heavily (Yan 2017). This observation can be captured by modeling the total penalty from

adulterating as convex increasing in m . We find that all of our results continue to hold in this alternative setup with updated threshold values.

Proposition C.1.12. All of our results in §4.3.1 and §4.3.2 continue to hold if the penalty that a farm incurred for adulterating is convex increasing in the total number of units, m , supplied by the farm.

C.2 Investing in Traceability and Testing Frequency to Mitigate EMA Risk

Given a supply network of farms, the manufacturer has two levers to mitigate the risk of EMA in the supply chain: increasing supply chain traceability and the frequency of testing the aggregated supply. Developing these capabilities can be costly. For example, it is very difficult to trace and inspect every individual farm in a supply chain sourcing from thousands of farms (Nestlé 2015). Therefore, the manufacturer needs to balance between the cost of investing in these capabilities and the benefit of reducing EMA risk in the supply chain. To address this tradeoff, we develop an optimization model from the manufacturer's perspective, where the objective is to minimize total investment costs while satisfying a constraint that the resulting risk of EMA in the supply chain cannot exceed a certain level.

First consider preemptive EMA. In this setting, the overall risk of EMA in the supply chain is measured by n_a^*/n under perfect testing and x^{PV^*} under imperfect testing (see §4.4). Define $l(q)$ and $g(t)$ as the manufacturer's investment costs for increasing testing frequency and traceability, both of which are convex and increasing functions. The manufacturer's optimization problem under preemptive EMA can be characterized as follows.

$$\Pi^{PP}(q, t) \equiv \min_{q,t} \{l(q) + g(t) \mid n_a^*/n \leq \alpha, q \in [0, 1], t \in [0, n]\}, \quad (\text{C.2})$$

$$\Pi^{PV}(q, t) \equiv \min_{q,t} \left\{ l(q) + g(t) \mid x^{PV^*} \leq \alpha, q \in [0, 1], t \in [0, n] \right\}, \quad (\text{C.3})$$

where n_a^* and x^{PV^*} are defined in Theorems 4.3.1 and 4.3.2, and α is the maximum level of risk allowed. For reactive EMA, the manufacturer's optimization problem can be modeled similarly as follows.

$$\Pi^{RI}(q, t) \equiv \min_{q,t} \{l(q) + g(t) \mid P_n \leq \alpha, q \in [0, 1], t \in [0, n]\}, \quad (\text{C.4})$$

$$\Pi^{RE}(q, t) \equiv \min_{q, t} \{l(q) + g(t) | E_n \leq \alpha, q \in [0, 1], t \in [0, n]\}, \quad (\text{C.5})$$

where P_n and E_n are defined in §4.4 given the farms' optimal adulteration strategies under perfect and imperfect testing, characterized in Theorems 4.3.3 and 4.3.4. The key difference is that we measure the risk of reactive EMA in the supply chain in two ways: the probability of an individual farm adulterating (i.e., P_n as in Model (C.4)) and the expected total amount of adulterated output in the supply chain (i.e., E_n as in Model (C.5)).

Before characterizing the manufacturer's optimal investment strategy under each of these model scenarios, we first define the following useful constants.

(i) For Model (C.2):

$$u^{PP} \equiv \frac{(r_H - r_L)(p_L^{\max} - p_L^{\min})n^2}{c(1 + \lfloor n\alpha \rfloor)}. \quad (\text{C.6})$$

(ii) For Model (C.3):

$$u^{PV} \equiv \frac{-h'(\alpha)(r_H - r_L)n}{\alpha c(n + 1)/n}. \quad (\text{C.7})$$

(iii) For Model (C.4) under perfect testing:

$$u^{RP} \equiv \left(\frac{n^2(r_H - r_L)}{ck} \right) \left(\frac{p_L k}{n} + \phi^{-1}(1 - \alpha) \sqrt{\frac{kp_L(1 - p_L)}{n}} \right). \quad (\text{C.8})$$

(iv) For Model (C.4) under imperfect testing:

$$u^{RV} \equiv \left(\frac{kn(r_H - r_L)}{c} \right) \left(\frac{p_L k}{n} + \phi^{-1}(\alpha) \sqrt{\frac{kp_L(1 - p_L)}{n}} + (n - 1) \int_0^{\frac{p_L k}{n} + \phi^{-1}(\alpha) \sqrt{\frac{kp_L(1 - p_L)}{n}}} x f(x, k/n, p_L) dx \right)^{-1}. \quad (\text{C.9})$$

The notation ϕ represents the PDF of the standard normal distribution. These constants are the values of qt when the risk constraint in the corresponding models indicated is binding. Note that in the optimal solution to these models, the risk constraint must be binding because n_a^* , x^{PV*} , and P_n are all decreasing in q and t , whereas the investment costs are increasing in q and t . The following theorem summarizes the manufacturer's optimal investment strategy for Models (C.2), (C.3), and (C.4) under perfect and imperfect testing.

Theorem C.2.1. Given the constants u^j with $j \in \{PP, PV, RP, RV\}$ defined in Equations (C.6)–(C.9), we have the following results for $\alpha \leq 0.5$.

- (i) If $u^j > n$, then the corresponding manufacturer problem is infeasible.
- (ii) If $u^j \leq n$, then the optimal solution to the corresponding manufacturer problem (q^*, t^*) can be characterized as follows.
 - (a) If $l'(1) \leq u^j g'(u^j)$, then $(q^*, t^*) = (1, u^j)$.
 - (b) If $g'(n) \leq \frac{u^j}{n^2} l' \left(\frac{u^j}{n} \right)$, then $(q^*, t^*) = \left(\frac{u^j}{n}, n \right)$.
 - (c) If $l'(1) > u^j g'(u^j)$ and $g'(n) > \frac{u^j}{n^2} l' \left(\frac{u^j}{n} \right)$, then $(q^*, t^*) \in (0, 1) \times (0, n)$ and satisfy the following first-order conditions: $q^* = \sqrt{\frac{u^j g'(u^j/q^*)}{l'(q^*)}}$ and $t^* = \frac{u^j}{q^*}$.

Theorem C.2.1 part (i) suggests that if the manufacturer cannot satisfy the risk constraint even when the supply chain is fully traceable and she always tests the aggregated supply, then additional levers are necessary to meet the risk constraint. Given our earlier discussions in §4.4, one possible solution is to reduce supply chain dispersion. Theorem C.2.1 part (ii) shows that when a feasible solution exists, the manufacturer always chooses the solution with the best cost-effectiveness. If the marginal cost at maximum testing frequency is lower than the marginal cost at the minimum necessary traceability to satisfy the risk constraint (i.e., increasing testing frequency is in general more cost effective than increasing traceability; Theorem C.2.1 part (ii-a)), then it is optimal for the manufacturer to always test the aggregated supply and build just enough traceability given the risk constraint. Conversely, if increasing traceability is in general more cost effective than increasing testing frequency (Theorem C.2.1 part (ii-b)), then it is optimal for the manufacturer to build full traceability in the supply chain and test just enough given the risk constraint. If neither of the above is true (Theorem C.2.1 part (ii-c)), then the optimal investment is an interior solution that achieves the best cost balance between investing in the two levers.

Our next proposition characterizes how the optimal investment solution (q^*, t^*) described in Theorem C.2.1 and the resulting optimal cost change with supply chain dispersion.

Proposition C.2.2. Let SC_H and SC_L be two supply chains such that the supply chain dispersion in SC_H is greater than that in SC_L (i.e., $n_H > n_L$). Consider each of the manufacturer's

optimization problems formulated in Models (C.2), (C.3), and (C.4) under perfect and imperfect testing. We have the following results for $\alpha \leq 0.5$.

- (i) If the manufacturer's problem is infeasible for SC_L , then it is also infeasible for SC_H .
- (ii) If the manufacturer's problem is feasible for SC_H , then it is also feasible for SC_L .
- (iii) Assume that the manufacturer's problem is feasible for SC_H . Let (q_H^*, t_H^*) and (q_L^*, t_L^*) be the optimal solution for SC_H and SC_L respectively. Then, $q_L^* \leq q_H^*$, $t_L^* \leq t_H^*$, and the resulting optimal cost for the manufacturer is lower in SC_L than in SC_H .

Proposition C.2.2 highlights two results. First, given a desirable risk constraint, it is always more difficult for a manufacturer with a more dispersed supply chain to satisfy the constraint (parts (i) and (ii)). Second, conditional on being able to satisfy the risk constraint, it is always more costly for a manufacturer with a more dispersed supply chain to do so (part (iii)). Therefore, higher supply chain dispersion results in greater challenges for a manufacturer to manage and mitigate the risk of individual farms adulterating, from both feasibility and financial standpoints.

Finally, we consider the manufacturer's problem formulated in Model (C.5). The key difference in this model versus the others is that the risk constraint is imposed on E_n , the expected total amount of adulterated output in the supply chain. Since E_n aggregates all farms' adulteration decisions, we cannot derive the manufacturer's optimal decisions analytically. Nevertheless, consistent with Proposition C.2.2, we show that higher supply chain dispersion again makes it more costly for the manufacturer to satisfy a desirable risk constraint, regardless of testing sensitivity (perfect or imperfect testing).

Proposition C.2.3. Let SC_H and SC_L be two supply chains such that the supply chain dispersion in SC_H is greater than that in SC_L (i.e., $n_H > n_L$). Consider the manufacturer's optimization problem formulated in Model (C.5) and assume that it is feasible for SC_H . If $\alpha \leq n \int_{mp/3} x f(x, m, p) dx$, then for both perfect and imperfect testing, the optimal cost for the manufacturer is lower in SC_L than in SC_H .

C.3 Investing in Testing Capabilities to Mitigate EMA Risk

We examine the effect of the manufacturer investing in perfect testing on mitigating EMA risk in the supply chain. To this end, we allow farms to engage in both preemptive and reactive

EMA in response to the manufacturer's testing capability. The model dynamics are similar to those described in §4.2 with the first two steps revised as follows. (i) The manufacturer chooses whether or not to adopt perfect testing for preemptive or reactive EMA respectively. The farms observe the manufacturer's choice. (ii) Each farm simultaneously and individually decides the amount of adulterants to add to reduce the likelihood of producing low-quality output from p_L^{\max} to some $p_L \leq p_L^{\max}$ (preemptive EMA). (iii) The uncertain quality of each unit of output is realized. (iv) Each farm simultaneously and individually decides whether or not to adulterate all of the realized low quality units n_L to create fake high-quality ones (reactive EMA). The remaining steps are exactly the same as in Figure 4.1. We are interested in analyzing how the total EMA risk, accounting for both preemptive and reactive EMA, is affected by the manufacturer's testing capability for either type of EMA. In particular, we analyze whether adopting perfect testing always reduces EMA risk in the supply chain. In this analysis, we take the farms to be short-term oriented (see footnote 8), and thus, they do not account for every possible realization of n_L and the corresponding reactive EMA decision when making their preemptive EMA decision. To simplify exposition, we also assume that when the farms adulterate preemptively with the maximum dosage, p_L becomes 0. Our results remain qualitatively the same without this simplifying assumption. Table C.1 summarizes the total EMA risk in the supply chain for the four different scenarios we analyze in §4.3. For example, the top left cell shows the total EMA risk if the manufacturer invests in perfect testing for both preemptive and reactive EMA. By Theorem 4.3.1 we know that under perfect testing, a subset of n_a^* farms adulterate preemptively with the maximum dosage while the remaining do not adulterate at all. Thus, the expected total amount of adulterants added preemptively is mn_a^* . In the reactive EMA stage, we again know from Theorem 4.3.3 that under perfect testing, farms adulterate when their low-quality units are greater than β^{RP} . Hence, the total EMA risk in the supply chain is equal to $\lambda mn_a^* + (n - n_a^*) \int_{\beta^{RP}}^m xf(x, m, p_L^{\max}) dx$. Note that λ here measures the importance of preemptive EMA relative to reactive EMA when the manufacturer evaluates the total EMA risk. We can similarly characterize the total EMA risk in the supply chain for the other three scenarios.

By Proposition 4.3.5, we know that reactive EMA risk is always lower when the manufacturer adopts perfect testing. Thus, we only need to compare the total EMA risk in the left two cells in Table C.1. We first focus on comparing the preemptive EMA risk between these two scenarios. Our results are summarized in the next proposition.

	Reactive EMA, perfect testing	Reactive EMA, imperfect testing
Preemptive EMA, perfect testing	$\lambda mn_a^* + (n - n_a^*) \int_{\beta_{RP}}^m x f(x, m, p_L^{\max}) dx$	$\lambda mn_a^* + (n - n_a^*) \int_0^{\beta^{RV}} x f(x, m, p_L^{\max}) dx$
Preemptive EMA, imperfect testing	$\lambda k x^{PV^*} + n \int_{\beta_{RP}}^m x f(x, m, h(x^{PV^*})) dx$	$\lambda k x^{PV^*} + n \int_0^{\beta^{RV}} x f(x, m, h(x^{PV^*})) dx$

Table C.1: Total EMA Risk When Farms Engage in Both Preemptive and Reactive EMA

Proposition C.3.1. Let $R_p^P \equiv mn_a^*$ and $R_{ip}^P \equiv kx^{PV^*}$ denote the preemptive EMA risk under perfect and imperfect testing respectively. Then,

- (i) If $c < -h'(1)(r_H - r_L)/[q(t/n)(n + 1)/n]$, then $R_p^P = k$ and $R_{ip}^P = k$. Thus, all farms adulterate to the maximum level under both perfect and imperfect testing cases.
- (ii) If $c \in [-h'(1)(r_H - r_L)/[q(t/n)(n + 1)/n, (r_H - r_L)(p_L^{\max} - p_L^{\min})/[q(t/n)]]$, then $x^{PV^*} \in (0, 1)$ and $n_a^* = n$. Thus, $R_p^P \geq R_{ip}^P$, i.e., preemptive EMA risk is higher under perfect testing than under imperfect testing.
- (iii) If $c \geq (r_H - r_L)(p_L^{\max} - p_L^{\min})/[q(t/n)]$, and
 - (a) If $c < (r_H - r_L)(p_L^{\max} - p_L^{\min})/[q(t/n)h^{-1}((p_L^{\min} - p_L^{\max})(1 + 1/n))]$, then $R_p^P \geq R_{ip}^P$, i.e., preemptive EMA risk is higher under perfect testing than under imperfect testing.
 - (b) If $c \geq (r_H - r_L)(p_L^{\max} - p_L^{\min})/[q(t/n)h^{-1}((p_L^{\min} - p_L^{\max})(1 + 1/n))]$ then $R_{ip}^P \geq R_p^P$, i.e., preemptive EMA risk is higher under imperfect testing than under perfect testing.

Proposition C.3.1 shows that when the per-unit penalty is not high enough, adopting perfect testing can in fact backfire and result in *higher* preemptive EMA risk. This is because under perfect testing, adulterating farms all adulterate with the maximum dosage, while under imperfect testing, they adulterate at a lower level to trade off revenue gain with the expected penalty. When the penalty is not high enough, more farms adulterate to the maximum dosage under perfect testing, therefore leading to a higher risk. Our next result shows that this observation remains true when considering the total EMA risk that also accounts for reactive EMA.

Theorem C.3.2. Let R_p^T denote the total EMA risk under perfect testing for both preemptive and reactive EMA, and R_{ip}^T the total EMA risk under imperfect testing for preemptive EMA and perfect testing for reactive EMA. We have the following results.

- (i) If $c < -h'(1)(r_H - r_L)/[q(t/n)(n + 1)/n]$, then $R_p^T = R_{ip}^T$.
- (ii) If $c \geq (r_H - r_L)/[q(t/n)]$, and
- (a) If $c < (r_H - r_L)(p_L^{max})/[q(t/n)h^{-1}(-p_L^{max}(1 + 1/n))]$, then $R_p^T \geq R_{ip}^T$.
- (b) If $c \geq (r_H - r_L)(p_L^{max})/[q(t/n)h^{-1}(-p_L^{max}(1 + 1/n))]$, then $R_p^T \geq R_{ip}^T$.

We complement Theorem C.3.2 with extensive numerical simulation for the range of c values that we cannot characterize the total risk analytically. Figure C.1 presents a representative pattern of how the total EMA risk changes with c under either perfect or imperfect testing for preemptive EMA and perfect testing for reactive EMA. Observe that adopting perfect testing for preemptive EMA in fact leads to higher total EMA risk inadvertently when c is not sufficiently high (for $c < r_H$ in this example).

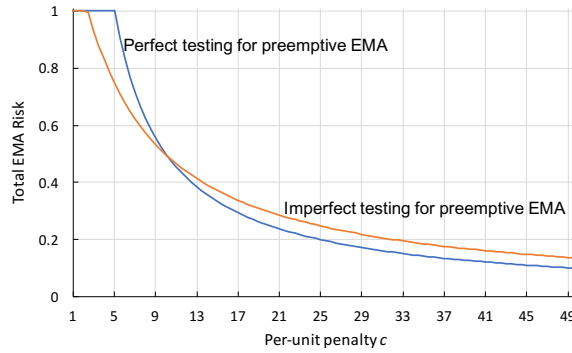


Figure C.1: Total EMA Risk under Perfect or Imperfect Testing for Preemptive EMA and Perfect Testing for Reactive EMA

Note: We use the following parameters in this example: $k = 100,000$, $m = 1,000$, $q = 1$, $r_H = 10$, $r_L = 0$, $t = n$, $c \in \{1, 1.5, \dots, 50\}$, $p_L^{max} = 0.5$, and $p_L^{min} = 0$.

C.4 Proofs of Analytical Results

Proof of Theorem 4.3.1

Let us assume that a subset n_a out of a total of n farms are adulterating. Consider a farm i that is adulterating at equilibrium. Then, revenue from adulteration (π_a^i) should be strictly greater than revenue from non adulteration (π_{na}^i). $\pi_a^i = r_H m(1 - p_L^{min}) + r_L m p_L^{min} - q \frac{n_a}{n} (t/n) cm$ and $\pi_{na}^i = r_H m(1 - p_L^{max}) + r_L m p_L^{max}$. Now consider a farmer i' who is not adulterating at equilibrium. Then, $\pi_{na}^{i'} \geq \pi_a^{i'}$. $\pi_a^{i'} = r_H m(1 - p_L^{min}) + r_L m p_L^{min} - q \frac{n_a + 1}{n} (t/n) cm$ and $\pi_{na}^{i'} = r_H m(1 - p_L^{max}) + r_L m p_L^{max}$. Simplifying the two conditions we get, $\frac{n^2(r_H - r_L)(p_L^{max} - p_L^{min})}{cqt} > n_a \geq$

$\frac{n^2(r_H - r_L)(p_L^{max} - p_L^{min})}{cqt} - 1$. Thus we get the desired result using the fact that n_a has to lie between 0 and n .

Proof of Theorem 4.3.2

We first define a Nash equilibrium in the game.

Definition C.4.1. The strategy profile $\{x_i^{PV^*}, i = 1, \dots, n\}$ constitutes a Nash equilibrium of the game if $x_i^{PV^*} \in \arg \max_{x_i \in [0,1]} \pi^{PV}(x_i, x_{-i}^{PV^*})$, where $\pi^{PV}(\cdot, \cdot)$ is defined in Equation (4.2).

At an NE, farm i chooses x_i^* so as to maximize his expected payoff in Equation (4.2) given the choices of all other farms. Note that $\frac{\partial \pi^{PV}(x_i, x_{-i})}{\partial x_i} = -mh'(x)(r_H - r_L) - q\frac{t}{n}cm(\frac{\sum_{-i} x_{-i} + 2x_i}{n})$ and $\frac{\partial^2 \pi^{PV}(x_i, x_{-i})}{\partial x_i^2} = -mh''(x)(r_H - r_L) - q\frac{t}{n}cm(\frac{2}{n}) < 0$. Further, since we consider only symmetric equilibrium, we have $x_{-i}^* = x_i^* = x^{PV^*}$. Substituting this, we get $\frac{\partial \pi^{PV}(x_i, x_{-i})}{\partial x_i} = -h'(x^*)(r_H - r_L) - q\frac{t}{n}c\frac{n+1}{n}x^*$. Since $\pi^{PV}(x_i, x_{-i})$ is a concave function, optimal solution is achieved either at the first order condition or at the boundary:

- (a) If $c < -h'(1)(r_H - r_L)/[q(t/n)((n+1)/n)]$, $x^{PV^*} = 1$.
- (b) If $c \geq -h'(1)(r_H - r_L)/[q(t/n)((n+1)/n)]$, then $x^{PV^*} \in (0, 1)$ and is the solution to the FOC: $-h'(x)/x = q(t/n)((n+1)/n)c/(r_H - r_L)$.

Proof of Theorem 4.3.3

Farmer's revenue is $\pi^{RP}(n_L) = \max \{r_H(m - n_L) + r_L n_L, r_H m - q(t/n)cm\}$ under perfect testing for reactive cases. Thus, he adulterates if $r_H m - q(t/n)cm > r_H(m - n_L) + r_L n_L$. This simplifies to $n_L > \frac{q(t/n)cm}{r_H - r_L} \equiv \beta^{RP}$. Observe that $\frac{\partial \beta^{RP}}{\partial (r_H - r_L)} = -\frac{q(t/n)cm}{(r_H - r_L)^2} \leq 0$, $\frac{\partial \beta^{RP}}{\partial q} = \frac{(t/n)cm}{(r_H - r_L)} \geq 0$, $\frac{\partial \beta^{RP}}{\partial t} = \frac{q(1/n)cm}{(r_H - r_L)} \geq 0$ and $\frac{\partial \beta^{RP}}{\partial c} = \frac{q(t/n)m}{(r_H - r_L)} \geq 0$.

Proof of Theorem 4.3.4

We first formally define a Bayesian Nash equilibrium in the game as follows.

Definition C.4.2. The strategy profile $\{a_i^*(n_{L,i}), i = 1, \dots, n\}$ constitutes a Bayesian Nash equilibrium (BNE) of the reactive adulteration game if,

$a_i^*(n_{L,i}) \in \arg \max_{a \in \{0,1\}} \mathbb{E}_{n_{L,-i}} [\pi_i^{RV}(a, n_{L,i}, a_{-i}^*(n_{L,-i}))]$, where

$$\mathbb{E}_{n_{L,-i}} [\pi_i^{RV}(a, n_{L,i}, a_{-i}^*(n_{L,-i}))] = \begin{cases} r_H(m - n_{L,i}) + r_L n_{L,i}, & \text{if } a = 0, \\ r_H m - \gamma_i(n_{L,i}, a_{-i}^*(n_{L,-i}))cm, & \text{if } a = 1, \end{cases} \quad (\text{C.10})$$

and $\gamma_i(\cdot, \cdot)$ is defined in Equation (4.3).

First we show that any optimal strategy in this game is a threshold policy. Farmer i 's revenue is $\max\{r_H(m - n_L) + r_L n_L, r_H m - \gamma_i(n_{L,i}, a_{-i}(n_{L,-i}))cm\}$, where $\gamma_i(n_{L,i}, a_{-i}(n_{L,-i})) \equiv q\left(\frac{t}{n}\right)\left(\frac{n_{L,i}}{m}\right)\left(\frac{n_{L,i} + \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]}{k}\right)$ from equation 4.3. Thus he adulterates if $r_H m - \gamma_i(n_{L,i}, a_{-i}(n_{L,-i}))cm > r_H(m - n_{L,i}) + r_L n_{L,i}$. This simplifies to

$$\frac{kn(r_H - r_L)}{cqt} - \mathbb{E}_{n_{L,-i}} \left[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i}) \right] \geq n_{L,i}$$

. Let $\beta \equiv \frac{kn(r_H - r_L)}{cqt} - \mathbb{E}(\sum_{j=1 \setminus i}^n a_{n_{L,j}} b_j)$. Under any BNE, farm i will adulterate only when his low quality units are less than the threshold β . Since all the farms behave symmetrically and they are all identical, all farms adulterate only when their realized low quality units are less than β . Thus, $a^*(n_{L,i}) = 1$ if $n_{L,i} \in [0, \beta)$ and $a^*(n_{L,i}) = 0$ if $n_{L,i} \in (\beta, m]$, for all i . Next, we characterize β and show that it is unique.

Using the fact that all farms have the same threshold β , we have that $\mathbb{E}_{n_{L,-i}} [\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})] = (n - 1) \int_0^{\beta^{RV}} x f(x, m, p_L) dx$. Note that $\max_{\beta} \left\{ \int_0^{\beta} x f(x, m, p_L) dx \right\} = mp_L$ and it is attained when $\beta \geq m$. If $\beta = m$ is a symmetric BNE, then given that $\beta_{-i} = m$, we should get that $\beta_i = m$ as well. This happens when $\beta_i \equiv \frac{kn(r_H - r_L)}{cqt} - (n - 1)mp_L \geq m$, because it implies that, given that all other farms have $\beta_{-i} = m$, $\beta_i = m$ is optimal for farm i . This condition simplifies to $\frac{kn^2(r_H - r_L)}{cqt} - (n - 1)kp_L - k \geq 0$, which is a quadratic in n and is satisfied when $n > 2(1 - p_L) \left/ \left(\sqrt{p_L^2 + \frac{4(1 - p_L)(r_H - r_L)}{cqt}} - p_L \right) \right.$. This proves part (b) of the proposition. If $n \leq 2(1 - p_L) \left/ \left(\sqrt{p_L^2 + \frac{4(1 - p_L)(r_H - r_L)}{cqt}} - p_L \right) \right.$, we will use the Intermediate Value Theorem to show the existence of β^{RV} . Define $F(\beta) \equiv \beta - \frac{kn(r_H - r_L)}{cqt} + (n - 1) \int_0^{\beta} x f(x, m, p_L) dx$ for $\beta \in [0, m]$. Note that $F(0) = -\frac{kn(r_H - r_L)}{cqt} < 0$ and $F(m) = m - \frac{kn(r_H - r_L)}{cqt} + (n - 1)mp_L > 0$. Since $F(\cdot)$ is continuous on $[0, m]$ and $F(0) < 0 < F(m)$, there exists $\beta^{RV} \in (0, m)$ s.t. $\beta^{RV} = \frac{kn(r_H - r_L)}{cqt} - (n - 1) \int_0^{\beta^{RV}} x f(x, m, p_L) dx$. Finally, since the RHS in the equation is monotonically decreasing in β^{RV} , the fixed point is also unique.

Using the Implicit function theorem, we have $\frac{\partial \beta^{RV}}{\partial (r_H - r_L)} = \frac{kn}{cqt(1 + (n - 1)\beta f(\beta, m, p_L))} > 0$ and $\frac{\partial \beta^{RV}}{\partial c} = -\frac{kn(r_H - r_L)}{c^2qt(1 + (n - 1)\beta f(\beta, m, p_L))} < 0$ and $\frac{\partial \beta^{RV}}{\partial q} = -\frac{kn(r_H - r_L)}{cq^2t(1 + (n - 1)\beta f(\beta, m, p_L))} < 0$ and $\frac{\partial \beta^{RV}}{\partial t} = -\frac{kn(r_H - r_L)}{cqt^2(1 + (n - 1)\beta f(\beta, m, p_L))} < 0$.

Proof of Proposition 4.3.5

Note that $\beta^{RP} = \frac{cqmt}{n(r_H - r_L)}$ from Theorem 4.3.3. First, we will show that if $\beta^{RP} \in (0, m)$

than $\beta^{RV} = m$. Note that $\beta^{RP} < m$ is equivalent to $\frac{cqt}{n(r_H - r_L)} < 1$. Further, we already know from theorem 4.3.4 that $\beta^{RV} = m$ if $\frac{kn(r_H - r_L)}{cqt} - (n-1)mp_L \geq m$. This condition simplifies to $\frac{n}{(n-1)p+1} > \frac{cqt}{n(r_H - r_L)}$. Since $p < 1$, $\frac{n}{(n-1)p+1} > 1 > \frac{cqt}{n(r_H - r_L)}$. Similarly if $\beta^{RV} \in (0, m)$, then $m < \frac{nm}{(n-1)p+1} < \frac{cqt m}{n(r_H - r_L)}$ and this implies $\beta^{RP} = m$.

Proof of Corollary 1

We have shown above that $\beta^{RV} = m$ if $\beta^{RP} \in (0, m)$. Since EMA risk is maximum when $\beta^{RV} = m$, we have that EMA risk under imperfect testing is always greater than that under perfect testing. Next, note that if $\beta^{RP} = m$, then EMA risk under perfect testing is 0 and is again always smaller than that under imperfect testing.

Next, we prove some lemmas that will be useful in proving the dispersion and quality uncertainty results in Proposition 4.5.1 under the reactive adulteration case. Again note that $f(x, m, p_L)$ is the probability density function of a normal distribution with mean mp and variance $mp(1-p)$ evaluated at x . Since normal distribution is a good approximation in our case, we will assume $\int_0^m f(x, m, p_L) dx = 1$. Define, $\beta' \equiv \frac{\partial \beta}{\partial m}$

Lemma C.4.3. $\int_0^\beta x f(x, m, p_L) dx = \int_0^\beta \int_t^\beta f(x, m, p_L) dx dt$

Using interchange in order of integration,

$$\int_0^\beta x f(x, m, p_L) dx = \int_0^\beta \int_0^x f(x, m, p_L) dt dx = \int_0^\beta \int_t^\beta f(x, m, p_L) dx dt$$

Lemma C.4.4. $\int_\beta^\infty x f(x, m, p_L) dx = \int_0^\beta \int_\beta^\infty f(x, m, p_L) dx dt + \int_\beta^\infty \int_t^\infty f(x, m, p_L) dx dt$

Using interchange in order of integration,

$$\begin{aligned} \int_\beta^\infty x f(x, m, p_L) dx &= \int_\beta^\infty \left(\int_0^\beta f(x, m, p_L) dt + \int_\beta^x f(x, m, p_L) dt \right) dx \\ &= \int_\beta^\infty \int_0^\beta f(x, m, p_L) dt dx + \int_\beta^\infty \int_\beta^x f(x, m, p_L) dt dx \\ &= \int_0^\beta \int_\beta^\infty f(x, m, p_L) dx dt + \int_\beta^\infty \int_t^\infty f(x, m, p_L) dx dt \end{aligned}$$

Lemma C.4.5. $\frac{\partial}{\partial m} \left(\int_0^\beta x f(x, m, p_L) dx \right) = \beta \frac{\partial}{\partial m} \left(\int_0^\beta f(x, m, p_L) dx \right) + \int_0^\beta \frac{(t+mp)f(t, m, p)}{2m} dt$

Using lemma C.4.3,

$$\frac{\partial}{\partial m} \int_0^\beta x f(x, m, p_L) dx = \frac{\partial}{\partial m} \left(\int_0^\beta \int_t^\beta f(x, m, p_L) dx dt \right)$$

$$\begin{aligned}
 &= \frac{\partial}{\partial m} \left(\int_0^\beta \int_0^\beta f(x, m, p_L) dx dt - \int_0^\beta \int_0^t f(x, m, p_L) dx dt \right) \\
 &= \frac{\partial}{\partial m} \left(\beta \int_0^\beta f(x, m, p_L) dx - \int_0^\beta \int_0^t f(x, m, p_L) dx dt \right) \\
 &= \beta' \int_0^\beta f(x, m, p_L) dx + \beta \frac{\partial}{\partial m} \int_0^\beta f(x, m, p_L) dx - \int_0^\beta \frac{\partial}{\partial m} \int_0^t f(x, m, p_L) dx dt - \\
 &\quad \beta' \int_0^\beta f(x, m, p_L) dx \\
 &= \beta \frac{\partial}{\partial m} \left(\int_0^\beta f(x, m, p_L) dx \right) - \int_0^\beta \frac{\partial}{\partial m} \left(\int_0^{\frac{t-mp}{\sqrt{mp(1-p)}}} \frac{e^{-(x^2/2)}}{\sqrt{2\pi}} dx \right) dt \\
 &= \beta \frac{\partial}{\partial m} \left(\int_0^\beta f(x, m, p_L) dx \right) + \int_0^\beta \frac{(t+mp)f(t, m, p)}{2m} dt
 \end{aligned}$$

Lemma C.4.6. $\frac{\partial}{\partial m} \left(\int_\beta^\infty x f(x, m, p_L) dx \right) = \beta \frac{\partial}{\partial m} \left(\int_\beta^\infty f(x, m, p_L) dx \right) + \int_\beta^\infty \frac{t+mp}{2m} f(t, m, p) dt$

Using lemma C.4.4,

$$\begin{aligned}
 \frac{\partial}{\partial m} \int_\beta^\infty x f(x, m, p_L) dx &= \frac{\partial}{\partial m} \left(\int_0^\beta \int_\beta^\infty f(x, m, p_L) dx dt + \int_\beta^\infty \int_t^\infty f(x, m, p_L) dx dt \right) \\
 &= \frac{\partial}{\partial m} \left(\beta \int_\beta^\infty f(x, m, p_L) dx + \int_\beta^\infty \int_t^\infty f(x, m, p_L) dx dt \right) \\
 &= \beta' \int_\beta^\infty f(x, m, p_L) dx + \beta \frac{\partial}{\partial m} \int_\beta^\infty f(x, m, p_L) dx + \int_\beta^\infty \frac{\partial}{\partial m} \int_t^\infty f(x, m, p_L) dx dt \\
 &\quad - \beta' \int_\beta^\infty f(x, m, p_L) dx \\
 &= \beta \frac{\partial}{\partial m} \left(\int_\beta^\infty f(x, m, p_L) dx \right) + \int_\beta^\infty \frac{\partial}{\partial m} \left(\int_{\frac{t-mp}{\sqrt{mp(1-p)}}}^\infty \frac{e^{-(x^2/2)}}{\sqrt{2\pi}} dx \right) dt \\
 &= \beta \frac{\partial}{\partial m} \left(\int_\beta^\infty f(x, m, p_L) dx \right) + \int_\beta^\infty \frac{(t+mp)f(t, m, p)}{2m} dt
 \end{aligned}$$

Lemma C.4.7. Let $P_m \equiv P(n_L \leq \beta_m)$. Then, $\frac{\partial P_m}{\partial m} \leq 0$ if $\frac{\partial \beta_m}{\partial m} \leq \frac{\beta+mp}{2m}$

$$\begin{aligned}
 P_m &= P(n_L \leq \beta_m) \\
 &= P\left(\mathcal{N}(0, 1) \leq \frac{\beta_m - mp}{\sqrt{mp(1-p)}}\right) \\
 \frac{\partial P_m}{\partial m} &= f(\beta, m, p) \left(\frac{\partial \beta_m}{\partial m} - \frac{\beta + mp}{2m} \right)
 \end{aligned} \tag{C.11}$$

Thus, if $\frac{\partial \beta}{\partial m} \leq \frac{\beta+mp}{2m}$ then $\frac{\partial P_m}{\partial m} \leq 0$ as well.

Lemma C.4.8. Let $P_m \equiv P(n_L \leq \beta_m)$. Then, $\frac{\partial P_m}{\partial p} \leq 0$ if $\frac{\partial \beta}{\partial p} \leq \frac{\beta(1-2p)+mp}{2p(1-p)}$

$$\begin{aligned}
 P_m &= P(n_L \leq \beta) = P\left(\mathcal{N}(0, 1) \leq \frac{\beta_m - mp}{\sqrt{mp(1-p)}}\right) \\
 \frac{\partial P_m}{\partial p} &= f(\beta, m, p) \left(\frac{\partial \beta}{\partial p} - \frac{\beta(1-2p) + mp}{2p(1-p)} \right)
 \end{aligned} \tag{C.12}$$

Thus, if $\frac{\partial \beta}{\partial p} \leq \frac{\beta(1-2p) + mp}{2p(1-p)}$ then $\frac{\partial P_m}{\partial p} \leq 0$ as well.

Proof of Proposition 4.4.1

First we show the results for the preemptive case under both perfect and imperfect testing.

First we prove the result for perfect testing. From Theorem 4.3.1, we have that if $t = n$ (full traceability) $n_a^* = \min \left\{ n, \left\lceil \frac{(r_H - r_L)(p_L^{\max} - p_L^{\min})n}{cq} - 1 \right\rceil \right\}$ and if $t < n$ (partial traceability) $n_a^* = \min \left\{ n, \left\lceil \frac{(r_H - r_L)(p_L^{\max} - p_L^{\min})n^2}{cqt} - 1 \right\rceil \right\}$. Thus, it is straightforward to see that n_a^*/n increases with n under partial traceability and remains constant under full traceability..

Next, we prove the results for reactive case. First we prove the results for P_n for both fully ($t = n$) and partially ($t < n$) traceable supply chains. From Theorem 4.3.3 we have that $\beta^{RP} = \frac{qcm(t/n)}{r_H - r_L}$ and $P_n = \text{Prob}(n_L > \beta^{RP})$. If $t = n$, $\beta^{RV} = \frac{qcm}{r_H - r_L}$ and $\frac{\partial P_n}{\partial n} = \binom{k}{n^2} f\left(\frac{cqk}{n(r_H - r_L)}, \frac{k}{n}, p_L\right) \left(\frac{cq - p_L(r_H - r_L)}{2(r_H - r_L)}\right)$. If $t < n$, $\beta^{RP} = \frac{qcm(t/n)}{r_H - r_L}$ and $\frac{\partial P_n}{\partial n} = \binom{k}{n^2} f\left(\frac{tcqk}{n^2(r_H - r_L)}, \frac{k}{n}, p_L\right) \left(\frac{3cqt - np(r_H - r_L)}{2n(r_H - r_L)}\right)$

Next, we prove the results for E_n . We will prove this result by showing that $\frac{\partial \mathbf{E}_m}{\partial m} \leq 0$. Since $m = k/n$, this implies that $\frac{\partial \mathbf{E}_n}{\partial n} = \frac{\partial m}{\partial n} \frac{\partial \mathbf{E}_m}{\partial m} = -\frac{k}{n^2} \frac{\partial \mathbf{E}_n}{\partial m} \geq 0$ in all the cases. Let X_i^m be the amount of supply adulterated by farm i when he supplies a total of m units. Then total supply that is adulterated is just $\sum_{i=1}^n X_i^m$. For fully traceable networks we have $\beta^{RP} = \frac{qcm}{r_H - r_L}$. Let $\alpha \equiv \frac{cq}{r_H - r_L}$. Using lemma C.4.6, if $c \geq \frac{p_L(r_H - r_L)}{q}$,

$$\begin{aligned}
 \mathbf{E}_m &= \mathbf{E}\left[\sum_{i=1}^n X_i^m\right] = \sum_{i=1}^n \mathbf{E}[X_i^m] = n\mathbf{E}[X_i^m] = \frac{k}{m} \int_{\beta}^{\infty} x f(x, m, p_L) dx \\
 \frac{\partial \mathbf{E}_m}{\partial m} &= \frac{k}{m} \frac{\partial}{\partial m} \left(\int_{\beta}^{\infty} x f(x, m, p_L) dx \right) - \frac{k}{m^2} \int_{\beta}^{\infty} x f(x, m, p_L) dx \\
 &= \frac{k}{m} \left(\beta \frac{\partial}{\partial m} \left(\int_{\beta}^{\infty} f(x, m, p_L) dx \right) + \int_{\beta}^{\infty} \frac{(x + mp_L) f(x, m, p_L)}{2m} dx \right) - \frac{k}{m^2} \int_{\beta}^{\infty} x f(x, m, p_L) dx \\
 &= k\alpha \frac{\partial P_m}{\partial m} - \frac{k}{m^2} \int_{\beta}^{\infty} x f(x, m, p_L) dx + \frac{k}{m} \int_{\beta}^{\infty} \frac{(mp_L + x)}{2m} f(x, m, p_L) dx \\
 &= k\alpha \frac{\partial P_m}{\partial m} + \frac{k}{2m} \int_{\beta}^{\infty} \left(p_L - \frac{x}{m}\right) f(x, m, p_L) dx \\
 &\leq k\alpha \frac{\partial P_m}{\partial m} + \frac{k}{2m} \int_{\beta}^{\infty} \left(p_L - \frac{\beta}{m}\right) f(x, m, p_L) dx \\
 &= k\alpha \frac{\partial P_m}{\partial m} + \frac{k}{2m} \left(p_L - \frac{cq}{r_H - r_L}\right) \int_{\beta}^{\infty} f(x, m, p_L) dx
 \end{aligned}$$

$$\leq 0$$

The last inequality follows because $\frac{\partial P_m}{\partial m} \leq 0$ when $c \geq \frac{p_L(r_H - r_L)}{q}$ from the fact that P_m is decreasing in m in this range.

Now we consider partially traceable networks. If $qc \geq \frac{p_L(r_H - r_L)}{3(t/n)q}$ then this implies $m \geq \frac{kp(r_H - r_L)}{3cqt}$. We will prove this result by considering two cases. Let $\alpha \equiv \frac{cqt}{k(r_H - r_L)}$. Here,

$$\beta^{RP} = \frac{qcm(t/n)}{r_H - r_L} = \frac{cqm^2t}{k(r_H - r_L)} = \alpha m^2.$$

Case 1. If $\frac{kp_L(r_H - r_L)}{3cqt} < m < \frac{kP_L(r_H - r_L)}{cqt}$

$$\begin{aligned} \mathbf{E}_m &= \mathbf{E}\left[\sum_{i=1}^n X_i^m\right] = \sum_{i=1}^n \mathbf{E}[X_i^m] = n\mathbf{E}[X_i^m] = \frac{k}{m} \int_{\beta}^{\infty} xf(x, m, p_L)dx \\ &= kp - \frac{k}{m} \int_0^{\beta} xf(x, m, p_L)dx \\ \frac{\partial \mathbf{E}_m}{\partial m} &= \frac{k}{m^2} \int_0^{\beta} xf(x, m, p_L)dx - \frac{k}{m} \frac{\partial \left(\int_0^{\beta} xf(x, m, p_L)dx \right)}{\partial m} \\ &= \frac{k}{m^2} \int_0^{\beta} xf(x, m, p_L)dx - \frac{k}{m} \left(\beta \frac{\partial}{\partial m} \left(\int_0^{\beta} f(x, m, p_L)dx \right) + \int_0^{\beta} \frac{(y + mp)f(y, m, p_L)}{2m} dy \right) \\ &= -\frac{k}{m} \beta \frac{\partial(1 - P_m)}{\partial m} + \frac{k}{2m} \int_0^{\beta} \left(\frac{x}{m} - p_L \right) f(x, m, p_L)dx \\ &\leq -\frac{k}{m} \beta \frac{\partial(1 - P_m)}{\partial m} + \frac{k}{2m} \left(\frac{\beta}{m} - p_L \right) \int_0^{\beta} f(x, m, p_L)dx \\ &= -\frac{k}{m} \beta \frac{\partial(1 - P_m)}{\partial m} + \frac{k}{2m} \left(\frac{cqt m}{k(r_H - r_L)} - p_L \right) \int_0^{\beta} f(t, m, p_L)dt \\ &\leq 0 \end{aligned}$$

The second equality follows from lemma C.4.5. The last inequality follows from the assumption that $\frac{kp_L(r_H - r_L)}{3cqt} < m < \frac{kP_L(r_H - r_L)}{cqt}$ and the fact that P_m is decreasing in m in this range.

Case 2. If $m \geq \frac{kP_L(r_H - r_L)}{cqt}$ Following the same procedure as above, we get

$$\begin{aligned} \frac{\partial \mathbf{E}_m}{\partial m} &= k\beta \frac{\partial P_m}{\partial m} + \frac{k}{2m} \int_{\beta}^{\infty} \left(p_L - \frac{x}{m} \right) f(x, m, p_L)dx \\ &\leq k\beta \frac{\partial P_m}{\partial m} + \frac{k}{2m} \int_{\beta}^{\infty} \left(p_L - \frac{\beta}{m} \right) f(x, m, p_L)dx \\ &= k\beta \frac{\partial P_m}{\partial m} + \frac{k}{2m} \left(p_L - \frac{cqm t}{k(r_H - r_L)} \right) \int_{\beta}^{\infty} f(x, m, p_L)dx \\ &\leq 0 \end{aligned}$$

The last inequality follows from the assumption that $m > \frac{kp(r_H - r_L)}{cqt}$ and the fact that P_m is decreasing in m in this range.

Proof of Proposition 4.4.2

Under imperfect testing with preemptive EMA, we have from Theorem 4.3.2 that x^{PV^*} can either be the solution to the first order equation or at the boundaries. We will show that $x_{n+1}^{PV^*} \geq x_n^{PV^*}$ for all n in all these cases.

Consider the case when $x_n^{PV^*}$ is the solution to the FOC and $t = n$. Then from theorem 4.3.2, we have $-h'(x_n^{PV^*})/x_n^{PV^*} = q((n+1)/n)c/(r_H - r_L)$. If $x_{n+1}^{PV^*}$ is a solution to the FOC then $-h'(x_{n+1})/x_{n+1} = q((n+2)/(n+1))c/(r_H - r_L)$. Let $F(x, n) \equiv xq((n+1)/(n))c/(r_H - r_L) + h'(x)$. Note that $F(x_n^{PV^*}, n) = 0$ and $F(x_{n+1}^{PV^*}, n+1) = 0$ and $F(0, n) < 0$ since h is a decreasing function. Since h is a convex function, $F(x, n)$ is monotonically increasing in x . Further, $F(x_{n+1}^{PV^*}, n) = c/(r_H - r_L)x_{n+1}^{PV^*}q((n+1)/n) - (n+2)/(n+1) = c/(r_H - r_L)x_{n+1}^{PV^*}q(1/(n(n+1))) > 0$ using the definition of F and the fact that $F(x_{n+1}^{PV^*}, n+1) = 0$. Since $F(x, n)$ is monotonically increasing in x , and $F(x_n^{PV^*}, n) = 0 < F(x_{n+1}, n)$, we have $x_{n+1}^{PV^*} \geq x_n^{PV^*}$ in this case as well. If $t < n$ then letting $F(x, n) \equiv xq(t/n)((n+1)/n)c/(r_H - r_L) + h'(x)$ and redoing the analysis yields the same result. If $x_{n+1}^{PV^*} = 1$ then since $x_n^{PV^*} \leq 1$, we again have $x_{n+1}^{PV^*} \geq x_n^{PV^*}$. Finally, if $x_n^{PV^*} = 1$ then since $-h'(1)(r_H - r_L)/[q(t/n)((n+1)/n)]$ is increasing in n , $x_{n+1}^{PV^*}$ will be 1 for $n+1$ as well.

Next we show the results for reactive EMA. We prove that $\frac{\partial P_n}{\partial n} \geq 0$ by showing that $\frac{\partial P_m}{\partial m} \leq 0$. From lemma C.4.7 we only need to show that $\frac{\partial \beta_m}{\partial m} \leq \frac{\beta + mp_L}{2m}$ in order to show $\frac{\partial P_m}{\partial m} \leq 0$.

For fully traceable supply chains ($t=n$) and $\beta^{RV} = \frac{k(r_H - r_L)}{cq} - (\frac{k}{m} - 1) \int_0^{\beta^{RV}} xf(x, m, p_L)dx$

Case 1: Assume $\beta^{RV} \leq mp_L$.

$$\begin{aligned} \text{Let } g(\beta^{RV}, m) &= \frac{k(r_H - r_L)}{cq} - (\frac{k}{m} - 1) \int_0^{\beta^{RV}} xf(x, m, p_L)dx - \beta^{RV} \\ \frac{\partial g}{\partial m} &= \frac{k}{m^2} \int_0^{\beta^{RV}} xf(x, m, p_L)dx - (\frac{k}{m} - 1) \frac{\partial}{\partial m} \left(\int_0^{\beta^{RV}} xf(x, m, p_L)dx \right) \\ &= \frac{k}{m^2} \int_0^{\beta^{RV}} xf(x, m, p_L)dx - \left(\frac{k}{m} - 1 \right) \left(\beta^{RV} \frac{\partial}{\partial m} \left(\int_0^{\beta^{RV}} f(x, m, p_L)dx \right) \right. \\ &\quad \left. + \int_0^{\beta^{RV}} \frac{(t + mp_L)f(t, m, p_L)}{2m} dt \right) \\ &= \left(\frac{k}{m} - 1 \right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m} \right) + \frac{k}{m^2} \int_0^{\beta^{RV}} tf(t, m, p_L)dt \end{aligned}$$

$$\begin{aligned}
 & -\left(\frac{k}{m} - 1\right) \int_0^{\beta^{RV}} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \\
 & = \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_0^{\beta^{RV}} (t - mp) f(t, m, p_L) dt \\
 & + \int_0^{\beta^{RV}} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \\
 \frac{\partial g}{\partial \beta^{RV}} & = -1 - \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \\
 \frac{\partial \beta^{RV}}{\partial m} & = \frac{\left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_0^{\beta^{RV}} (t - mp) f(t, m, p_L) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)} \\
 & + \frac{\int_0^{\beta^{RV}} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)}
 \end{aligned}$$

Since $\beta^{RV} \leq mp_L$ from assumption, we have

$$\begin{aligned}
 \frac{k}{2m^2} \int_0^{\beta^{RV}} (t - mp) f(t, m, p) dt + \int_0^{\beta^{RV}} f(t, m, p) \left(\frac{t + mp_L}{2m}\right) dt & \leq \int_0^{\beta^{RV}} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \\
 & \leq \left(\frac{\beta^{RV} + mp_L}{2m}\right) \int_0^{\beta^{RV}} f(t, m, p_L) dt \leq \left(\frac{\beta^{RV} + mp_L}{2m}\right)
 \end{aligned}$$

Adding $\left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right)$ on both sides, we get

$$\begin{aligned}
 & = \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_0^{\beta^{RV}} (t - mp_L) f(t, m, p_L) dt \\
 & + \int_0^{\beta^{RV}} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \leq \frac{\beta^{RV} + mp_L}{2m} \left(1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)\right) \\
 & = \frac{\left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_0^{\beta^{RV}} (t - mp_L) f(t, m, p_L) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)} \\
 & + \frac{\int_0^{\beta^{RV}} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)} \\
 & \leq \frac{\beta^{RV} + mp_L}{2m}
 \end{aligned}$$

This implies $\frac{\partial \beta^{RV}}{\partial m} \leq \frac{\beta^{RV} + mp}{2m}$. From lemma C.4.7 we get that since $\frac{\partial \beta^{RV}}{\partial m} \leq \frac{\beta^{RV} + mp}{2m}$, $\frac{\partial P_m}{\partial m} \leq 0$ as well.

Case 2: Assume $\beta^{RV} > mp_L$.

$$\begin{aligned}
 \beta^{RV} & = \frac{k(r_H - r_L)}{cq} - \left(\frac{k}{m} - 1\right) \int_0^{\beta^{RV}} x f(x, m, p_L) dx \\
 & = \frac{k(r_H - r_L)}{cq} - kp + mp_L + \left(\frac{k}{m} - 1\right) \left(\frac{k}{m} - 1\right) \left(\int_{\beta^{RV}}^{\infty} x f(x, m, p_L) dx\right)
 \end{aligned}$$

$$\begin{aligned}
 g(\beta^{RV}, m) &= \frac{k(r_H - r_L)}{cq} - kp_L + mp_L + \left(\frac{k}{m} - 1\right) \left(\int_{\beta^{RV}}^{\infty} xf(x, m, p_L)dx\right) - \beta^{RV} \\
 \frac{\partial g}{\partial m} &= p_L + \left(\frac{k}{m} - 1\right) \frac{\partial}{\partial m} \left(\int_{\beta^{RV}}^{\infty} xf(x, m, p_L)dx\right) - \frac{k}{m^2} \left(\int_{\beta^{RV}}^{\infty} xf(x, m, p_L)dx\right) \\
 &= p_L + \left(\frac{k}{m} - 1\right) \left(\beta^{RV} \frac{\partial}{\partial m} \left(\int_{\beta^{RV}}^{\infty} f(x, m, p_L)dx\right) + \int_{\beta^{RV}}^{\infty} \frac{t + mp_L}{2m} f(t, m, p_L)dt\right) \\
 &\quad - \frac{k}{m^2} \left(\int_{\beta^{RV}}^{\infty} xf(x, m, p_L)dx\right) \\
 &= p_L + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) - \frac{k}{m^2} \int_{\beta^{RV}}^{\infty} tf(t, m, p_L)dt \\
 &\quad + \left(\frac{k}{m} - 1\right) \int_{\beta^{RV}}^{\infty} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \\
 &= p_L + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) \\
 &\quad + \frac{k}{2m^2} \int_{\beta^{RV}}^{\infty} (mp_L - t) f(t, m, p_L) dt - \int_{\beta^{RV}}^{\infty} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \\
 \frac{\partial g}{\partial \beta^{RV}} &= -1 - \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \\
 \frac{\partial \beta^{RV}}{\partial m} &= \frac{p_L + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_{\beta^{RV}}^{\infty} (mp_L - t) f(t, m, p_L) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)} \\
 &\quad - \frac{\int_{\beta^{RV}}^{\infty} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)}
 \end{aligned}$$

Since $\beta^{RV} > mp$ from assumption, we have

$$\begin{aligned}
 &p_L + \frac{k}{2m^2} \int_{\beta^{RV}}^{\infty} (mp_L - t) f(t, m, p_L) dt - \int_{\beta^{RV}}^{\infty} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \\
 &\leq p_L = \frac{mp_L + mp_L}{2m} < \frac{\beta^{RV} + mp_L}{2m}
 \end{aligned}$$

Adding $\left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right)$ on both sides, we get

$$\begin{aligned}
 &= p_L + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_{\beta^{RV}}^{\infty} (mp_L - t) f(t, m, p_L) dt \\
 &\quad - \int_{\beta^{RV}}^{\infty} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt \leq \frac{\beta^{RV} + mp_L}{2m} \left(1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)\right) \\
 &= \frac{p_L + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L) \left(\frac{\beta^{RV} + mp_L}{2m}\right) + \frac{k}{2m^2} \int_{\beta^{RV}}^{\infty} (mp_L - t) f(t, m, p_L) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)} \\
 &\quad - \frac{\int_{\beta^{RV}}^{\infty} f(t, m, p_L) \left(\frac{t + mp_L}{2m}\right) dt}{1 + \left(\frac{k}{m} - 1\right) \beta^{RV} f(\beta^{RV}, m, p_L)} \leq \frac{\beta^{RV} + mp_L}{2m}
 \end{aligned}$$

This implies $\frac{\partial \beta^{RV}}{\partial m} \leq \frac{\beta^{RV} + mp_L}{2m}$. From lemma C.4.7 we get that since $\frac{\partial \beta^{RV}}{\partial m} \leq \frac{\beta^{RV} + mp_L}{2m}$, $\frac{\partial P_m}{\partial m} \leq 0$ as well.

Let β_f^{RV} be the threshold for fully traceable systems. For partially traceable systems, $\beta_p^{RV} = \frac{k^2(r_H-r_L)}{cqmt} - (\frac{k}{m}-1) \int_0^{\beta^{RV}} xf(x, m, p_L)dx = \frac{k^2(r_H-r_L)}{cqmt} + \beta_f^{RV} - \frac{k(r_H-r_L)}{cq}$. Then $\frac{\partial \beta_p^{RV}}{\partial m} = -\frac{k^2(r_H-r_L)}{cqm^2t} + \frac{\partial \beta_f^{RV}}{\partial m} < \frac{\beta^{RV} + mp_L}{2m}$. Thus for both partially and fully traceable systems P_n increases as n increases if β^{RV} is a solution to the fixed point equation.

Next, we will show that $\frac{\partial E_m}{\partial m} \leq 0$ to prove that $\frac{\partial E_n}{\partial n} \geq 0$. Let X_i^m be the total supply adulterated by supplier $i=1, \dots, n$ when each supplier supplies m units. For fully traceable systems, $\beta = \frac{k(r_H-r_L)}{cq} - (\frac{k}{m}-1) \int_0^{\beta} xf(x, m, p_L)dx$. Let $\alpha = \frac{k(r_H-r_L)}{cq}$. Then $\int_0^{\beta} xf(x, m, p_L)dx = \frac{\alpha-\beta}{n-1}$. Then for fully traceable systems,

$$\begin{aligned} E_m &= E\left[\sum_{i=1}^n X_i^m\right] = \sum_{i=1}^n E[X_i^m] = nE[X_i^m] = \frac{k}{m} \int_0^{\beta} xf(x, m, p_L)dx \\ \frac{\partial E_m}{\partial m} &= -\frac{k}{m^2} \int_0^{\beta} xf(x, m, p_L)dx + \frac{k}{m} \left(\beta \frac{\partial}{\partial m} \left(\int_0^{\beta} f(x, m, p_L)dx \right) + \int_0^{\beta} \frac{(t+mp_L)f(t, m, p_L)}{2m} \right) \\ &= -\frac{k}{m^2} \int_0^{\beta} xf(x, m, p_L)dx + \frac{k}{m} \beta \frac{\partial P_m}{\partial m} + \frac{k}{m} \int_0^{\beta} f(t, m, p) \left(\frac{t+mp_L}{2m} \right) dt \\ &= -\frac{k}{2m^2} \int_0^{\beta} xf(x, m, p_L)dx + \frac{k}{m} \beta \frac{\partial P_m}{\partial m} + \frac{kp_L}{2m} \int_0^{\beta} f(t, m, p_L)dt \\ &= -\frac{k}{2m^2} \int_0^{\beta} xf(x, m, p_L)dx + \frac{k}{m} \beta \frac{\partial P_m}{\partial m} + \frac{kp_L}{2m} \int_0^{\beta} f(t, m, p_L)dt \end{aligned}$$

Replacing the value of $\frac{\partial P_m}{\partial m}$ and $\frac{\partial \beta}{\partial m}$, we get that showing $\frac{\partial E_m}{\partial m} \leq 0$ is equivalent to showing:

$$\begin{aligned} 0 &\leq \beta^2 f(\beta, m, p_L) + mp \left(\beta f(\beta, m, p_L) - \int_0^{\beta} f(x, m, p_L)dx \right) + \\ &\quad \left(1 - 2\beta f(\beta, m, p_L) \right) \int_0^{\beta} xf(x, m, p_L)dx \end{aligned} \quad (\text{C.13})$$

Note that $\int_0^{\beta} xf(x, m, p_L)dx = mp_L(1-p_L)f(\beta, m, p_L) + mp_L \int_0^{\beta} f(x, m, p_L)dx$ using integration by parts and the assumption that $\int_{-\infty}^0 xf(x, m, p_L)dx = 0$. Replacing the value in equation C.13, and after some algebraic simplifications, we get that $\frac{\partial E_m}{\partial m} \leq 0$ is equivalent to showing:

$$0 \leq \beta \left(\beta + mp_L - 2mp_L \int_0^{\beta} f(x, m, p_L)dx \right) + mp_L(1-p_L) \left(1 - 2\beta f(\beta, m, p_L) \right) \quad (\text{C.14})$$

$$\Leftrightarrow 0 \leq \beta + mp_L - 2mp_L \int_0^{\beta} f(x, m, p_L)dx + mp_L(1-p_L) \left(\frac{1}{\beta} - 2f(\beta, m, p_L) \right) \quad (\text{C.15})$$

We will show that the inequalities hold by considering different cases:

Case 1: Assume $\beta \leq mp_L$.

(i) If $1 - 2\beta f(\beta, m, p_L) > 0$, then second term in equation C.14 is non negative. Since

$\beta \leq mp_L$, $\int_0^\beta f(x, m, p_L) dx \leq 1/2$ and the first term is also non negative because $\beta + mp_L - 2mp_L \int_0^\beta x f(x, m, p_L) dx \geq \beta + mp_L - mp_L \geq 0$.

- (ii) If $1 - 2\beta f(\beta, m, p_L) \leq 0$, then $1 - 2mp_L f(\beta, m, p_L) \leq 1 - 2\beta f(\beta, m, p_L) \leq 0$. Differentiating the RHS in equation C.15 w.r.t. β , we get $1 - 2mp_L f(\beta, m, p_L) - mp_L(1 - p_L) \left(\frac{1}{\beta^2} + 2f(\beta, m, p_L) \frac{(mp_L - \beta)}{mp_L(1 - p_L)} \right)$. Since $1 - 2mp_L f(\beta, m, p_L) < 0$ and $\beta \leq mp_L$, the differential is always negative and the expression is decreasing in β . Thus the smallest value of the RHS is when $\beta = \min(mp_L, \beta^*)$ where β^* is the largest β s.t. $1 - 2\beta f(\beta, m, p_L) \leq 0$. Since $\beta f(\beta, m, p_L)$ is increasing in β for $\beta \leq mp_L$ and $mp_L > 5$ from assumption, it is easy to check that $1 - 2mp_L f(mp_L, m, p_L) > 0$ and $\min(mp_L, \beta^*) = \beta^*$. Note that at β^* , $1 - 2\beta^* f(\beta^*, m, p_L) = 0$ since $\beta f(\beta, m, p_L)$ is increasing in β and we want the largest β s.t. $1 - 2\beta f(\beta, m, p_L) \leq 0$. Evaluating at β^* , the second term in equation C.15 is 0 while the first term is again non negative. Finally, since the RHS is non negative at β^* , it must be non negative for all $\beta < \beta^*$.

Case 2: Assume $\beta > mp_L$.

- (i) If $1 - 2\beta f(\beta, m, p_L) \geq 0$, then the second term equation C.15 is non negative and the first term is increasing in β (this is because $1 - 2mp_L f(\beta, m, p_L) \geq 1 - 2\beta f(\beta, m, p_L) \geq 0$). Since the first term evaluated at $\beta = mp_L$ is positive, it must be positive for all $\beta > mp_L$.
- (ii) If $1 - 2\beta f(\beta, m, p_L) < 0$, then since $\int_0^\beta x f(x, m, p_L) dx \leq mp_L$ and $\int_0^\beta f(x, m, p_L) dx \leq 1$, from equation C.13 we have that $\beta^2 f(\beta, m, p_L) + mp \left(\beta f(\beta, m, p_L) - \int_0^\beta f(x, m, p_L) dx \right) + \left(1 - 2\beta f(\beta, m, p_L) \right) \int_0^\beta x f(x, m, p_L) dx > \beta^2 f(\beta, m, p_L) + mp_L \left(\beta f(\beta, m, p_L) - 1 \right) + \left(1 - 2\beta f(\beta, m, p_L) \right) mp_L = \beta f(\beta, m, p_L) (\beta - mp_L) \geq 0$.

For partially traceable systems let $\alpha = \frac{k^2(r_H - r_L)}{cqmt}$. Note that α is decreasing in m and since we have already shown that $\frac{\partial E_m}{\partial m} \leq 0$ under full traceability (with a constant α), E_m must be decreasing in m in the partial traceability case as well.

Proof of Proposition 4.5.1

Under preemptive EMA with perfect testing, as p_L^{\max} increases, it is straightforward to see that the threshold $n_a^* = \min \left\{ n, \left\lceil \frac{(r_H - r_L)(p_L^{\max} - p_L^{\min})n^2}{cqt} - 1 \right\rceil \right\}$ also increases. Under imperfect testing, since $h(1) = p_L^{\min}$ stays constant and $h(x)$ is a convex decreasing function, this implies that $\frac{\partial h'(x)}{\partial p_L^{\max}} \leq 0$ for all x . Thus, $-h'(1)(r_H - r_L)/[q(t/n)((n+1)/n)]$ increases as p_L^{\max} increases and the range in which $x^{PV^*} = 1$ also increases.

Under reactive EMA, we will first prove the results for P_n . Note that we will use p and p_L interchangeably in all the proofs that follow. Under perfect testing, $\beta^{RP} = \frac{cqt m^2}{k(r_H - r_L)}$. Thus $\frac{\partial \beta^{RP}}{\partial p} = 0$. Using lemma C.4.8, we have that $\frac{\partial(1-P_m)}{\partial p} \leq 0$ if $\frac{\beta(1-2p)+mp}{2p(1-p)} \geq \frac{\partial \beta}{\partial p} = 0$. We will show that this is indeed the case for both fully and partially traceable networks. Under fully traceable networks, if $cq > (r_H - r_L)$ then $P_m = 0$ as $\beta^{RP} = m$. If $cq \leq (r_H - r_L)$, $\frac{\beta(1-2p)+mp}{2p(1-p)} = \frac{m(cq(1-2p)+p(r_H-r_L))}{2p(1-p)(r_H-r_L)}$. If $p \leq 0.5$ then $cq(1-2p) + p(r_H - r_L)$ is always non-negative. If $p > 0.5$ then $\frac{p(r_H-r_L)}{2p-1} - cq$ is again non-negative since $\frac{p}{2p-1} \geq 1$ and $r_H - r_L \geq cq$ from assumption.

For partially traceable networks, if $m > \frac{k(r_H-r_L)}{cqt}$ then $\beta^{RP} = m$ and $P_m = 0$. If $m \leq \frac{k(r_H-r_L)}{cqt}$, $\frac{\beta(1-2p)+mp}{2p(1-p)} = \frac{m(cqmt(1-2p)+pk(r_H-r_L))}{2kp(1-p)(r_H-r_L)}$. If $p \leq 0.5$ then $m(cqmt(1-2p) + pk(r_H - r_L))$ is always non-negative. If $p > 0.5$ then $\frac{kp(r_H-r_L)}{2p-1} - cqmt$ is again non-negative since $\frac{p}{2p-1} \geq 1$ and $k(r_H - r_L) \geq cqmt$ from assumption.

Under imperfect testing, the threshold is the solution to the following fixed point equation: $\beta^{RV} = \frac{k(r_H-r_L)}{cq} - (\frac{k}{m} - 1) \int_0^\beta x f(x, m, p_L) dx$ Let $g(\beta, p) = \frac{k(r_H-r_L)}{cq} - (\frac{k}{m} - 1) \int_0^\beta x f(x, m, p_L) dx - \beta$ for fully traceable systems and $g(\beta, p) = \frac{k^2(r_H-r_L)}{cqmt} - (\frac{k}{m} - 1) \int_0^\beta x f(x, m, p_L) dx - \beta$ for partially traceable systems.

$$\begin{aligned} \frac{\partial g}{\partial p} &= -\left(\frac{k}{m} - 1\right) \frac{\partial \left(\int_0^\beta x f(x, m, p_L) dx \right)}{\partial p} \\ &= -\left(\frac{k}{m} - 1\right) \frac{\partial \left(\beta \int_0^\beta f(x, m, p_L) dx \right)}{\partial p} - \frac{\partial \left(\int_0^\beta \int_0^t f(x, m, p_L) dx dt \right)}{\partial p} \\ &= \left(\frac{k}{m} - 1\right) \left(\frac{\beta f(\beta, m, p)(\beta(1-2p) + mp)}{2p(1-p)} - \frac{\int_0^\beta (t - 2tp + mp) f(t, m, p) dt}{2p(1-p)} \right) \\ \frac{\partial g}{\partial \beta} &= -1 - \left(\frac{k}{m} - 1\right) \beta f(\beta, m, p) \\ \frac{\partial \beta}{\partial p} &= \frac{\left(\frac{k}{m} - 1\right) \left(\frac{\beta f(\beta, m, p)(\beta(1-2p) + mp)}{2p(1-p)} - \frac{\int_0^\beta (t - 2tp + mp) f(t, m, p) dt}{2p(1-p)} \right)}{1 + \left(\frac{k}{m} - 1\right) \beta f(\beta, m, p)} \end{aligned}$$

Note that we have $-\left(\frac{k}{m} - 1\right) \left(\int_0^\beta (t - 2tp + mp) f(t, m, p) dt \right) \leq \beta(1-2p) + mp$. This is because if $0 < p < 0.5$ then RHS is always positive while LHS is always negative and the inequality is satisfied. If $1 > p \geq 0.5$ then

$$\begin{aligned} -\left(\frac{k}{m} - 1\right) \left(\int_0^\beta (t - 2tp + mp) f(t, m, p) dt \right) &= \left(\frac{k}{m} - 1\right) \left(\int_0^\beta (t(2p-1) - mp) f(t, m, p) dt \right) \\ &\leq \left(\frac{k}{m} - 1\right) \left((\beta(2p-1) - mp) \int_0^\beta f(t, m, p) dt \right) \leq \left(\frac{k}{m} - 1\right) \left(\beta(2p-1) - mp \right) \end{aligned}$$

$$\leq 0 \leq mp + \beta(1 - 2p)$$

Adding $(\frac{k}{m} - 1) \frac{\beta f(\beta, m, p)(\beta(1-2p)+mp)}{2p(1-p)}$ on both sides, we get

$$\begin{aligned} & \left(\frac{k}{m} - 1\right) \left(\frac{\beta f(\beta, m, p)(\beta(1-2p)+mp)}{2p(1-p)} - \frac{\int_0^\beta (t-2tp+mp)f(t, m, p)dt}{2p(1-p)} \right) \\ & \leq \frac{\beta(1-2p)+mp}{2p(1-p)} \left(1 + \left(\frac{k}{m} - 1\right)\beta f(\beta, m, p)\right) \\ & = \frac{\left(\frac{k}{m} - 1\right) \left(\frac{\beta f(\beta, m, p)(\beta(1-2p)+mp)}{2p(1-p)} - \frac{\int_0^\beta (t-2tp+mp)f(t, m, p)dt}{2p(1-p)} \right)}{1 + \left(\frac{k}{m} - 1\right)\beta f(\beta, m, p)} \leq \frac{\beta(1-2p)+mp}{2p(1-p)} \end{aligned}$$

Using lemma C.4.8, we have that $\frac{\partial P_m}{\partial p} \leq 0$ since $\frac{\beta(1-2p)+mp}{2p(1-p)} \geq \frac{\partial \beta}{\partial p}$.

Next we prove results for E_n . Under perfect testing, $\frac{\partial E_m}{\partial p} = \frac{k}{m} \frac{\partial \int_{\beta}^{\infty} x f(x, m, p_L) dx}{\partial p}$.

$$\begin{aligned} \frac{\partial \int_{\beta}^{\infty} x f(x, m, p_L) dx}{\partial p} &= \frac{\partial \left(\int_0^\beta \int_\beta^\infty f(x, m, p_L) dx dt + \int_\beta^\infty \int_t^\infty f(x, m, p_L) dx dt \right)}{\partial p} \\ &= \int_0^\beta \frac{\partial \left(\int_\beta^\infty f(x, m, p_L) dx \right)}{\partial p} dt + \int_\beta^\infty \frac{\partial \left(\int_t^\infty f(x, m, p_L) dx \right)}{\partial p} dt \\ &= \int_0^\beta f(\beta, m, p) \frac{\beta(1-2p)+mp}{2p(1-p)} dt + \int_\beta^\infty f(t, m, p) \frac{t(1-2p)+mp}{2p(1-p)} dt \\ &= \beta f(\beta, m, p) \frac{\beta(1-2p)+mp}{2p(1-p)} + \int_\beta^m f(t, m, p) \frac{t(1-2p)+mp}{2p(1-p)} dt \end{aligned}$$

If $p < .5$ then both the terms in the last equation are positive and we have that $\frac{\partial E_n}{\partial n} \geq 0$. If $p > 0.5$, then the first term is positive since we have already shown $\frac{\partial P_n}{\partial p} = f(\beta, m, p) \frac{\beta(1-2p)+mp}{2p(1-p)} \geq 0$ under perfect testing. Further, $\int_\beta^m f(t, m, p) \frac{t(1-2p)+mp}{2p(1-p)} dt \geq \frac{-m(2p-1)+mp}{2p(1-p)} \int_\beta^m f(t, m, p) dt = \frac{m}{2p} \int_\beta^m f(t, m, p) dt \geq 0$ and again we have $\frac{\partial E_n}{\partial p} \geq 0$.

Proof of Theorem C.1.1

Let us assume that a subset n_a out of a total of n farms are adulterating. Consider a farm i that is adulterating at equilibrium. Then, revenue from adulteration (π_a^i) should be strictly greater than revenue from non adulteration (π_{na}^i). $\pi_a^i = r_H m(1 - p_L^{\min}) + r_L m p_L^{\min} - q \frac{n_a}{n} (t/n) cm - \lambda \bar{F}(\gamma, m, p_L^{\min})$ and $\pi_{na}^i = r_H m(1 - p_L^{\max}) + r_L m p_L^{\max} - \lambda \bar{F}(\gamma, m, p_L^{\max})$. Now consider a farmer i' who is not adulterating at equilibrium. Then, $\pi_{na}^{i'} \geq \pi_a^{i'}$. $\pi_a^{i'} = r_H m(1 - p_L^{\min}) + r_L m p_L^{\min} - q \frac{n_a+1}{n} (t/n) cm - \lambda \bar{F}(\gamma, m, p_L^{\min})$ and $\pi_{na}^{i'} = r_H m(1 - p_L^{\max}) + r_L m p_L^{\max} - \lambda \bar{F}(\gamma, m, p_L^{\max})$. Simplifying the two conditions we get the desired result in the theorem.

Proof of Theorem C.1.2

Differentiating the revenue function for risk averse farmers w.r.t x , we get:

$$\begin{aligned}\frac{\partial \pi_L^{PP}(x)}{\partial x} &= -mh'(x)(r_H - r_L) - q\frac{t}{n}cm\left(\frac{\sum_{-i} x_{-i}^* + 2x_i}{n}\right) \\ &\quad - h'(x_i)\frac{\lambda mf(\gamma m, m, p)(\gamma(1 - 2h(x_i)) + h(x_i))}{2h(x_i)(1 - h(x_i))} \\ \frac{\partial^2 \pi_L^{PV}(x_i, x_{-i})}{\partial x_i^2} &= -mh''(x)(r_H - r_L) - q\frac{t}{n}cm\left(\frac{2}{n}\right) - h''(x_i)\frac{\lambda mf(\gamma m, m, p)(\gamma(1 - 2h(x_i)) + h(x_i))}{2h(x_i)(1 - h(x_i))} \\ &\quad - h'(x_i)^2 \lambda mf(\gamma m, m, p)\frac{-\gamma + h(x)(2\gamma + h(x)(1 - 2\gamma))}{2(1 - h(x))^2 h(x)^2}\end{aligned}$$

Note that the first three terms are less than 0 because of the convexity of $h(x)$, and the last term is again less than 0 because $-\gamma + h(x)(2\gamma + h(x)(1 - 2\gamma)) \geq -\gamma + p_L^{\min}(2\gamma + p_L^{\min}(1 - 2\gamma)) \geq 0$ from the assumption that $\gamma \leq \frac{p_L^{\min}}{1 - 2p_L^{\min}(1 - p_L^{\min})}$. Since $\pi_L^{PV}(x_i, x_{-i})$ is a concave function and we are considering symmetric equilibrium, replacing $x_{-i}^* = x^* = x_L^{PV*}$ gives us the desired result.

Proof of Proposition C.1.3

Let $T^O = n^2 \frac{(R_H - r_L)(p_L^{\max} - p_L^{\min})}{cqt}$. From Theorem 4.3.1 we have that $n_a^* = \min\{n, \lceil T^O - 1 \rceil\}$. From lemma C.4.8 we have that $\bar{F}(\gamma, m, p)$ is increasing in p since $\gamma m(1 - 2p) + mp$ is positive for all p . This implies that $\lambda(\bar{F}(\gamma, m, p_L^{\max}) - \bar{F}(\gamma, m, p_L^{\min}))$ is positive and $T^{RA} > T$. Since $n_a^{RA*} = \min\{n, \lceil T^{RA} - 1 \rceil\}$ and $T^{RA} > T$ we have that $n_a^{RA*} > n_a^*$. Similarly, since $(\gamma(1 - 2h(x)) + h(x))$ is positive for all x , we again have that $\lambda f(\gamma m, m, h(x))\frac{(\gamma(1 - 2h(x)) + h(x))}{2h(x)(1 - h(x))}$ is positive. Following the same logic as above, we have that $x_{RA}^{PV*} > x^{PV*}$.

Proof of Proposition C.1.4

Define $F(x) \equiv x + h'(x)\frac{(r_H - r_L) + \lambda f(\gamma m, m, h(x))\frac{(\gamma(1 - 2h(x)) + h(x))}{2h(x)(1 - h(x))}}{[cq(t/n)((n + 1)/n)]}$. Since $\gamma \leq \frac{p_L^{\min}}{1 - 2p_L^{\min}(1 - p_L^{\min})}$ by assumption, the function is concave and we have that $\frac{\partial F(x)}{\partial n} \leq 0$. Note that $[cq(t/n)((n + 1)/n)]$ is decreasing in n and $\frac{\partial f(\gamma m, m, h(x))}{\partial m} = -f(\gamma m, m, p)\frac{m(\gamma - p)^2 + p(1 - p)}{2mp(1 - p)} < 0$. This implies that $\frac{\partial F(x)}{\partial n} > 0$. By implicit function theorem, $\frac{\partial x_{RA}^{pv*}}{\partial n} = -\frac{(\partial F(x))/(\partial n)}{(\partial F(x))/(\partial x_{RA}^{pv*})} \geq 0$.

Proof of Theorem C.1.5

For part (i), it is easy to check that farms' payoff from adulteration is $r_H m - q(t/n)cm$ and non adulteration is $r_L m$. Comparing the two payoffs gives us the desired result in the theorem.

For part (ii), note that expected payoff for a farm from adulteration is $mr_H -$

$q(t/n) \frac{(m + m(n-1)p_L^{max} p_{ad})}{k} cm$ and from non adulteration is mr_L . For part (a), It is easy to check that if $q(t/n)c \leq \frac{(r_H - r_L)n}{p_L^{max}(n-1) + 1}$, then even if all other farms always adulterate (i.e., $p_{ad} = 1$), it is optimal for the given farm to also always adulterate. Similarly, for part (c), it is easy to check that if $(r_H - r_L)n \leq q(t/n)c$, then even if none of the other farms adulterate (i.e., $p_{ad} = 0$), it is optimal for the given farm to also not adulterate. For part (b), it is easy to find p_{ad} using the fact that the expected payoff from adulteration should be equal to that of non adulteration at equilibrium.

Proof of Proposition C.1.6

Checking that E_n is increasing in n is equivalent to checking that p_{ad} is increasing in n because $E_n = nmp_L^{max} p_{ad} = kp_L^{max} p_{ad}$. First, consider the case when $t = n$. Consider n_1 and n_2 such that $n_2 > n_1$. We will show that $p_{ad}^{n_2} > p_{ad}^{n_1}$ for all n_1 and n_2 . If $(r_H - r_L)n \leq qc$ then both $p_{ad}^{n_1} = p_{ad}^{n_2} = 0$ from theorem C.1.5. Next, if both n_1 and n_2 are such that $cq \in [\frac{(r_H - r_L)n_i}{p_L^{max}(n_i - 1) + 1}, (r_H - r_L)n_i]$ for $i = 1, 2$, then it is easy to check that $\frac{\partial p_{ad}}{\partial n} \geq 0$. Since $n_2 > n_1$, this implies that $p_{ad}^{n_2} \geq p_{ad}^{n_1}$. Finally, note that since $\frac{(r_H - r_L)n}{p_L^{max}(n-1) + 1}$ is increasing in n , either $qc \leq \frac{(r_H - r_L)n_i}{p_L^{max}(n_i - 1) + 1}$ for both $i = 1, 2$ or $\frac{(r_H - r_L)n_1}{p_L^{max}(n_1 - 1) + 1} \leq qc \leq \frac{(r_H - r_L)n_2}{p_L^{max}(n_2 - 1) + 1}$. In both the cases it is easy to check that $p_{ad}^{n_2} \geq p_{ad}^{n_1}$. The proof for partial traceability follows similarly and we avoid redoing it here for the sake of brevity.

Proof of Theorem C.1.7

Let x_{n_L} be the fraction of n_L low quality units adulterated at equilibrium and let $a_i(n_{L,i}) : \{1, \dots, m\}$ be farm i 's adulteration strategy. then revenue for a farm with n_L low quality units is given by $\pi_F^{RV} = (m - n_L + n_L x_{n_L})r_H + (n_L - x_{n_L})r_L - \gamma_i(n_{L,i}, a_{-i}(n_{L,-i}))cm$, where $\gamma_i(n_{L,i}, a_{-i}(n_{L,-i})) \equiv q\left(\frac{t}{n}\right) \binom{x_{n_L,i}}{m} \left(\frac{x_{n_L,i} + \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]}{k}\right)$. Note that $\frac{\partial \pi_F^{RV}}{\partial x_{n_L}} = n_L(r_H - r_L) - cq(t/n)(2n_L^2 x_{n_L} + n_L \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})])$ and $\frac{\partial \pi_F^{RV}}{\partial^2 x_{n_L}^2} = -cq(t/n)(2n_L^2) \leq 0$. Since the revenue function is concave, the optimal value of $x_{n_L}^*$ is either at the boundary points or at the FOC. Thus $x_{n_L}^* = \max\{0, \min\{1, \frac{1}{2n_L} \left(\frac{kn(r_H - r_L)}{cqt} - \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]\right)\}\}$. Further, note that the optimal value of $x_{n_L}^* n_L$ is a constant if the optimal solution is not at the boundary. Rewriting the condition, we get that if $2n_L \leq \left(\frac{kn(r_H - r_L)}{cqt} - \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]\right)$ then $x^*(n_L) = 1$ and otherwise it is $\frac{1}{2n_L} \left(\frac{kn(r_H - r_L)}{cqt} - \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]\right)$ as long as $\frac{kn(r_H - r_L)}{cqt} \geq \mathbb{E}_{n_{L,-i}}[\sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})]$. Since we are considering symmetric BNE, we can rewrite the

above FOC to determine the threshold β as $2\beta = \frac{nk(r_H - r_L)}{cqt} - (n-1) \left(\int_0^\beta x f(x, \frac{k}{n}, p_L) dx + \int_\beta^m \beta f(x, \frac{k}{n}, p_L) dx \right)$. Note that $\max_\beta \left\{ \int_0^\beta x f(x, \frac{k}{n}, p_L) dx + \int_\beta^m \beta f(x, \frac{k}{n}, p_L) dx \right\} = mp$ and it is attained when $\beta \geq m$. If $\beta = m$ is a symmetric BNE, then given that $\beta_{-i} = m$, we should get that $\beta_i = m$ as well. If $\frac{1}{2} \left(\frac{nk(r_H - r_L)}{cqt} - (n-1)mp \right) \geq m$ then this implies that $\beta_i = m$. This proves part (b) of the theorem. If $\frac{1}{2} \left(\frac{nk(r_H - r_L)}{cqt} - (n-1)mp \right) < m$ then we can use the intermediate value theorem and the fact that the fixed point equation is monotonic in β to show the existence and uniqueness of β as we did in Theorem 4.3.4.

Proof of Proposition C.1.8

Proceeding in the same way as in proposition ?? first for fully traceable systems, we get that proving $\frac{\partial E_m}{\partial m} \leq 0$ is equivalent to showing the following:

$$0 \leq \int_0^\beta f(x, m, p) dx \left(\int_0^\beta f(x, m, p) dx + \int_\beta^m \beta f(x, m, p) dx - mp \right) + \int_\beta^m \beta f(x, m, p) dx \quad (\text{C.16})$$

$$\Leftrightarrow 0 \leq \left(\int_0^\beta f(x, m, p) dx \right) \left(\int_\beta^m (\beta - x) f(x, m, p_L) dx \right) + \int_\beta^m \beta f(x, m, p) dx \quad (\text{C.17})$$

Differentiating the RHS w.r.t to β , we get:

$$-f(\beta, m, p) \left(\int_\beta^m (x - \beta) f(x, m, p_L) dx + \beta \right) + 1 - \left(\int_0^\beta f(x, m, p) dx \right)^2 \quad (\text{C.18})$$

Differentiating equation C.18, we get:

$$-f(\beta, m, p) \left(2 \int_0^\beta f(x, m, p) dx + \int_0^\beta f(x, m, p) dx \right) - \left(\int_\beta^m (x - \beta) f(x, m, p_L) dx \right) \quad (\text{C.19})$$

$$+ \beta) f(\beta, m, p) \frac{mp - \beta}{mp(1 - p)} \quad (\text{C.20})$$

We will again consider two cases:

Case 1: Assume $\beta \leq mp$: We wish to show that the minimum value of the RHS in equation C.17 for any $\beta \leq mp$ is greater than 0. Note that the first term in equation C.18 is always non positive. Further, since $\int_\beta^m (x - \beta) f(x, m, p_L) dx + \beta$ is increasing in β and it is non-negative at $\beta = 0$ it must always be non negative for $\beta \geq 0$. If $\beta \leq mp$ then $\frac{mp - \beta}{mp(1 - p)}$ is also non negative and the second term is also non positive. Thus since the double differential of the RHS is negative the function is concave in β . Since we are trying to find the minimum value of a

concave function, we only need to check the boundary points. It is easy to check that equation C.17 is non negative when evaluated at $\beta = 0$ and $\beta = mp$. Thus it must be non negative at all $\beta \in (0, mp)$.

Case 2: Assume $\beta > mp$: Replacing the value of $\int_0^\beta xf(x, m, p)dx = mp(1 - p)f(\beta, m, p) + mp \int_0^\beta f(x, m, p)dx$ in equation C.16, and after some algebraic simplifications, we get that $\frac{\partial E_m}{\partial m} \leq 0$ is equivalent to showing:

$$\begin{aligned} 0 \leq & (\beta - mp) \int_0^\beta f(x, m, p)dx \int_\beta^m f(x, m, p)dx + \beta \int_\beta^m f(x, m, p)dx \\ & + mp(1 - p)f(\beta, m, p)dx \int_0^\beta f(x, m, p)dx \end{aligned} \quad (\text{C.21})$$

Note that the second and the third term in equation C.21 are always non negative. If $\beta > mp$ then the first term is also non negative and thus RHS must be non negative when $\beta > mp$.

For partially traceable systems, since the $\frac{\partial E_m}{\partial m} \leq 0$ under full traceability, E_m must be decreasing in m in the partial traceability case as well. This is because partial traceability can only increase risk since traceability factor decreases as n increases.

Proof of Theorem C.1.9

At NE, farm i chooses x_i^* so as to maximize his expected payoff given the choices of all other farms. Note that $\frac{\partial \pi^{PV}(x_i, x_{-i})}{\partial x_i} = -mh'(x)(r_H - r_L) - q\frac{t}{n}cmS'(x_i) \left(\frac{\sum_{-i} x_{-i} + 2S(x_i)}{n} \right)$. It is easy to check that this is a concave function. Further, since we consider only symmetric equilibrium, we have $x_{-i}^* = x_i^* = x_{PV}^*$. Substituting this, and noting that x cannot be greater than 1 gives us the desired result.

Note that in the second case there are two regions of the S_2 function. We will compare the optimal solutions in both the regions and pick the one which is optimal. For the first region i.e. $x \leq \tau$, differentiating farmers' revenue w.r.t. x gives us a first order condition equivalent to one in theorem C.1.9. Since it is again a concave function, the optimal solution is obtained either at the first order condition or at the boundary. If $x > \tau$, revenue function increases linearly since the detection probability stays constant while the revenue from adulteration increases linearly. This implies that if the optimal solution in the first region is at the boundary, the revenue is always increasing and $x^{PV^*} = 1$ is optimal. This proves the first part. Otherwise we have two points to compare: x^* of the first order condition and $x^* = 1$. Using the symmetry condition, we get that if $r_H m - cmq(t/n) > mr_H(1 - h(x^*) + mr_L h(x^*) - cmq(t/n)S_2(x^*)^2)$, $x^{PV^*} = 1$ is optimal, otherwise $x^{PV^*} = 1$ is optimal. This proves the second part of the equation.

Because of the piecewise linear nature of the testing sensitivity in S_3 , we need to consider four scenarios depending on where the adulterants in mixed and individual supply lie respectively. For example, if $\sum x_j/n \leq 1/\alpha$ then detection probability in the aggregated supply is $\alpha \sum x_j/n$ and it is 1 otherwise. Similarly, for farmer i , if $x_i \leq 1/\alpha$ then detection probability for individual sample is $x_i\alpha$ and it is 1 otherwise. Since we are considering symmetric equilibrium, $x_j = x^*$ for all j at equilibrium. Thus, either both $\sum x^*/n$ and x^* are less than $1/\alpha$ or they are both greater than $1/\alpha$. The revenue for farmer i if $x_i \leq 1/\alpha$ is given by $\pi_i^{ip} = mr_H - mh(x)(r_H - r_L) - cmq(t/n) \frac{\sum_{-i} x_{-i} + x_i}{n} x_i \alpha^2$. Similarly if $x_i > 1/\alpha$, then $\pi_i^p = mr_H - mh(x)(r_H - r_L) - cmq(t/n)$. It is easy to check that π_i^{ip} is concave in x_i . Using the fact that $x_i = x_{-i}$ for a symmetric equilibrium, we get that the optimal solution is either at the first order condition or at the boundary points. Checking that the value of $\frac{\partial \pi_i^{ip}}{\partial x_i}$ at $1/\alpha$ is positive simplifies to $c < -h'(1/\alpha)(r_H - r_L)/[\alpha q(t/n)((n+1)/n)]$. Next, note that π_i^p is increasing in x thus $x^* = 1$ is a local optimal in this region. Thus while the revenue function is concave in the first region it is monotonically increasing in the second region. Comparing the optimal solution in the first and the second region gives us the desired result in the theorem.

Proof of Theorem C.1.10

We prove the first part when detection probability is modeled as S_1 . It is optimal for a farm to adulterate only if revenue with adulteration is greater than revenue without adulteration i.e.

$$mr_H - q \left(\frac{t}{n} \right) S_1 \left(\frac{n_L}{m} \right) \mathbb{E}_{n_L, -i} \left[S_1 \left(\frac{n_L + \sum_{-i} n_{L, -i} a_{-i}(n_{L, -i})}{k} \right) \right] cm \geq mr_H - n_L(r_H - r_L) \quad (\text{C.22})$$

$$\Leftrightarrow \frac{n(r_H - r_L)}{qtc m} \geq \frac{1}{n_L} S_1 \left(\frac{n_L}{m} \right) \mathbb{E}_{n_L, -i} \left[S_1 \left(\frac{n_L + \sum_{-i} n_{L, -i} a_{-i}(n_{L, -i})}{k} \right) \right] \quad (\text{C.23})$$

Notice that the LHS in equation C.23 does not depend on n_L . If we show that the RHS is increasing in n_L , then the LHS and RHS can only be equal at a unique point. Next, we will show that the RHS is indeed increasing in n_L .

$$\frac{\partial \left(\frac{1}{n_L} S_1 \left(\frac{n_L}{m} \right) \right)}{\partial n_L} \geq 0 \Leftrightarrow \frac{n_L}{m} S_1' \left(\frac{n_L}{m} \right) - S \left(\frac{n_L}{m} \right) \geq 0 \quad (\text{C.24})$$

$$\frac{\partial \left(n_L m S_1' \left(\frac{n_L}{m} \right) - S_1' \left(\frac{n_L}{m} \right) \right)}{\partial n_L} = \frac{n_L}{m^2} S_1'' \left(\frac{n_L}{m} \right) \geq 0 \quad (\text{C.25})$$

The last inequality in equation C.25 follows from the convexity of S . Since equation C.25 is

always non negative, it implies that the LHS in C.24 is increasing in n_L and we only need to evaluate the smallest n_L to check that it is always non negative. Since $S(0) = 0$ and S is an increasing function, LHS in equation C.24 evaluated at $n_L = 0$ is indeed non negative. Thus we have that $\frac{1}{n_L}S\left(\frac{n_L}{m}\right)$ is increasing in n_L . Finally note that since $S(\cdot)$ is an increasing function and $\frac{n_{L,i} + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k}$ increases as n_L increases, $\mathbb{E}_{n_{L,-i}} \left[S\left(\frac{n_L + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k}\right) \right]$ is also increasing in n_L . Since both the terms in the RHS of equation C.23 are increasing in n_L , their product must also increase as n_L increases. Thus, since there is a unique point at which the LHS is equal to RHS, and the RHS is increasing in n_L , farm's adulteration strategy is indeed a threshold strategy with farms adulterating only when there low quality units are less than a threshold. We characterize the threshold under a symmetric BNE. If $\beta^S = m$, under a symmetric BNE, then given $\beta_{-i}^S = m$, $\beta^S = m$ should be optimal. This implies equation C.23 evaluated at $\beta^S = m$ should hold true i.e. $\frac{n(r_H - r_L)}{qtc m} \geq \frac{1}{m} \mathbb{E}_{n_{L,-i}} \left[S\left(\frac{m + \sum_{-i} n_{L,-i}}{k}\right) \right]$. Simplifying this gives us the desired result in part (b) of the theorem. If $\frac{n(r_H - r_L)}{qtc m} < \frac{1}{m} \mathbb{E}_{n_{L,-i}} \left[S\left(\frac{m + \sum_{-i} n_{L,-i}}{k}\right) \right]$, then using the fact that all farms have the same threshold at which condition in C.23 is satisfied at equality under the symmetric case, we can rewrite the condition in C.23 as follows: $\frac{n(r_H - r_L)}{qtc m} = \frac{1}{\beta^S} S\left(\frac{\beta^S}{m}\right) \mathbb{E}_{n_{L,-i}} \left[S\left(\frac{\beta^S + \sum_{-i} n_{L,-i} \mathbb{I}\{n_{L,-i} < \beta^S\}}{k}\right) \right] cm$. We will use the Intermediate Value Theorem to show the existence of β^S . Define $F(\beta) \equiv \frac{1}{\beta^S} S\left(\frac{\beta^S}{m}\right) \mathbb{E}_{n_{L,-i}} \left[S\left(\frac{\beta^S + \sum_{-i} n_{L,-i} \mathbb{I}\{n_{L,-i} < \beta^S\}}{k}\right) \right] cm - \frac{n(r_H - r_L)}{qtc m}$ for $\beta \in (0, m)$. Note that $F(0) = -\frac{n(r_H - r_L)}{qtc m} < 0$ and $F(m) = \frac{1}{m} \mathbb{E}_{n_{L,-i}} \left[S\left(\frac{m + \sum_{-i} n_{L,-i}}{k}\right) \right] - \frac{n(r_H - r_L)}{qtc m} > 0$ from assumption. Since $F(\cdot)$ is continuous and $F(0) < 0 < F(m)$, there exists $\beta^{RV} \in (0, m)$ s.t. the condition is satisfied at equality. Finally, since the RHS in the equation is monotonic in β^S , the fixed point is also unique.

Next, we characterize the optimal solution in the case when detection probability is modeled by S_2 . It is optimal for a farmer to adulterate only if revenue with adulteration is greater than revenue without adulteration i.e.

$$mr_H - q \left(\frac{t}{n}\right) S_2\left(\frac{n_L}{m}\right) \mathbb{E}_{n_{L,-i}} \left[S_2\left(\frac{n_L + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k}\right) \right] cm \geq mr_H - n_L(r_H - r_L) \quad (C.26)$$

$$\Leftrightarrow n_L \frac{n(r_H - r_L)}{qtc m} \geq S_2\left(\frac{n_L}{m}\right) \mathbb{E}_{n_{L,-i}} \left[S_2\left(\frac{n_L + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k}\right) \right] \quad (C.27)$$

We characterize the threshold under a symmetric BNE. If $\beta^S = m$, under a symmetric BNE,

then given $\beta_{-i}^S = m$, $\beta^S = m$ should be optimal. This implies equation C.27 evaluated at $\beta^S = m$ should hold true i.e. $\frac{n(r_H - r_L)}{qtc m} \geq \frac{1}{m} \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{m + \sum_{-i} n_{L,-i}}{k} \right) \right]$. Simplifying this gives us the desired result in part (b) of the theorem.

If $\frac{n(r_H - r_L)}{qtc m} < \frac{1}{m} \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{m + \sum_{-i} n_{L,-i}}{k} \right) \right]$: Using the fact that RHS is convex increasing initially and a constant finally and LHS is linearly increasing, we have that the LHS in equation C.27 can be equal to RHS only at most two points (the first point is when LHS is convex increasing and the second point is when LHS is a constant). We will show that both β^l and β^u cannot be less than τm using contradiction. Assume it is indeed the case β^l and β^u are less than τm . Using equation C.27, this implies that $\frac{n(r_H - r_L)}{qtc m} = \frac{S_2 \left(\frac{\beta^l}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{\beta^l + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right) \right]}{\beta^l} = \frac{S_2 \left(\frac{\beta^u}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{\beta^u + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right) \right]}{\beta^u}$.

For all $n_L \in (\beta^l, \tau m)$, note that $S_2 \left(\frac{n_L}{m} \right)$ is convex increasing and $\mathbb{E}_{n_{L,-i}}[\cdot]$ is also increasing. This implies that $\frac{S_2 \left(\frac{n_L}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{n^l + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right) \right]}{n^l}$ is increasing for all $n_L \in (\beta^l, \tau m)$ and $\frac{S_2 \left(\frac{\beta^l}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{\beta^l + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right) \right]}{\beta^l} < \frac{S_2 \left(\frac{\beta^u}{m} \right) \mathbb{E}_{n_{L,-i}} \left[S_2 \left(\frac{\beta^u + \sum_{-i} n_{L,-i} a_{-i}(n_{L,-i})}{k} \right) \right]}{\beta^u}$.

This is a contradiction and thus we have shown that both β^l and β^u cannot be less than τm simultaneously.

Finally, we consider the case when detection probability is modeled as S_3 . We will first show that when $cq \leq \max \left(\frac{r_H - r_L}{\alpha}, \frac{k(r_H - r_L)}{\alpha(p\alpha(k - m) + m)} \right)$ then always adulterating is indeed an equilibrium. Let us first consider the case when $\frac{r_H - r_L}{\alpha} < \frac{k(r_H - r_L)}{\alpha(p\alpha(k - m) + m)}$. This implies $\frac{m}{\alpha} \leq \frac{k}{\alpha} - \left(\frac{k}{m} - 1 \right) mp$ because $\alpha p \leq 1$. If $n_L \leq \frac{m}{\alpha}$, then the probability of detection in both mixed (p_m) (in expectation) and individual supply (p_i) is less than 1. This is because $p_i = \frac{n_L \alpha}{m} \leq 1$

and $p_m = \frac{\alpha(n_L + \left(\frac{k}{m} - 1 \right) mp)}{k} \leq 1$ for all $n_L \leq \frac{m}{\alpha}$. It is easy to check that revenue from adulteration is more than revenue from non-adulteration when $n_L = \frac{m}{\alpha}$. Further, note that the adulteration strategy when $n_L \leq \frac{m}{\alpha}$ is to adulterate below a threshold as already proved in theorem 4.3.4. Thus, if it is optimal to adulterate when $n_L = \frac{m}{\alpha}$ then it should indeed be optimal to adulterate for all $n_L \leq \frac{m}{\alpha}$. Next, if $\frac{m}{\alpha} < n_L < \frac{k}{\alpha} - \left(\frac{k}{m} - 1 \right) mp$, then while $p_i = 1$,

$p_m = \frac{\alpha(n_L + \left(\frac{k}{m} - 1 \right) mp)}{k}$. It is easy to check that in this range, it is optimal to adulterate only when the number of low quality units are greater than threshold. However, since we have already shown that it is optimal to adulterate at $n_L = \frac{m}{\alpha}$, then it should indeed be optimal to adulterate

for all $\frac{m}{\alpha} < n_L \leq \frac{k}{\alpha} - (\frac{k}{m} - 1)mp$. Finally, note that if $n_L > \frac{k}{\alpha} - (\frac{k}{m} - 1)mp$, then $p_i = 1$ and $p_m = 1$. This is similar to the perfect testing case we have already analyzed and the adulteration strategy in perfect testing case is to adulterate above a threshold. Since adulteration is indeed optimal at $\frac{k}{\alpha} - (\frac{k}{m} - 1)mp$, it should be optimal for all $n_L \geq \frac{k}{\alpha} - (\frac{k}{m} - 1)mp$. Thus we have shown that it is indeed optimal for this farm to always adulterate when all other farms are adulterating. The proof for the other case when $\frac{r_H - r_L}{\alpha} > \frac{k(r_H - r_L)}{\alpha(p\alpha(k - m) + m)}$ follows similarly. Next, note that if $cq \in \left(\max\left(\frac{r_H - r_L}{\alpha}, \frac{k(r_H - r_L)}{\alpha(p\alpha(k - m) + m)}\right), \max\left(\frac{k(r_H - r_L)}{m\alpha}, r_H - r_L\right) \right)$, then as in the previous case, it is optimal to adulterate till a threshold (β^S) when $n_L \leq tr \equiv \min\left(\frac{k}{\alpha} - \mathbb{E}_{n_L, -i} [\sum_{-i} n_{L, -i} a_{-i}(n_{L, -i})], \frac{m}{\alpha}\right)$ as the detection probability in this range is similar to the imperfect testing case we have already analyzed. It is again easy to check that if $n_L > tr$, it is optimal to adulterate above a threshold (β^U). Further, note that β^U has to be smaller than $\frac{cqm(t/n)}{r_H - r_L}$ because we already know from theorem 4.3.3 that even under perfect testing, which has the highest risk of being penalized, farms still adulterate when $n_L \geq \frac{cqm(t/n)}{r_H - r_L}$. The two conditions imply that in any symmetric BNE, equilibrium adulteration strategy is to adulterate below a threshold (β^S) and then adulterate above a threshold (β^U). The exact values of the thresholds in the theorem follow from the symmetric BNE assumption and the constraints that $\beta^S \geq 0$ and $\beta^U \leq \min\left(m, \frac{cqm(t/n)}{r_H - r_L}\right)$.

Proof of Theorem C.1.11

First, we prove the result for part (i). Farm i chooses x_i^* so as to maximize his expected payoff $\pi^{PP} = mr_H h(x) + mr_L(1 - h(x)) - cm^2 x q(t/n)$. Let us assume that a subset n_a out of a total of n farms are adulterating. Consider a farm i that is adulterating at equilibrium. Let x^* be the optimal amount of adulterants for farms that are adulterating. Then, revenue from adulteration (π_a^i) should be strictly greater than revenue from non adulteration (π_{na}^i). $\pi_a^i = r_H m - mh(x^*)(r_H - r_L) - q \frac{n_a}{n} (t/n) cm^2 x^*$ and $\pi_{na}^i = r_H m - mh(0)(r_H - r_L)$. This is equivalent to checking that $\frac{\partial \pi_a^i}{\partial x}$ at $x = 0$ is greater than 0. Now consider a farmer i' who is not adulterating at equilibrium. Then, $\pi_{na}^{i'} \geq \pi_a^{i'}$. $\pi_a^{i'} = r_H m - mh(x^*)(r_H - r_L) - q \frac{n_a + 1}{n} (t/n) cm^2 x^*$ and $\pi_{na}^{i'} = r_H m - mh(0)(r_H - r_L)$. This is equivalent to checking that $\frac{\partial \pi_{na}^{i'}}{\partial x}$ at $x = 0$ is less than 0. Simplifying the two conditions we get, $\frac{nh'(0)(r_H - r_L)}{cmq(t/n)} - 1 \leq n_a < \frac{nh'(0)(r_H - r_L)}{cmq(t/n)}$. Next, we also know that $n_a^* \leq n$. Thus we need to have $\frac{-h'(0)(r_H - r_L)n^2}{mcqt} - 1 < n$. Simplifying this condition gives us the condition on c for part (a). Condition for part (b) follows similarly from the condition on n_a^* . Finally, in both the cases optimal x^{PP*} can be calculated using the fact

that $\pi_a = mr_H - mh(x)(r_H - r_L) - cm^2q(t/n)x$ is concave in x , and the optimal x^* is either the solution to the first order condition or at the boundary point.

The proof for the second part follows exactly in the same way as in theorem 4.3.2 with cm replaced with cm^2x . At an NE, farm i chooses x_i^* so as to maximize his expected payoff in Equation (4.2) given the choices of all other farms. Note that $\frac{\partial \pi^{PV}(x_i, x_{-i})}{\partial x_i} = -mh'(x)(r_H - r_L) - q\frac{t}{n}cm\left(\frac{2x_i \sum_{-i} x_{-i} + 3x_i^2}{n}\right)$ and $\frac{\partial^2 \pi^{PV}(x_i, x_{-i})}{\partial x_i^2} = -mh''(x)(r_H - r_L) - q\frac{t}{n}cm\left(\frac{2 \sum_{-i} x_{-i} + 6x_i}{n}\right) < 0$. Further, since we consider only symmetric equilibrium, we have $x_{-i}^* = x_i^* = x_{PV}^*$. Substituting this, we get that $\pi^{PV}(x_i, x_{-i})$ is a concave function and the optimal solution is achieved either at the first order condition or at the boundary.

For the third part, farms adulterate if $r_H m - q(t/n)cmn_L > r_H m - n_L(r_H - r_L)$. This simplifies to $r_H - r_L > q(t/n)cm$.

For the last part, the proof for showing that any optimal strategy in this game is a threshold policy follows exactly as in Theorem 4.3.4 and we omit it here for brevity. Using the fact that all farms have the same threshold β , we have that $\mathbb{E}_{n_L, -i} [\sum_{-i} n_{L, -i} a_{-i}(n_{L, -i})] = (n - 1) \int_0^{\beta^{RV}} xf(x, m, p_L)dx$. Note that $\max_{\beta} \left\{ \int_0^{\beta} xf(x, m, p_L)dx \right\} = mp_L$ and it is attained when $\beta \geq m$. If $\beta = m$ is a symmetric BNE, then given that $\beta_{-i} = m$, we should get that $\beta_i = m$ as well. This happens when $\pi(a, \beta_{-i} = m) > \pi(na, \beta_{-i} = m)$, because it implies that, given that all other farms have $\beta_{-i} = m$, $\beta_i = m$ is optimal for farm i . This condition simplifies to $cq \left(\frac{t}{n}\right) < \frac{k(r_H - r_L)}{m(m + (n - 1)mp)}$. This proves part (b). If $cq \left(\frac{t}{n}\right) \geq \frac{k(r_H - r_L)}{m(m + (n - 1)mp)}$, it is again easy to show the existence of β^{RV} using the Intermediate Value Theorem and the fact that the threshold should be the same for all farms in a symmetric equilibrium as in Theorem 4.3.4 .

Proof of Proposition C.1.12

It is easy to check that all the results for both reactive and preemptive adulteration follow exactly as in the above proofs with cm replaced by $cf(m)$ where $f(m)$ is a convex increasing function. We omit the proofs here for brevity.

Proof of Theorem C.2.1

We will first write the risk constraint in terms of qt for all the problems. For instance, the risk constraint in the original problem for reactive scenarios under perfect testing can be written as $P_n \equiv \int_0^{\frac{\beta^{RP} - mp}{\sqrt{mp(1-p)}}} f(x, 0, 1)dx \leq \alpha$. This constraint can be equivalently written as $qt \geq u^{RP} \equiv \left(\frac{n^2(r_H - r_L)}{ck}\right) \left(\frac{pk}{n} + \phi^{-1}(1 - \alpha)\sqrt{\frac{kp(1-p)}{n}}\right)$ using the fact that $\beta^{RP} = \frac{cqt m^2}{k(r_H - r_L)}$. In all the cases, we can similarly rewrite the risk constraint in terms of qt . Since $q \leq 1$ and $t \leq n$, qt cannot be

greater than n . Thus if $u^j > n$ in any of the cases, then the problem becomes infeasible. This proves the first part of the proposition. Now if $u^j \leq n$, then note that since cost is increasing in both q and t and P_m decreases as q or t increases, the risk constraint will be strict in all the cases and $qt = u^j$ at optimality. We can thus replace t in the problem as $\frac{u^j}{q}$ and rewrite the problem as

$$\Pi^j(q, t) \equiv \min_q \left\{ l(q) + g\left(\frac{u^j}{q}\right) \mid q \in [0, \frac{u_j}{n}], \right\}, \quad (\text{C.28})$$

Note that $\frac{\partial(l(q)+g(\frac{u^j}{q}))}{\partial q} = l'(q) - \frac{u^j}{q^2}g'(\frac{u^j}{q})$ and $\frac{\partial^2(l(q)+g(\frac{u^j}{q}))}{\partial q^2} = l''(q) + \frac{u_j^2}{q^4}g''(\frac{u_j}{q}) + 2\frac{u_j}{q^3}g'(\frac{u_j}{q}) > 0$ since l and g are convex increasing functions. This problem is just a constrained convex optimization problem. Thus the optimal solution is either at the point where $\frac{\partial(l(q)+g(\frac{u^j}{q}))}{\partial q} = l'(q) - \frac{u^j}{q^2}g'(\frac{u^j}{q}) = 0$ or at the boundary points i.e. $q = 1$ or $q = \frac{mu_1}{k}$. There are three cases:

- (i) Case 1: If $l'(1) - u^j g'(u^j) < 0$ then setting q to 1 is optimal since cost is always decreasing in $q \in [\frac{u^j}{n}, 1]$. Thus $(q^*, t^*) = (1, u_1)$
- (ii) Case 2: If $g'(n) \leq \frac{u^j}{n^2}l'(\frac{u^j}{n})$, then setting q to be minimum is optimal since cost is always increasing in $q \in [\frac{u^j}{n}, 1]$. Thus $(q^*, t^*) = (\frac{u^j}{n}, n)$
- (iii) Case 3: If $l'(1) > g'(u^j)$ and $g'(n) > \frac{u^j}{n^2}l'(\frac{u^j}{n})$ there exists q where $f'(q) - \frac{u_1}{q^2}g'(\frac{u_1}{q}) = 0$ and that q is optimal. Simplifying, we get $q^* = \sqrt{\frac{u^j g'(u^j/q^*)}{l'(q^*)}}$ and $t^* = \frac{u^j}{q^*}$.

Proof of Proposition C.2.2

From Theorem C.2.1, we know that if $\frac{u^j}{n} > 1$, then the manufacturer's problem with n suppliers is infeasible. It is easy to check that $\frac{u^j}{n}$ is increasing in n for all cases and we present the proof only for $\frac{u^{RV}}{n}$ here for the sake of brevity. Define $d = cmp + c\phi^{-1}(\alpha)\sqrt{mp(1-p)} + (\frac{k}{m} - 1)c \int_0^\beta xf(x, m, p_L)dx$. Then $\frac{u^{RV}}{n} = \frac{k(r_H - r_L)}{d}$ using the fact that $m = \frac{k}{n}$. Define $h = p + \phi^{-1}(\alpha)\frac{\sqrt{p(1-p)}}{2\sqrt{m}}$ We will show that $\frac{\partial d}{\partial m} \geq 0$.

$$\begin{aligned} \frac{\partial d}{\partial m} &= p + \phi^{-1}(\alpha)\frac{\sqrt{p(1-p)}}{2\sqrt{m}} - \frac{k}{m^2} \int_0^\beta xf(x, m, p_L)dx + \left(\frac{k}{m} - 1\right)\frac{\partial}{\partial m} \left(\int_0^\beta xf(x, m, p_L)dx \right) \\ &= h - \frac{k}{m^2} \int_0^\beta xf(x, m, p_L)dx + \left(\frac{k}{m} - 1\right) \left(\beta \frac{\partial}{\partial m} \left(\int_0^\beta f(x, m, p_L)dx \right) + \int_0^\beta \frac{(t+mp)f(t, m, p)}{2m} dt \right) \\ &= h - \frac{k}{m^2} \int_0^\beta xf(x, m, p_L)dx + \left(\frac{k}{m} - 1\right) \left(\beta \frac{\partial \alpha}{\partial m} + \int_0^\beta \frac{(t+mp)f(t, m, p)}{2m} dt \right) \\ &= h - \frac{k}{m^2} \int_0^\beta xf(x, m, p_L)dx + \left(\frac{k}{m} - 1\right) \left(\int_0^\beta \frac{(t+mp)f(t, m, p)}{2m} dt \right) \end{aligned}$$

$$\begin{aligned}
 &= h + \frac{k}{2m^2} \int_0^\beta (mp - x)f(x, m, p_L)dx - \int_0^\beta \frac{(t + mp)f(t, m, p)}{2m} dt \\
 &\geq h - \frac{\beta + mp}{2m} + \frac{k}{2m^2} \int_0^\beta (mp - x)f(x, m, p_L)dx \\
 &= h - \frac{\phi^{-1}(\alpha)\sqrt{mp(1-p)} + 2mp}{2m} + \frac{k}{2m^2} \int_0^\beta (mp - x)f(x, m, p_L)dx \\
 &= p + \phi^{-1}(\alpha)\frac{\sqrt{p(1-p)}}{2\sqrt{m}} - \frac{\phi^{-1}(\alpha)\sqrt{mp(1-p)} + 2mp}{2m} + \frac{k}{2m^2} \int_0^\beta (mp - x)f(x, m, p_L)dx \\
 &= \frac{k}{2m^2} \int_0^\beta (mp - x)f(x, m, p_L)dx \\
 &\geq 0
 \end{aligned}$$

The last inequality follows because $\phi^{-1}(\alpha) < 0$ for $\alpha < 0.5$ and thus $\beta \leq mp$. Finally, since d is increasing in m , d is decreasing in n .

Now let us assume that the manufacturer's problem is infeasible for SC_L , then $\frac{u_{n_L}^j}{n_L} > 1$ by Theorem C.2.1. But since $n_H > n_L$ and $\frac{u_n^j}{n}$ is increasing in n , we have that $\frac{u_{n_H}^j}{n_H} \geq \frac{u_{n_L}^j}{n_L} > 1$. This proves the first part of the proposition. Similarly if we assume that the manufacturer's problem is feasible for SC_H , then $\frac{u_{n_H}^j}{n_H} < 1$ by Theorem C.2.1 and we have that $\frac{u_{n_L}^j}{n_L} \leq \frac{u_{n_H}^j}{n_H} < 1$. This proves the second part of the proposition.

Next, we prove the third part of the proposition. From Theorem C.2.1 we also know that for any value of u_n^j there are three possibilities for optimal solutions. We will consider n_H and n_L and compare optimal solutions for all the nine possible cases. Let u_{n_H} and u_{n_L} be the value of the constant u_n^j for n_H and n_L respectively. Note that $u_{n_L} < u_{n_H}$ since $\frac{\partial u_n^j}{\partial n} \leq 0$. We will show that in all the cases $q_{n_L}^* \leq q_{n_H}^*$ and $t_{n_L}^* \leq t_{n_H}^*$.

Let's first assume that $(q_{n_H}^*, t_{n_H}^*) = (1, u_{n_H})$. There are three possible solutions for n_L .

- (i) Case 1: $(q_{n_L}^*, t_{n_L}^*) = (1, u_{n_L})$ then since $u_{n_L} < u_{n_H}$ we have $q_{n_L}^* = q_{n_H}^*$ and $t_{n_L}^* < t_{n_H}^*$.
- (ii) Case 2: $(q_{n_L}^*, t_{n_L}^*) = (\frac{u_{n_L}}{n_L}, n_L)$ If $u_{n_H} \geq n_L$ then we are done since $q_{n_L}^* \leq q_{n_H}^*$ and $t_{n_L}^* \leq t_{n_H}^*$. We will show that $u_{n_H} < n_L$ is not possible by contradiction. Assume that $u_{n_H} < n_L$. If $(q_{n_L}^*, t_{n_L}^*) = (\frac{u_{n_L}}{n_L}, n_L)$ then this implies $l'(\frac{u_{n_L}}{n_L}) - \frac{n_L^2}{u_{n_L}}g'(n_L) > 0$ from Theorem C.2.1. Also since l is convex and $\frac{u_{n_L}}{n_L} \leq 1$ from feasibility conditions, we have $l'(1) \geq l'(\frac{u_{n_L}}{n_L})$. Similarly since g is also convex, $u_{n_H} < n_L$ and $1 \leq \frac{n_L}{u_{n_L}}$, we have $\frac{n_L^2}{u_{n_L}}g'(n_L) > u_{n_H}g'(u_{n_H})$. This implies $l'(1) - u_{n_H}g'(u_{n_H}) > l'(\frac{u_{n_L}}{n_L}) - \frac{n_L^2}{u_{n_L}}g'(n_L) > 0$. This is a contradiction because we have $l'(1) - u_{n_H}g'(u_{n_H}) < 0$ from the assumption that $(q_m^*, t_m^*) = (1, u_{n_H})$.

- (iii) Case 3: $(q_{n_L}^*, t_{n_L}^*) = (\sqrt{\frac{u_{n_L} g'(u_{n_L}/q^*)}{l'(q^*)}}, \frac{u_{n_L}}{q^*})$. We have that $q_{n_L}^* \leq q_{n_H}^* = 1$. We need to check if $t_{n_L}^* = \frac{u_{n_L}}{q^*} \leq u_{n_H}$. Let $h = \frac{u_{n_L}}{u_{n_H}}$. Note that if we show $l'(h) - \frac{u_{n_L}}{h^2} g'(\frac{u_{n_L}}{h}) \leq 0$ then $q_{n_L}^* \geq h$ because we are minimizing a convex function whose derivative is less than 0 at h . We have $l'(h) - \frac{u_{n_L}}{h^2} g'(\frac{u_{n_L}}{h}) \leq 0$ because $l'(h) - \frac{u_{n_L}}{h^2} g'(\frac{u_{n_L}}{h}) = l'(\frac{u_{n_L}}{u_{n_H}}) - \frac{u_{n_H}^2}{u_{n_L}} g'(u_{n_H}) < l'(1) - u_{n_H} g'(u_{n_H}) < 0$. The first inequality follows because $h < 1$ and the second inequality follows from the assumption that $(q_m^*, t_m^*) = (1, u_{n_H})$.

Now let's assume that $(q_{n_H}^*, t_{n_H}^*) = (\frac{u_{n_H}}{n_H}, n_H)$. We will consider all the three possible optimal solutions for n_L .

- (i) Case 1: Let's assume $(q_{n_L}^*, t_{n_L}^*) = (1, u_{n_L})$, then $l'(1) - u_{n_L} g'(u_{n_L}) < 0 < l'(\frac{u_{n_H}}{n_H}) - \frac{n_H^2}{u_{n_H}} g'(n_H)$. Also note that $u_{n_H} \leq n_H$ and $u_{n_L} \leq n_L < n_H$ from feasibility conditions. This implies $l'(1) \geq l'(\frac{u_{n_H}}{n_H})$ and $u_{n_L} g'(u_{n_L}) < \frac{n_H^2}{u_{n_H}} g'(n_H)$. Thus $l'(1) - u_{n_L} g'(u_{n_L}) > l'(\frac{u_{n_H}}{n_H}) - \frac{n_H^2}{u_{n_H}} g'(n_H)$ which is a contradiction.
- (ii) Case 2: $(q_{n_L}^*, t_{n_L}^*) = (\frac{m' u_{n_L}}{k}, n_L)$. We have already shown that $\frac{\partial u_{n_L}^j}{\partial n} \geq 0$ and this implies $\frac{u_{n_L}}{n_L} \leq \frac{u_{n_H}}{n_H}$. We also have $n_L < n_H$ by construction.
- (iii) Case 3: $(q_{n_L}^*, t_{n_L}^*) = (\sqrt{\frac{u_{n_L} g'(u_{n_L}/q^*)}{l'(q^*)}}, \frac{u_{n_L}}{q^*})$. We need to show $q_{n_L}^* \leq \frac{u_{n_H}}{n_H}$ and $t_{n_L}^* \leq n_H$. We already have $t_{n_L}^* \leq n_L < n_H$. We also have $l'(\frac{u_{n_H}}{n_H}) - \frac{n_H^2}{u_{n_H}} g'(n_H) > 0$ because $(q_m^*, t_m^*) = (\frac{u_{n_H}}{n_H}, n_H)$. Let $h = \frac{u_{n_H}}{n_H}$. Then $h \leq 1$ because the problem is feasible for m . Since $n_H \geq 1$ and $u_{n_L} < u_{n_H}$ we have that $\frac{u_{n_L}}{h^2} g'(\frac{u_{n_L}}{h}) \leq \frac{n_H^2}{u_{n_H}} g'(n_H)$. This implies $l'(h) - \frac{u_{n_L}}{h^2} g'(\frac{u_{n_L}}{h}) > l'(\frac{u_{n_H}}{n_H}) - \frac{n_H^2}{u_{n_H}} g'(n_H) > 0$ and thus $q_{n_L}^* < \frac{u_{n_H}}{n_H}$.

Now let's assume that $(q_{n_H}^*, t_{n_H}^*) = (\sqrt{\frac{u_{n_H} g'(u_{n_H}/q^*)}{l'(q^*)}}, \frac{u_{n_H}}{q^*})$. We will consider all three possible optimal solutions for m' .

- (i) Case 1: Let's assume $(q_{n_L}^*, t_{n_L}^*) = (1, u_{n_L})$ then $l'(1) - u_{n_L} g'(u_{n_L}) < 0$. Further, we also have that $l'(q_{n_H}^*) - \frac{u_{n_H}}{q_{n_H}^*} g'(\frac{u_{n_H}}{q_{n_H}^*}) = 0$, $q_{n_H}^* \leq 1$ and $u_{n_L} < u_{n_H}$. This implies $l'(1) \geq l'(q_{n_H}^*)$ and $u_{n_L} g'(u_{n_L}) < \frac{u_{n_H}}{q_{n_H}^*} g'(\frac{u_{n_H}}{q_{n_H}^*})$. Combining the two we get that $l'(q_{n_H}^*) - \frac{u_{n_H}}{q_{n_H}^*} g'(\frac{u_{n_H}}{q_{n_H}^*}) < l'(1) - u_{n_L} g'(u_{n_L}) < 0$. This is a contradiction because we have that $l'(q_{n_H}^*) - \frac{u_{n_H}}{q_{n_H}^*} g'(\frac{u_{n_H}}{q_{n_H}^*}) = 0$. Thus this case is not possible.
- (ii) Case 2: Let's assume $(q_{n_L}^*, t_{n_L}^*) = (\frac{u_{n_L}}{n_L}, n_L)$. We will prove that $q_{n_H}^* \geq \frac{u_{n_L}}{n_L}$ and $\frac{u_{n_H}}{q^*} \geq n_L$. We have already shown that $\frac{u_{n_L}}{n_L} \leq \frac{u_{n_H}}{n_H}$. By feasibility, $q_{n_H}^* \geq \frac{u_{n_H}}{n_H} \geq \frac{u_{n_L}}{n_L}$. Further note that since $u_{n_H} > u_{n_L}$ we have that $l'(\frac{u_{n_H}}{n_L}) > l'(\frac{u_{n_L}}{n_L})$ and $\frac{n_L^2}{u_{n_H}} g'(n_L) < \frac{n_L^2}{u_{n_L}} g'(n_L)$.

This implies that $l'(\frac{u_{nH}}{n_L}) - \frac{n_L^2}{u_{nH}}g'(n_L) > l'(\frac{u_{nL}}{n_L}) - \frac{n_L^2}{u_{nL}}g'(n_L) > 0$. The last inequality follows because we have assumed that $(q_{nL}^*, t_{nL}^*) = (\frac{u_{nL}}{n_L}, n_L)$. Thus we have shown that $q_{nH}^* < \frac{u_{nH}}{n_L}$. This implies that $t_{nH}^* = \frac{u_{nH}}{q^*} \geq n_L$ as well.

(iii) Case 3: $(q_{nL}^*, t_{nL}^*) = (\sqrt{\frac{u_{nL}g'(u_{nL}/q^*)}{l'(q^*)}}, \frac{u_{nL}}{q^*})$. Using implicit function theorem,

$$\frac{\partial q}{\partial m} = \frac{\frac{1}{q^2} \frac{\partial u^j}{\partial m} [\frac{u^j}{q} g''(\frac{u^j}{q}) + g'(\frac{u^j}{q})]}{l''(q) + 2\frac{u^j}{q^3} g'(\frac{u^j}{q}) + \frac{u^j^2}{q^4} g'(\frac{u^j}{q})}$$

Since l and g are convex functions all the terms in this expression are positive except $\frac{\partial u^j}{\partial m}$ which we have already shown is non positive. Thus $\frac{\partial q}{\partial m} < 0$ and this implies that $q_{nL}^* < q_{nH}^*$. Next we will show that $\frac{u_{nL}}{q_{nL}^*} < \frac{u_{nH}}{q_m^*}$. Since $u_{nH} < u_{nL}$ we have that $l'(q_m^*) < l'(\frac{u_{nL}q_m^*}{u_{nH}})$ and $\frac{u_{nH}}{q_m^*} g'(\frac{u_{nH}}{q_m^*}) < \frac{u_{nL}^2}{u_{nL}q_m^*} g'(\frac{u_{nH}}{q_m^*})$. This implies that $l'(\frac{u_{nL}q_m^*}{u_{nH}}) - \frac{u_{nL}^2}{u_{nL}q_m^*} g'(\frac{u_{nH}}{q_m^*}) < l'(q_m^*) - \frac{u_{nH}}{q_m^*} g'(\frac{u_{nH}}{q_m^*}) = 0$ Thus $q_{nL}^* > \frac{u_{nL}q_m^*}{u_{nH}}$ which proves that $t_{nL}^* < t_{nH}^*$.

Thus we have shown that in all possible cases optimal q and t increases as n increases.

Proof of Proposition C.2.3

We will consider perfect testing case first. Let (q_{nH}^*, t_{nH}^*) be the optimal solutions for SC_{nH} and (q_{nL}^*, t_{nL}^*) be the optimal solutions for SC_L . Note that $E_m = \frac{k}{m} \int_{\beta}^x f(x, m, p_L) dx$ and $\beta = \frac{q(t/n)cm}{(r_H - r_L)}$. There are two cases to consider:

(i) If $t_{nH}^* \leq n_L$ then $q_{nH}^* \leq 1$ then (q_{nH}^*, t_{nH}^*) is a feasible solution for SC_L as long as $E_{nL}(q_{nH}^*, t_{nH}^*) \leq \alpha$. Since $\alpha \leq n \int_{\frac{mp}{3}}^x f(x, m, p_L) dx$ we have that $\frac{q(t/n_L)cm}{(r_H - r_L)} \geq mp/3$. From proposition ??, this implies that $\frac{\partial E_n}{\partial n} \geq 0$. Thus $E_{nL}(q_H^*, t_H^*) \leq E_{nH}(q_{nH}^*, t_{nH}^*) \leq \alpha$ and (q_{nH}^*, t_{nH}^*) is a feasible solution for SC_L . Since (q_{nL}^*, t_{nL}^*) is the optimal solution, this implies $\pi_{nL}^{PP}(q_{nL}^*, t_{nL}^*) \leq \pi_{nL}^{PP}(q_{nH}^*, t_{nH}^*) \leq \pi_{nH}^{PP}(t_{nH}^*, q_{nH}^*)$.

(ii) If $n_H \geq t_{nH}^* > n_L$ then we have that (q_{nH}^*, t_{nH}^*) is infeasible for SC_L . We will prove by contradiction that we can find a $t' \leq n_L$ s.t. (q_{nH}^*, t') is feasible for SC_L . Assume there does not exist a $t' < n_L$ s.t. (q_{nH}^*, t') is feasible. This implies that $n_L \int_{\frac{ckq_{nH}^*}{n_L(r_H - r_L)}}^x f(x, k/n_L, p) dx > \alpha$ since (q_{nH}^*, n_L) is also infeasible. Further since $t_{nH}^* \leq n_H$ we also have that $n_H \int_{\frac{ckq_{nH}^*}{n_H(r_H - r_L)}}^x f(x, m, p_L) dx \leq \alpha$. This implies that $E_{nH} = n_H \int_{\frac{ckq_{nH}^*}{n_H(r_H - r_L)}}^x f(x, k/n_H, p_L) dx < n_L \int_{\frac{ckq_{nH}^*}{n_L(r_H - r_L)}}^x f(x, k/n_L, p) dx = E_{nL}$. This is a contradiction since by Proposition 4.5.1 we know that $\frac{\partial E_n}{\partial n} \geq 0$. Thus we have that a feasible (q_{nH}^*, t') for SC_L exists. This implies $\pi_{nL} \leq l(q_{nH}^*) + g(t') < f(q_{nH}^*) + g(t_m^*) = \pi_{nH}$.

The proof follows similarly for imperfect testing scenarios since we have already proved that $\frac{\partial E_n}{\partial n} \geq 0$.

Proof of Proposition C.3.1

From theorem 4.3.1 and theorem 4.3.2 we already know farms' equilibrium adulteration strategy. Note that n_a^* can be equivalently interpreted as a mixed strategy equilibrium with each farm adulterating with probability n_a^*/n and not adulterating otherwise. The total (expected) amount of adulterants in the two cases are just kn_a^*/n and kx^{PV^*} . If $c < -h'(1)(r_H - r_L)/[q(t/n)(n + 1)/n]$, then we have that $n_a^* = n$ and $x^{PV^*} = 1$ from theorem 4.3.1 and theorem 4.3.2. Thus all farms adulterate to the maximum level. The result for the remaining parts in the proposition follows similarly by comparing the two values under different values for c .

Proof of Theorem C.3.2

If $c < -h'(1)(r_H - r_L)/[q(t/n)(n + 1)/n]$, then we have that all farms adulterate to the maximum level from theorem 4.3.2. This implies that the realized low quality likelihood for all farms is p_L^{min} . Since we have assumed $p_L^{min} = 0$, none of the farms will have any low quality units and the reactive EMA risk is 0. Thus $R_p^T = R_{ip}^T$. Next, note that if $c \geq (r_H - r_L)/[q(t/n)]\beta^{RP} = \frac{cqm(t/n)}{r_H - r_L} > m$ and reactive EMA risk is again 0 from theorem 4.3.3. The result in the theorem then follows directly from proposition C.3.1.

Appendix D

Appendix of Chapter 5

D.1 Model Notations

Variable	Definition
b_c	Consumers' private, heterogeneous budgets in Period 1; trader knows $b_c \sim \text{Uniform}[0, 1]$
q_i	Quantity sold by the trader in Period $i = 1, 2$
$p_i(q_i)$	Period i price when quantity sold in the market is q_i
g	Consumers' budget adjustment factor
λ	Consumption utility per consumer
$\delta(q_1)$	Relative increase in consumers' budgets when quantity sold in Period 1 is q_1
c	Amount of cash transferred to each eligible consumer under cash subsidy
f	Fraction of consumers subsidized under cash subsidy
b	Total budget of the government for intervention
q_g	Total quantity sold by the government under supply allocation schemes
p_g	Price of the commodity charged by the government under supply allocation schemes

Table D.1: Model Notations

D.2 Additional Analytical Results

Proposition D.2.1. Let the government's CS policy be such that f fraction of consumers are each subsidized with an amount c . Then the market price in Period 2 given that the trader sells quantity q_2 is as follows.

- (i) If $q_2 \leq |f + \frac{c}{1 + g(1 - q_1)} - 1|$, then $p_2 = (1 + g(1 - q_1))(\max(f + \frac{c}{1 + g(1 - q_1)}, 1) - q_2)$.
- (ii) If $q_2 \in (|f + \frac{c}{1 + g(1 - q_1)} - 1|, \min(1 + \frac{c}{1 + g(1 - q_1)} - f, 1 + f - \frac{c}{1 + g(1 - q_1)})]$, then $p_2 = (1 + g(1 - q_1))(1 + f + \frac{c}{1 + g(1 - q_1)} - q_2)/2$.

(iii) If $q_2 \in (\min(1 + \frac{c}{1+g(1-q_1)} - f, 1 + f - \frac{c}{1+g(1-q_1)}), 1]$, then $p_2 = (1 + g(1 - q_1))(1 + \min(f, \frac{c}{1+g(1-q_1)}) - q_2)$.

Proposition D.2.2. Let $z \equiv \frac{1}{2} + \frac{c}{2} - \sqrt{\frac{2c(1+\delta) + c^2}{4(1+\delta)^2}}$. The optimal q_2^* under CS for the wholesaler is the following:

(i) If $\frac{c}{1+\delta} \in [0, \frac{1}{2\sqrt{2}})$:

(a) If $f \leq \min(z, \sqrt{2} - 1 - \frac{c}{1+\delta})$, then $q_2^* = 1/2$

(b) If $f \in (\min(z, \sqrt{2} - 1 - \frac{c}{1+\delta}), \max(z, \frac{1}{3} + \frac{c}{3(1+\delta)})]$ then $q_2^* = \frac{1}{2}$ if $\frac{c}{1+\delta} \leq \frac{3\sqrt{2}}{4} - 1$ and $q_2^* = \frac{1+f + \frac{c}{(1+\delta)}}{2}$ otherwise.

(c) If $f \in (\max(z, \frac{1}{3} + \frac{c}{3(1+\delta)}), \frac{1}{2} + \frac{c}{2(1+\delta)})]$ then $q_2^* = 1 - f + \frac{c}{(1+\delta)}$

(ii) If $\frac{c}{1+\delta} \in [\frac{1}{2\sqrt{2}}, 1)$:

(a) If $f \leq \max(\frac{2c}{1+\delta} - 1, \frac{1+\delta}{4c} - 1 + \frac{c}{1+\delta})$ then $q_2^* = \frac{1 + f \mathbf{1}_{\{2c > (1+\delta)\}}}{2}$

(b) If $f \in (\max(\frac{2c}{1+\delta} - 1, \frac{1+\delta}{4c} - 1 + \frac{c}{1+\delta}), \min(\frac{3c}{(1+\delta)} - 1, \frac{c}{1+\delta})]$ then $q_2^* = 1 + f - \frac{c}{(1+\delta)}$

(c) If $f \in (\min(\frac{3c}{(1+\delta)} - 1, \frac{c}{1+\delta}), \max(\frac{1}{3} + \frac{c}{3(1+\delta)}, \frac{c}{1+\delta})]$ then $q_2^* = \frac{1+f + \frac{c}{(1+\delta)}}{2}$

(d) If $f \in (\max(\frac{1}{3} + \frac{c}{3(1+\delta)}, \frac{c}{1+\delta}), \frac{1}{2} + \frac{c}{2(1+\delta)})]$ then $q_2^* = 1 - f + \frac{c}{1+\delta}$

(iii) If $f \in (\frac{1}{2} + \frac{c}{2(1+\delta)}, 1]$ then $q_2^* = \frac{1}{2} + \frac{c}{2(1+\delta)}$

Lemma D.2.3. Define t_1 to be the solution to the following fixed point equation:

$$\frac{(80 + 8g + g^2)}{64} = \max_{t_1 \in [0,1]} \left\{ t_1(1-t_1) + 2 \frac{1 + g(1-t_1) + \sqrt{(1+g(1-t_1))(1+12b+g(1-t_1))^2}}{36(1+g(1-t_1))} \right\}$$

We have that $t_1 \leq \frac{8+4g+g^2}{32}$ for all $q_1 \in [0, 1]$. Further, if $(1+g)\frac{3-2\sqrt{2}}{8} \leq b$ then $t_1 \leq 0$

Proposition D.2.4. Let $q_d \equiv 1 - p_g/(1 + g(1 - q_1))$. The market price in Period 2, given that the trader sells q_2 and the government intervenes by selling q_g inventory at a subsidized price p_g , is as follows.

- (i) Under PA: (a) If $q_2 < q_d - q_g$, then $p_2 = (1 + g(1 - q_1))(1 - q_2)$; (b) If $q_2 \geq q_d - q_g$, then $p_2 = (1 + g(1 - q_1))(1 - q_2 - q_g)$.
- (ii) Under RA: (a) If $q_2 < q_d - q_g$, then $p_2 = (1 + g(1 - q_1))(1 - q_2/(1 - q_g/q_d))$; (b) If $q_2 \geq q_d - q_g$, then $p_2 = (1 + g(1 - q_1))(1 - q_2 - q_g)$.

- Proposition D.2.5.** (i) Under PA: (a) If $q_d \leq (1 + 2q_g - \sqrt{2q_g - q_g^2})/2$ then $q_2^* = (1 - q_g)/2$. (b) If $q_d \in [(1 + 2q_g - \sqrt{2q_g - q_g^2})/2, q_g + 1/2)$ then $q_2^* = q_d - q_g$. (c) If $q_d \geq q_g + 1/2$ then $q_2^* = 1/2$.
- (ii) Under RA: (a) If $q_d \leq 1/(2 - q_g)$ then $q_2^* = (1 - q_g)/2$. (b) If $q_d > 1/(2 - q_g)$ then $q_2^* = (1 - q_g/q_d)/2$

Proposition D.2.6. The aggregate consumer surplus in the two periods under all three interventions increases as b increases. Further $CS_I > CS_\phi$, for all $b \geq 0$ and $I \in \{CS, PA, RA\}$.

Proposition D.2.7. Let R_t^I be the aggregate revenue of the trader under each intervention $I \in \{\phi, CS, PA, RA\}$. Comparing R_ϕ^t with R_I^t , we have the following:

- (i) Under CS: $R_\phi^t \leq R_{CS}^t$ for all $b \geq 0$. Further, as b increases R_{CS}^t also increases.
- (ii) Under PA: (a) If $b \leq p_b/2$, $R_\phi^t = R_{PA}^t$ and R_{PA}^t is constant in b . (b) If $b > p_b/2$, $R_\phi^t > R_{PA}^t$ and R_{PA}^t is decreasing in b .
- (iii) Under RA: $R_\phi^t \leq R_{RA}^t$ for all $b \geq 0$. Further, as b increases R_{RA}^t decreases.

D.3 Model Extensions

We examine multiple model extensions to demonstrate the robustness of our conclusions to a number of modeling assumptions.

D.3.1 Skewed Budget Distribution

In §5.3, we assume that the budget distribution of consumers in Period 1 is uniformly distributed. We now analyze an extension in which consumers' budget in Period 1 follows an exponential distribution with mean $1/\lambda$. The larger probability mass over smaller values in the exponential distribution allows us to capture the setting where low-income consumers dominate the retail market. Proposition D.3.1 confirms that it is again optimal for the trader to create artificial shortage in the market.

Proposition D.3.1. Under no government intervention, the optimal quantity released and the optimal price charged by the trader in each period is as follows: $(q_{1,E}^*, p_{1,E}^*) = (e^{-(1+g/e)}, \frac{1}{\lambda} + \frac{g}{\lambda e})$, and $(q_{2,E}^*, p_{2,E}^*) = (e^{-1}, \frac{1 + g(1 - e^{-(1+g/e)})}{\lambda})$. Further, if $g > 0$, $q_1^* \leq q_2^*$ and $p_1^* \leq p_2^*$.

The next result highlights that the intensity of shortage at equilibrium is intensified when the budget is distributed exponentially. Let $q^*(g)_{1,D}$ be the optimal quantity released by the trader in Period 1 when budget adjustment factor is g and consumers' budget follows $D \in \{Uniform(U), Exponential(E)\}$. Relative shortage intensity (compared to the case when $g = 0$) is, $S_D = 1 - \frac{q^*(g)_{1,D}}{q^*(0)_{1,D}}$. Our next result confirms that the relative shortage is always larger when consumers' budget follows an exponential distribution.

Lemma D.3.2. The relative shortage under no intervention is more when the budget follows an exponential distribution (S_E) than when it follows a uniform distribution (S_U) for any $g \in [0, 1]$ and $\lambda \geq 0$.

Let $q_{I,E}^t$ be the optimal quantity released by the trader in period t under intervention $I \in \{\phi, RA, PA\}$. The next result confirms that the key policy insights related to the two supply allocation schemes discussed in §6 continue to hold in this setting.

Proposition D.3.3. Assume $\mu = 0$,

- (i) Under PA , if $b \leq p_b(1 - q_{\phi,E}^2)$, then $q_{PA,E}^1 = q_{\phi,E}^1$. Otherwise if $b > p_b(1 - q_{\phi,E}^2)$, $q_{PA,E}^1 \geq q_{\phi,E}^1$.
- (ii) Under RA , for all $b > 0$, $q_{RA,E}^1 > q_{\phi,E}^1$.

Although we cannot analytically show that the results under CS also continue to hold, we do extensive numerical simulations to characterize the solution. We solve the tri-level optimization numerically and find that similar to Proposition 5.6.1, equilibrium quantity released by the trader is in fact smaller under cash subsidy scheme relative to the no intervention case (Figure D.1).

D.3.2 Divisible Good

In §5.3, we assume that the consumers' purchase behavior is binary: consumers buy a unit if $b_i < p$, and do not buy otherwise. In this extension, we analyze an additional setting where consumers may buy fractional units if their budget is less than the market price. In particular,

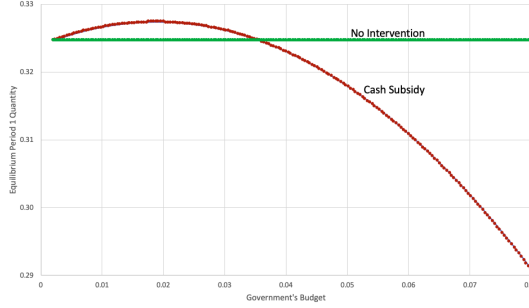


Figure D.1: Equilibrium Period 1 Quantity with Exponential Budget Distribution

Note: We use the following parameters in this example: $g = 0.43$, $b \in \{0.002, \dots, 0.08\}$. For computational reasons, we first evaluate the optimal solution $(q_{1,e}^*)$ at 20 points in this range of b and then fit a polynomial of degree 2 to approximate results on other points.

consumers with budget $b_i \in [p/n, p]$ purchase b_i/p units and those with $b_i > p$ purchase 1 unit. We use the parameter n to model the consumption of a minimal unit (e.g., one onion) because consumers cannot practically buy an infinitesimal unit of an essential commodity. That is, when consumers' budget is lower than the market price, consumers purchase a minimum of $1/n$ units if they can afford it (with $n \geq 1$); otherwise, consumers make no purchase. $n = 1$ corresponds to our base model of a binary purchase decision; as $n \rightarrow \infty$, we converge to the case when consumers are willing to purchase an infinitesimal unit. In line with the original formulation, the updated budget of a consumer in period 2 with initial budget b_i is modeled as $b_i(1 + g(1 - q_1))$. Our first result confirms that it is again optimal for the trader to create artificial shortage under no intervention.

Proposition D.3.4. Under no government intervention, the optimal quantity released and the optimal price charged by the trader in each period are as follows:

- (i) If $n < \frac{2}{\sqrt{g}}$, then $(q_1^*, p_1^*) = \left(\frac{1}{2} - \frac{g}{8}, \frac{n^2}{1+n^2} \left(1 + \frac{g}{4}\right)\right)$ and $(q_2^*, p_2^*) = \left(\frac{1}{2}, \frac{n^2}{1+n^2} \left(1 + \frac{g}{2} + \frac{g^2}{8}\right)\right)$
- (ii) If $n \geq \frac{2}{\sqrt{g}}$, then $(q_1^*, p_1^*) = \left(\frac{1}{2} - \frac{1}{2n^2}, 1\right)$ and $(q_2^*, p_2^*) = \left(\frac{1}{2}, \frac{g}{2} + \frac{2n^2}{2(1+n^2)}\right)$

Further, $q_1^* < q_2^*$ and $p_1^* < p_2^*$ in both the cases.

Our key insights continue to hold in this extension. Observe from Proposition D.3.4 that, as long as $g > 0$, we have both $q_1^* \leq q_2^*$ and $p_1^* < p_2^*$. The above result confirms the importance of effective government interventions to reduce the effect of artificial shortage in the markets. Next, we confirm that the key policy insights related to the supply allocation schemes in Proposition 5.6.2 continue to hold in this setting.

Proposition D.3.5. We have the following:

- (i) Under PA , if b is less (greater) than a threshold, $q_{PA}^1 = q_\phi^1 (q_{PA}^1 \geq q_\phi^1)$.
- (ii) Under RA , $q_{RA}^1 > q_\phi^1$ for all $b > 0$.

While we cannot show the result from Proposition 5.6.1 analytically for CS , extensive numerical simulations confirm that the equilibrium shortage increases when the government budget is larger than a threshold (Figure D.2).

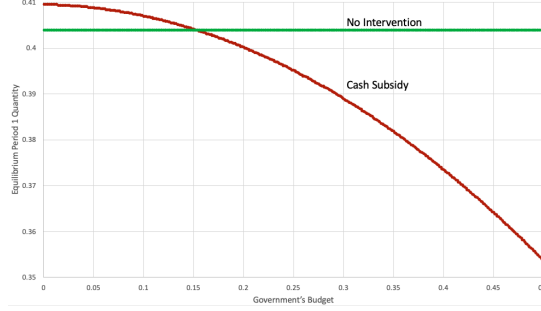


Figure D.2: Equilibrium Period 1 Quantity with Divisible Good

Note: We use the following parameters in this example: $n = 2$, $g = 0.43$, $\lambda = 0.43$, $b \in \{0.001, \dots, 0.5\}$. For computational reasons, we first evaluate the optimal solution ($q_{1,e}^*$) at 20 points in this range of b and then fit a polynomial of degree 2 to approximate results on other points.

D.3.3 Trader with Inventory Holding Cost and Inventory Constraint

In §5.3, we assume that the trader does not incur an inventory holding cost when hoarding supply in Period 1. To confirm the robustness of our results, we analyze an extension in which the trader procures Q units of inventory *at the beginning of* the two periods and incurs a per-unit holding cost, h , for all units that are unsold in Period 1.

Proposition D.3.6. (i) If $Q \geq 1 - g/8 + h/2$, $q_{\phi,c}^1 = (1 + h)/2 - g/8$, and $q_{\phi,c}^2 = 1/2$

(ii) If $Q < 1 - g/8 + h/2$, $q_{\phi,c}^1 = \max\left\{0, \frac{(2 + 2gQ - \sqrt{4 + 3g^2 + 2gQ - 3g^2Q + g^2Q^2 - 3gh})}{3g}\right\}$
and $q_{\phi,c}^2 = Q - q_{\phi,c}^1$.

Further, if $h \leq g/4$, $q_{\phi,c}^1 < q_{\phi,c}^2$ and $p_{\phi,c}^1 < p_{\phi,c}^2$ in both the cases.

Consistent with intuition, Proposition D.3.6 shows that if the per-unit inventory holding cost is small relative to the budget adjustment factor ($h \leq g/4$), then it is optimal for the trader to create artificial shortage under the no-intervention case. Otherwise, if $h > g/4$, then it is no longer optimal for the trader to hoard supply for Period 2. Nevertheless, given repeated and persistent occurrences of traders illegally hoarding supply of essential commodities in practice

(The New Indian Express 2019), we expect that the holding cost is not prohibitively high that it would stop the traders from creating artificial shortage in the market. Similar to Proposition 5.6.1, Proposition D.3.7 characterizes how shortage changes under PA and RA relative to the case when there is no government intervention.

Proposition D.3.7. We have the following:

- (i) Under PA , if $\mu \leq \frac{2\lambda}{1+g}$,
 - (a) If $b \leq (1 - q_{\phi,c}^2)pb$, then the optimal $q_{PA,c}^1 = q_{\phi,c}^1$.
 - (b) If $b > (1 - q_{\phi,c}^2)pb$, the optimal $q_{PA,c}^1 \geq q_{\phi,c}^1$.
- (ii) Under RA , for all $b > 0$, the optimal $q_{RA,c}^1 > q_{\phi,c}^1$.

The above proposition shows that supply allocation schemes are effective in reducing artificial shortage even when there is *objective* shortage in the market. While PA becomes effective only when government's budget is large enough, RA is always effective in reducing shortage in the market. Although we cannot analytically show that the results under CS also continue to hold, we do extensive numerical simulations to characterize the solution. We solve the tri-level optimization numerically and find that similar to Proposition 5.6.1, equilibrium quantity released by the trader is in fact smaller under cash subsidy scheme relative to the no intervention case (Figure D.4).. Thus, excessive spending from the government further increases shortage even when the trader faces supply constraints.

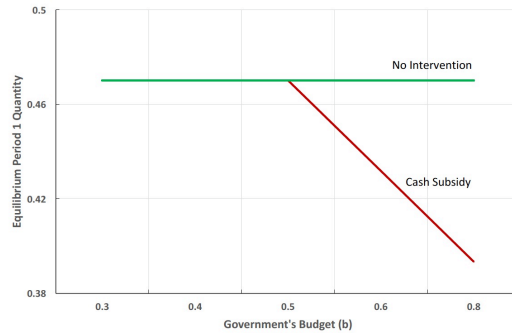


Figure D.3: Equilibrium Period 1 Quantity under Inventory Constraint

Note: We use the following parameters in this example: $Q = 0.5$, $\lambda = 2$, $g = 0.43$, $b \in \{0.3, \dots, 0.75\}$, $h = 0$

D.3.4 Unequal Relative Budget Increase in Period 2 Across Consumers

In §5.3, we assume that all consumers increase their budgets in Period 2 by the same relative amount. In this section, we analyze a setting where consumers with a higher initial budget increase their budgets by a higher relative amount. To model this setting, we assume that $\delta(\cdot)$, denoting the relative increase of a consumer's budget, is a function of both b_c and q_1 , i.e. $\delta(q_1, b_c) = g(1 - q_1)b_c$. Consumers with initial budget b_c update it to $b_c(1 + \delta(q_1, b_c))$ in Period 2. The Period 2 price-demand curve is no longer linear but convex in this setting.

Proposition D.3.8. $(q_1^*, p_1^*) = (\frac{1}{2} - \frac{g(1 - q_2^*)^2 q_2^*}{2}, \frac{1}{2} + \frac{g(1 - q_2^*)^2 q_2^*}{2})$. $q_2^* \in (0, \frac{1}{2})$ and is the solution to the following equation: $1 - 2q_2^* + \frac{g(1 + gq_2(1 - q_2)^2)(1 - 4q_2 + 3q_2^2)}{2} = 0$. Finally, $p_2^*(q_2^*) \geq \frac{1}{2}$.

The above proposition shows that it is indeed optimal for the trader to create artificial shortage in the first period under no government intervention. Note that because the demand in the second period is now convex, the optimal quantity in the second period is smaller than second period quantity under the base model analyzed in §5.3. Thus in addition to reducing supply in the first period, the trader finds it beneficial to also reduce supply in the second period relative to the base case.

Similar to the previous extension, we confirm that the key insights regarding the effectiveness of the three government interventions continue to hold in this setting (Proposition D.3.9).

Proposition D.3.9. If $\mu \leq \frac{1}{2}$. We have the following:

- (i) Under PA ,
 - (a) If $b \leq (1 - q_2^*)p_b$, then the optimal $q_{PA}^1 = q_\phi^1$.
 - (b) If $b > (1 - q_2^*)p_b$, the optimal $q_{PA}^1 \geq q_\phi^1$.
- (ii) Under RA , for all $b \geq 0$, the optimal $q_{RA}^1 > q_\phi^1$.

While we cannot analyze the cash subsidy scheme analytically, we do extensive numerical simulations to understand its impact on equilibrium shortage. We find that similar to Proposition 5.6.2, if the government's budget is larger than a threshold then the equilibrium shortage is indeed larger compared to the base case when there is no government intervention (Figure D.4).

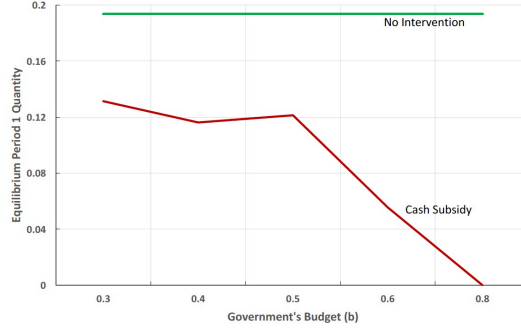


Figure D.4: Equilibrium Period 1 Quantity under Under Unequal Relative Budget Increase

Note: We use the following parameters in this example: $\lambda = 2$, $g = 0.43$, $b \in \{0.3, \dots, 0.75\}$

D.3.5 Alternative Government Interventions

In §5.3, we assume that the government can only intervene *reactively* in Period 2. In this extension, we consider cases where government considers intervening in both periods with an objective of maximizing the total consumer welfare in both periods. We assume that the government procures inventory at price p_b per unit *at the beginning of* the two periods and decides how to optimally release the inventory to the market across the two periods. Since the objective of any such intervention would be to increase consumer welfare, we assume that μ is 0. Recall from §5.5.2 that μ captures government's relative valuation for money earned back from the intervention. Since $\mu = 0$, it is optimal for the government to buy as much as possible, ($q_g = \frac{b}{p_b}$) and distribute it (optimally) across the two periods at the lowest price possible ($q_d = 1$). We analyze optimal government strategies for both *PA* and *RA*. Let q_g^1 and q_g^2 be the quantity released by the government in the two periods at equilibrium.

Proposition D.3.10. (i) Under *PA*

(a) If $q_g \in [1 + \frac{g}{8}, 2]$, the the optimal $(q_g^1, q_g^2) = (\min\{q_g - 1/2, 1\}, \max\{1/2, q_g - 1\})$

(b) If $q_g \in [0, 1 + \frac{g}{8}]$, the optimal $(q_g^1, q_g^2) = (\min\{q_g, 1\}, \max\{0, q_g - 1\})$.

(ii) Under *RA*, the optimal $(q_g^1, q_g^2) = (\max\{0, q_g - 1\}, \min\{q_g, 1\})$

The above result shows that the government's optimal strategy depends on whether it adopts prioritized allocation (PA) or randomized allocation (RA). Under PA, it is optimal for the government to intervene *preemptively* by releasing more quantity in Period 1 than in Period 2. However, under RA, it is optimal for the government to intervene *reactively* by releasing more quantity in Period 2 than in Period 1. These contrasting results are due to the following reasons. Under PA, the government's intervention effectively increases total supply in the market

by serving lower-budget consumers who could not afford to purchase from the trader anyways. Therefore, intervening preemptively has the dual benefit of both mitigating consumers' potential budget increase and serving more consumers, whereas intervening reactively benefits consumers primarily through serving more consumers. In contrast, under RA, the government's intervention creates supply competition for the trader over *all* consumers. As a result, the trader responds to the government's quantity decision in either period to ensure that the total supply in the market across both periods remains the same regardless of whether the government intervenes preemptively or reactively. In this case, if the government were to intervene preemptively, then the trader would further reduce first-period quantity to induce consumers to increase their budgets and then benefit by selling at a higher price in Period 2 (as the government would have limited supply in Period 2). Conversely, if the government intervenes reactively, then it creates the maximum extent of supply competition in Period 2. Such competition would force the trader to sell more quantity in Period 1, which would help to mitigate consumers' budget increase in the first place, thus bringing more benefit to the consumers.

D.3.6 Effect of Government Interventions on Consumers' Budget Adjustment

In §5.3, we assume that the consumer's budget adjustment factor is exogenous and constant across different scenarios. In this extension we analyze the cases where consumers' budget adjustment factor is affected by government interventions. In particular, we assume that consumers' budget adjustment factor under no intervention g is reduced to $g' < g$ if the government intervenes by *CS*, *PA* or *RA*. This extension captures scenarios where the knowledge (or expectation) about potential government interventions relieves some anxiety among consumers.

Proposition D.3.11. Under all government interventions, q_1^* is decreasing in g . Further, we have the following:

- (i) Comparing q_ϕ^1 to q_{CS}^1 under *CS*, we have: (a) If $g' \leq \frac{g}{4}$, then $q_{CS}^1 \geq q_\phi^1$ for all $b > 0$. (b) If $g' > \frac{g}{4}$, $q_{CS}^1 < q_\phi^1$ if and only if b is greater than a threshold.
- (ii) Comparing q_ϕ^1 to q_{PA}^1 under *PA*, we have $q_{PA}^1 \geq q_\phi^1$ for all $b > 0$. If b is less (greater) than a threshold, q_{PA}^1 is constant (increasing) in b .
- (iii) Comparing q_ϕ^1 to q_{RA}^1 under *RA*, we have $q_{RA}^1 \geq q_\phi^1$ for all $b > 0$. q_{RA}^1 is always increasing in b .

The above proposition shows that the additional benefit of reduction in consumer budget adjustment from g to g' can further reduce trader's incentives to create shortage in the first period under all three policies. The key policy insights discussed in §5.6 continue to hold in this case with the following changes. First, if the reduction in consumers' budget adjustment factor is substantially large, then cash subsidy would be beneficial (i.e., mitigating artificial shortage) for all government budget levels. Second, similar to randomized allocation, prioritized allocation also always leads to a larger quantity released by the trader in Period 1. Nevertheless, when the government's budget is small, randomized allocation remains more effective than prioritized allocation as the former leads to a larger Period 1 quantity.

D.3.7 Reference Price Model

In this section, we show the inability of reference-dependent demand models (Thaler 1985, Tversky and Kahneman 1991) to capture the phenomenon of strategic shortages in agricultural markets. Mental accounting theory posits that consumers utility from purchase consists of two components: consumption utility and transaction utility. Transaction utility captures the psychological value of the purchase relative to a reference point. A larger (smaller) price in the market relative to the reference price leads to a loss (gain) in transactional utility that consumers account for when making a purchase decision. It is reasonable to assume that the reference price for consumers in the second period is just the price they observe in the first period in a temporal model. We assume that the reference price effects are linear in nature following Nasiry (2010). Under this assumption, demand in the second period (q_2), is the following: $q_2 = 1 - p_2 - \gamma_l(p_2 - p_1)^+ + \gamma_g(p_1 - p_2)^+$. The last two terms capture the transaction utility in the following way. If period 2 price is higher (lower) than period 1, then consumers feel a psychological loss (gain) and demand suffers (increases). Finally, $\gamma_l > \gamma_g$ to capture the fact that consumers are loss averse. Our next lemma shows that a demand model with linear reference effects is unable to capture the phenomenon of strategic shortage in agricultural market. In particular, under the reference dependent model, the trader cannot increase price in the second period *and* also sell more quantity in period 2.

Lemma D.3.12. Under a model with reference dependent demand, it will never be the case that both $p_2 > p_1$ and $q_2 > q_1$. Further $p_2^* < p_1^*$ at equilibrium.

The above lemma shows that it can never be the case that *both* prices and quantity released in period 2 are higher than that in period 1 if we assume a reference-dependent demand model.

This result highlights the need to develop a novel model to capture consumer and trader behavior that drive artificial shortage in agricultural markets.

D.4 Proof of Analytical Results

Proof of Proposition 5.3.1

Note that the total revenue of the trader is given by $\pi_\phi(q_1, q_2) = q_1(1 - q_1) + q_2(1 - q_2)(1 + \delta(q_1))$. It is easy to check that the revenue function is optimized at $q_2 = \frac{1}{2}$ for all values of q_1 because the function is concave in q_2 . Replacing $q_2^* = \frac{1}{2}$ and $\delta(q_1) = g(1 - q_1)$ and differentiating with respect to q_1 we get $\frac{\partial \pi_\phi(q_1, 1/2)}{\partial q_1} = 1 - 2q_1 - \frac{g}{4}$ and $\frac{\partial^2 \pi_\phi(q_1, 1/2)}{\partial q_1^2} = -2$. This implies that the revenue function is concave in q_1 and the optimal solution can be obtained by setting the first order condition to 0.

Proof of Proposition 5.4.1

Note that the consumer surplus in both the periods is given by the following $CS_\phi^i = \int_0^{q_i} (1 - q_i) - q_i p_i + \lambda(2q_i - 1)$. Replacing the values of q_i and p_i from the proposition 5.3.1 gives us the result.

Proof of Theorem 5.5.1

Let us again consider the case when $b < \frac{3 - 2\sqrt{2}}{8}(1 + \delta)$. We first consider the problem of finding an optimal f for a given c . Note that $f \leq \frac{b}{c}$ has to be true to satisfy government's budget constraint. If $\frac{c}{1 + \delta} \in [0, \frac{3\sqrt{2}}{4} - 1)$, then we already have the optimal second period trader's quantity from proposition D.2.2. Define $th \equiv \frac{1}{2} + \frac{c}{2} - \frac{\sqrt{\frac{2c(1+\delta)+c^2}{(1+\delta)^2}}}{2}$. If $a < th$, then consumer surplus is not affected by f since $q^* = \frac{1}{2}$. If $f \in (th, \frac{1}{2} + \frac{c}{2(1+\delta)}]$ then $q_2 = 1 - f + \frac{c}{(1+\delta)}$, $ns = 1 - f$ and $s = \frac{c}{1+\delta}$. Replacing these values in consumer surplus's definition, we have $\frac{\partial CW(f, c)}{\partial f} \leq 0$. The inequality follows from the fact that $f \leq \frac{1}{2} + \frac{c}{2(1+\delta)}$ and assumption that $\lambda \geq 1 + \delta$. This implies that the consumer surplus is always decreasing in this region. If $f \in (\frac{1}{2} + \frac{c}{2(1+\delta)}, 1]$, then $q_2 = \frac{1}{2} + \frac{c}{2(1+\delta)}$ and it is again easy to check that the consumer surplus increases this region. Note that the consumer surplus is discontinuous at th because of a jump in q^* but it is easy to check that the surplus is more when $q_2 = th$ than when $q_2 = 1/2$ and consumer surplus is greater at $f = th$ compared to when $f = 1$. Thus the optimal f for a given c in this region is $f^* = \min(th, \frac{b}{c})$. Further, note that the c s.t. $th = \frac{b}{c}$ is given by $c_1 \equiv 2 \frac{b(1+\delta) + 2b\sqrt{b(1+\delta)}}{1+\delta - 4b}$ and $c_1 \leq (1+\delta) \frac{3\sqrt{2}}{4} - 1$ is always true when $b < \frac{3 - 2\sqrt{2}}{8}(1 + \delta)$.

Thus $f = \frac{b}{c}$ for all $c \geq c_1$. Now that we have characterized the optimal f , we find optimal c .

Note that it is easy to check that the consumer surplus is increasing for all $c \leq c_1$. For $c \in (c_1, (1 + \delta)\frac{3\sqrt{2}}{4} - 1)$, since $f \leq th$, $q^* = \frac{1}{2}$ and consumer surplus is smaller than that at c_1 . Next, if $\frac{c}{1 + \delta} \in [\frac{3\sqrt{2}}{4} - 1, \frac{1}{2\sqrt{2}})$ then it is easy to check that $\frac{b}{c} \leq \sqrt{2} - 1 - \frac{c}{1 + \delta}$ is always true and $q_2^* = \frac{1}{2}$ again. If $\frac{c}{1 + \delta} \in [\frac{1}{2\sqrt{2}}, \frac{1}{2})$, then note that $q_2 = 1 + \frac{b}{c} - \frac{c}{1 + \delta}$ is decreasing in c and $\frac{b}{c} \leq \frac{3c}{1 + \delta} - 1$ is always true. Thus, in this region, consumer surplus is maximum at the smallest c such that the $q_2^* > \frac{1}{2}$. Thus the optimal c is such that $\frac{b}{c} = -1 + \frac{c}{1 + \delta} + \frac{1 + \delta}{4c}$. Solving this equation gives us that $c_2 = \frac{1 + \delta - 2\sqrt{b(1 + \delta)}}{2}$. It is easy to check that $q(c_2) = q(c_1)$, $s(c_2) = s(c_1)$ and $p(c_2) = p(c_1)$ because $c_1 c_2 = b(1 + \delta)$. Thus, all the terms in the consumer surplus definition at the two points are same except sc . Since $c_2 > c_1$, we have that $CS(c_2, f) > CS(c_1, f)$. Finally, note that if $c \geq \frac{1 + \delta}{2}$, $\frac{b}{c} \leq \frac{c}{1 + \delta}$ is always true when $b < \frac{3 - 2\sqrt{2}}{8}(1 + \delta)$. Since $1 + \frac{b}{c} - \frac{c}{1 + \delta}$ and $\frac{1}{2} + \frac{b}{2c}$ are both decreasing in c , this implies that the consumer surplus is always decreasing when $c \geq \frac{1 + \delta}{2}$. Thus consumer surplus is maximum at c_2 when $b < \frac{3 - 2\sqrt{2}}{8}(1 + \delta)$. The result for other ranges of b follow similarly and we avoid redoing the proof here for the sake of brevity.

Proof of Theorem 5.5.2

It is easy to check that the total revenue of the trader is a bimodal function. It is piecewise concave: if $q_1 \geq \delta^{-1}(\frac{8b}{3 - 2\sqrt{2}} - 1) = \frac{1}{g} + 1 - \frac{8b}{g(3 - 2\sqrt{2})}$ and if $q_1 < \delta^{-1}(\frac{8b}{3 - 2\sqrt{2}} - 1)$, then the revenue function is again concave. However, note that while the revenue function increases as b increases if $q_1 \in (\delta^{-1}(\frac{8b}{3 - 2\sqrt{2}} - 1), \delta^{-1}(4b - 1)]$, it remains unaffected if $q_1 \leq (\delta^{-1}(\frac{8b}{3 - 2\sqrt{2}} - 1))$. This implies that there exists a threshold value such that for any b greater than that threshold, the optimal revenue will be obtained at $q_1 \geq \delta^{-1}(\frac{8b}{3 - 2\sqrt{2}} - 1)$. t_1 in the proposition D.2.3 is thus obtained by comparing the optimal revenue function value at the two modes and finding the threshold at which optimal value at the second mode is greater than that of the first one. For all other parts, we can use the fact that the revenue function is concave for all $q_1 \geq \delta^{-1}(\frac{8b}{3 - 2\sqrt{2}} - 1)$. If the total revenue is decreasing at $\delta^{-1}(4b - 1)$, then this implies that the function is decreasing for all $q_1 \geq \delta^{-1}(4b - 1)$ from concavity of the revenue function. Thus the optimal q_1 is in the second region. The condition in the second part is thus obtained by reducing the condition that $\frac{\partial \pi(\delta^{-1}(4b - 1))}{\partial q_1} \leq 0$. We can similarly find conditions under which the optimal solution is in the other two regions to obtain the remaining parts of the proposition.

Proof of Theorem 5.5.3

Note that if $q_d \leq \frac{1+2q_g}{2} - \frac{\sqrt{2q_g - q_g^2}}{2}$ then it is easy to check that the constraint that price charged by the government (p_g) is lower than the price charged by the trader is never satisfied. Thus, we restrict our analysis to the cases where $q_d \geq \frac{1+2q_g}{2} - \frac{\sqrt{2q_g - q_g^2}}{2}$. Note that the government is maximizing $\pi_g = CS(q_g, q_d) + \mu q_g p_g$. First, consider the case when $q_d - q_g \leq \frac{1}{2}$. In this case $q_2 = q_d - q_g$ from trader's optimal solution. Note that $\frac{\partial \pi_g^2}{\partial^2 q_d} = 1 + 2\delta$ and we only need to compare the boundary points to find the optimal q_d^* . Under the condition that $\mu \leq \frac{2\lambda}{1+\delta}$, it is easy to check that for all q_g , $q_d = 1$ is indeed optimal. Using, $q_d = 1$, it is easy to check that the consumer surplus is always increasing in q_g under the condition that $q_g \geq \frac{1}{2}$. Thus $q_g^* = \frac{b}{p_b}$ and $q_d^* = 1$. We can similarly consider the case when $q_d - q_g > \frac{1}{2}$, in which case $q_2 = \frac{1}{2}$. In this case, π_g is always increasing in q_d and q_g under the assumption that $\mu \leq (2\lambda)/(1+\delta)$. This proves the first part.

Next, we show the result for RA . We will first prove that the government's objective function, $\pi_g(q_g, q_d) = CS^{RA}(q_g, q_d) + \mu p_g q_g$ is again increasing in q_g for all $q_d \geq \frac{1}{2 - q_g}$. Note that $\frac{\partial \pi_g(q_g, q_d)}{\partial q_g} = \frac{(-1 + 2q_d)(1 + 8\lambda + 2q_d + \delta(-2 + 4q_d))}{8q_d} + \mu(1+\delta)(1-q_d) \geq 0 + \mu(1+\delta)(1-q_d)$. The last inequality follows from the constraint that $q_d \geq \frac{1}{2 - q_g} > \frac{1}{2}$ and the assumption that $\lambda - 1 \geq \delta$. Thus we have that $q_g^* = \frac{b}{p_b}$. Next we show that consumer surplus is concave in q_d . $\frac{\partial^2 CS^{RA}(q_g^*, q_d)}{\partial q_d^2} = \frac{(-1 + 2\delta - 8\lambda)q_g}{4q_d^3} \leq 0$. The last inequality again follows from the assumption that $\lambda - 1 \geq \delta$. The result in the proposition follows directly from the fact that we are maximizing a concave function with constraints.

Proof of Theorem 5.5.4

Note that the first period revenue of the trader is always $q(1 - q)$. The second period revenue follows directly by replacing the optimal q_g^* and q_d^* obtained in proposition 5.5.3. The optimal solution is obtained by setting the first order condition to 0 since the revenue function is concave in q_1^* .

Next, we show the result for RA . We first show that the total revenue is concave in each of the three regions. In the region (i), $\frac{\partial^2 \pi^{RA}}{\partial q^2} = -2 + \frac{(1 - q_g^*)}{4} \delta'' \leq 0$. The last inequality follows from the assumption that δ is concave in q_1 (Note that our result in this proof holds for more general form of anxiety than the linear form we have assumed throughout the paper). In region

(ii),

$$\begin{aligned} \frac{\partial^2 \pi^{RA}}{\partial q^2} = & -2 - \delta^2 \frac{(q_g(-2 + 8\lambda(-1 + \mu) + 3\mu)(-2(2 + 4\lambda + \delta(5 + 12\lambda)) + 3(1 + \delta)(3 + 8\lambda)\mu)}{2(1 - 2\delta + 8\lambda)^{5/2}(-1 + 2\delta(-1 + \mu) + 2\mu)(3/2)} \\ & - \delta'' \left(\frac{1}{4} + \frac{2q_g(3 + 2\delta - 4\delta^2 + 16\lambda + 24\delta\lambda + (1 + \delta)(-5 + 4\delta - 24\lambda)\mu)}{4(1 - 2\delta + 8\lambda)^{3/2}\sqrt{-1 + 2\delta(-1 + \mu) + 2\mu}} \right) \leq 0 \end{aligned} \quad (\text{D.1})$$

The last inequality follows in the following way: It is easy to check that the second term is always positive in region (ii) because we have assumed $\lambda - 1 \geq \delta$. Next, we bound the last term in the following way: Define $pa \equiv \frac{1}{4} + \frac{2q_g(3 + 2\delta - 4\delta^2 + 16\lambda + 24\delta\lambda + (1 + \delta)(-5 + 4\delta - 24\lambda)\mu)}{4(1 - 2\delta + 8\lambda)(3/2)\sqrt{-1 + 2\delta(-1 + \mu) + 2\mu}}$.

We will lower bound pa . Note that $\frac{\partial pa}{\partial \lambda} = \frac{-2(5 + 8\delta + 8(2 + 3\delta)\lambda)q_g + 6(1 + \delta)(3 + 8\lambda)q_g\mu}{((1 - 2\delta + 8\lambda)^{5/2}\sqrt{-1 + 2\delta(-1 + \mu) + 2\mu})}$. It is easy to check that pa is increasing in λ for all parameter values from our assumption that

$\delta + 1 \leq \lambda$. Replacing the smallest possible value of $\lambda = 1 + \delta$ and differentiating with respect to δ , we find the $\frac{\partial pa}{\partial \delta} = \frac{q_g(-4(3 + \delta(13 + 8\delta)) + 2(27 + 58\delta + 28\delta^2)\mu - 3(1 + \delta)(11 + 8\delta)\mu^2)}{6\sqrt{3}(3 + 2\delta)^{5/2}(-1 + 2\delta(-1 + \mu) + 2\mu)^{3/2}}$. It is

again easy to check that pa is decreasing in δ from the assumption that $\delta + 1 \leq \lambda$ and the condition that $\frac{5 + 6\delta + 8\lambda}{8 + 8\delta} < \mu$. Replacing the value of $\delta = \delta(0)$ and now differentiating with respect

to q_g , we get $\frac{\partial pa}{\partial q_g} = \frac{19 + 42\delta(0) + 20\delta(0)^2 - (1 + \delta(0))(29 + 20\delta(0))\mu}{6(3 + 2\delta(0))^{3/2}\sqrt{-3 + 6\delta(0)(-1 + \mu) + 6\mu}}$. It is again easy to check

that pa is always decreasing in q_g . Since $q_g \leq 1$, replacing $q_g = 1$ and finally differentiating with respect to μ , we get that $\frac{\partial pa}{\partial \mu} = -\frac{(1 + \delta(0))(-10 + 20\delta(0)^2(-1 + \mu) + 29\mu + \delta(0)(-36 + 49\mu))}{2(-1 + 2\delta(0)(-1 + \mu) + 2\mu)^{3/2}}$.

We can again check that pa is increasing for all $\mu \geq \frac{13 + 14\delta}{8 + 8\delta}$. Thus we have that the smallest

value of pa is achieved for the following vector $(\lambda, \delta, q_g, \mu) = (1 + \delta(0), \delta(0), 1, \frac{13 + 14\delta(0)}{8 + 8\delta(0)})$.

Replacing these parameter values in pa , we again have that pa is always decreasing in $\delta(0)$.

Thus taking the limit of pa as $\delta(0)$ goes to infinity, we have that the $\lim_{\delta(0) \rightarrow \infty} pa(\delta(0)) = -\frac{1}{6}$.

If we assume that $|\delta''| < 12|$ (this is trivially true in the linear case), the sum of the first and the last term in equation D.1 is always non positive. Since we have already shown that the second

term is also non positive, we have that revenue is concave in this region as well. In region (iii),

we again have that the $\frac{\partial^2 \pi^{RA}}{\partial q^2} = -2 + \frac{(1 - q_g^*)^2}{4} \delta'' \leq 0$ because of the concavity assumption of

δ .

Next we show that the function is unimodal. In order to show unimodality, we will consider different cases depending on how the revenue function behaves in each region. In each

region, the revenue function can be either increasing, decreasing or have the optimal point from the f.o.c in the region. Define $q_l \equiv \delta^{-1} \left(\frac{8 + 32\lambda - 4q_g - 32\lambda q_g + q_g^2 + 8\lambda q_g^2 - 8\mu}{-8q_g + 2q_g^2 + 8\mu} \right)$ and $q_h \equiv$

$\delta^{-1} \left(\frac{5 + 8\lambda - 8\mu}{-6 + 8\mu} \right)$. First, we start with the case when the function is always decreasing or has the

optimal solution in the first region. For both the cases we must have i.e. $1 - 2q_l + \frac{(1 - q_g^*)^2 \delta'}{4} \leq 0$ because of the concavity of the revenue function. We show that the function is always decreasing in this case. In region (ii), we have that $1 - 2q_l + \frac{(1 - q_g^*/q_d^*) \delta'}{4} \leq 1 - 2q_l + \frac{(1 - q_g^*)^2 \delta'}{4} \leq 0$. The first inequality follows from the fact that $q_d^* \leq 1$. Thus the function is decreasing in region (ii) as well. In region (iii), again we have that $1 - 2q_h + \frac{(1 - q_g^*)^2 \delta'}{4} < 1 - 2q_h + \frac{(1 - q_g^*/q_d^*) \delta'}{4} \leq 0$. The first inequality follows from the assumption that $q_d^* \geq \frac{1}{2 - q_g}$. Thus in both the cases the function is unimodal. We can similarly show that in all the cases the revenue function is unimodal. We avoid redoing the other cases here for the sake of brevity.

Proof of Proposition 5.6.1

Note that from proposition 5.3.1 we have that $q_\phi^1 = \frac{1}{2} - \frac{g}{8}$. If $b < t_1$, then the two values are equal. We only need to consider the case when $t_1 < b \leq \frac{8 + 4g + g^2}{32}$. Define $f(q_1) \equiv 1 - 2q_1 + \frac{a^2(-1 + q_1) - g(1 + 6b + \sqrt{1 + g(1 - q_1)(1 + g + 12b - gq_1)})}{9\sqrt{(-1 + g(-1 + q_1))(-1 - 12b + g(-1 + q_1))}}$. Note that since the revenue function is concave in this region, $f'(x)$ is decreasing in q_1 . This implies that in order to show that $q_{CS}^1 > \frac{1}{2} - \frac{g}{8}$, it is sufficient to show that $f(\frac{1}{2} - \frac{g}{8}) \geq 0$ for any $t_1 < b \leq \frac{8 + 4g + g^2}{32}$. Next, note that $f(\frac{1}{2} - \frac{g}{8}) = -\frac{g(32 + 16g + 4g^2 + 192b - 5\sqrt{(8 + 4g + g^2)(8 + 4g + g^2 + 96b)})}{32}$. Showing this value is greater than 0 is equivalent to checking that $16(8 + 4g + g^2 + 48b)^2 \leq 25(8 + 4g + g^2)(8 + 4g + g^2 + 96b)$. Simplifying this further gives us that this is equivalent to $b \leq \frac{8 + 4g + g^2}{32}$. This proves that $q_{CS}^1 \geq q_\phi^1$. Using a similar argument, we can again show that if $\frac{8 + 4g + g^2}{32} < b$, then $q_{CS}^1 < q_\phi^1$.

Now, we prove part (ii) of the proposition. If $b < t_1$, then from theorem 5.5.2, we have that q_1 is constant.

If $\frac{8 + 4g + g^2}{32} > b > t_1$, then q_1 is the solution to the following fixed point equation: $1 - 2q_1 + \frac{g^2(-1 + q_1) - g(1 + 6b + \sqrt{1 + g(1 - q_1)(1 + g + 12b - gq_1)})}{9\sqrt{(-1 + g(-1 + q_1))(-1 - 12b + g(-1 + q_1))}} = 0$ and $\frac{\partial q_1^*}{\partial b} = -\frac{4gb(1 + g(1 - q_1))}{(1 + g(1 - q_1))(1 + 12b + g(1 - q_1))^{3/2}} \leq 0$. Since we already have that the revenue function is concave, by implicit function, this implies that the q_1^* decreases as b increases. We can similarly show that q_1 is decreasing with b in the other regions as well.

Proof of Proposition 5.6.2

First consider PA . If $b \leq \frac{pb}{2}$, then q_{PAP}^1 is the solution to the following $1 - 2q^* + \frac{\delta'(q^*)}{4} = 0$. This is equivalent to the f.o.c for q_ϕ^1 . Since $q_\phi^1 = q_{PAP}^1$, it follows trivially that $p_{PAP}^1 = p_\phi^1$ and $CS_{PAP}^1 = CS_\phi^1$. If $b \geq \frac{pb}{2}$, then q_{PAP}^1 is the solution to the following $1 - 2q^* + (q_g^*)(1 -$

$q_g^*) \frac{\delta'(q^*)}{4} = 0$. Since $q_g \geq \frac{1}{2}$, this implies that $1 - 2q^* + (q_g^*)(1 - q_g^*)\delta'(q^*) \leq 1 - 2q^* + \frac{\delta'(q^*)}{4}$. Thus, $q_1^{PAP} \geq q_1^\phi$ which implies that $p_{PAP}^1 \leq p_\phi^1$ and $CS_{PAP}^1 \geq CS_\phi^1$.

Next consider RA . Note that the optimality of q_ϕ^1 implies that $1 - 2q_\phi^1 + \frac{\delta'}{4} = 0$. Since the function $f(q) \equiv 1 - 2q + \frac{\delta'(q)}{4} = 0$ is monotonic in q , it suffices to show that $f(q_{RAP}^1)$ is smaller than 0 in order to show that $q_{RAP}^1 > q_\phi^1$. Replacing the value of q_{RAP}^1 using the first order condition from proposition D.2.5 in all the three regions, we get that $f(q_{RAP}^1) \leq 0$. The last inequality follows because in all the three regions from proposition D.2.5, we have that the revenue contribution from the second period is less than that under the case without intervention $((1 + \delta)/4)$. Since quantity in the first period is higher, this implies price is lower and consumer surplus is higher under the randomized allocation policy. Next, note that $q_{RAP}^2 = (1 + q_g)/2$ or $q_{RAP}^2 = 1/2 + q_g(1 - 1/2q_d)$ from proposition D.2.4. Since $q_d \geq 1/(2 - q_g)$ is always true (from the constraint in the government's maximization problem), we have that the $q_{RAP}^2 \geq 1/2$.

Now we prove part (ii) of the proposition. First consider PA . Note that if $b \leq \frac{pb}{2}$, then $q_g^* \leq 1/2$ and the second period revenue of the trader is $(1 + \delta)/4$. Thus $\frac{\partial \pi^{pa}}{\partial b} = 0$. Since q_1 does not change, δ also remains the same. If $b > \frac{pb}{2}$, then $q_g^* > \frac{1}{2}$ and the second period revenue for the trader is $q_g(1 - q_g)(1 + \delta)$. Since $\delta = g(1 - q_1)$, the trader's revenue is a concave function. Note that $\frac{\partial^2 \pi^{pa}}{\partial q \partial b} = \frac{\delta'(1 - 2q_g^*)}{p_b} \geq 0$. The last inequality follows from the fact that $q_g^* > \frac{1}{2}$ and $\delta' \leq 0$. Thus first period optimal quantity increases as b increases and the anxiety decreases.

Next, consider RA . We will start by showing that if the optimal solution is an interior point in any of the four regions, then it is indeed decreasing in b . Note that $\frac{\partial q^*}{\partial b} = -\frac{(\partial^2 \pi)/\partial q \partial b}{(\partial^2 \pi)/(\partial q^2)}$ from implicit function theorem. Since the function is concave in each of the region and unimodal, we can easily check that for each of the regions, $\frac{\partial^2 \pi}{\partial q \partial b} \geq 0$. Similarly, we can check that the threshold point $t_1 \equiv \delta^{-1}\left(\frac{8 + 32\lambda - 4q_g - 32\lambda q_g + q_g^2 + 8\lambda q_g^2 - 8\mu}{-8q_g + 2q_g^2 + 8\mu}\right)$ is also increasing in b since $\frac{\partial t_1}{\partial b} = -\frac{4(2 - q_g)(-2 + 8\lambda(-1 + \mu) + 3\mu)}{\delta'(t_1)(-4q_g + q_g^2 + 4\mu)^2} \geq 0$. Since we have already shown that the revenue function is unimodal, it is easy to show that q_1^* increases as b increases. This can be shown in the following way. Let us consider the case when $q^*(b_1)$ is an interior point of one of the regions (let's assume it is in region (ii)). Consider the optimal $q^*(b_2)$ for a $b_2 > b_1$. Note that we have already proved that $t_1(b_2) > t_1(b_1)$. Let $q_1^*(b)$ ($q_2^*(b)$) be the solution to the foc in the first (second) region. There are multiple cases to consider. Let us first consider the case when $q^*(b_2) = q_1^*(b_2)$. We have already shown that $q_1^*(b_1) \leq q_1^*(b_2)$. Note that since $(1 - q_g^*) \geq (1 - \frac{q_g}{q_d})$, we also have that $q_2^*(b_1) \leq q_1^*(b_1)$. Combining the two equations, we have that $q_2^*(b_1) \leq q_1^*(b_2)$. Next, consider

the case when $q^*(b_2) = t_1(b_2)$. This implies that the revenue function is decreasing in the second region i.e. $q_2^*(b_2) < t_1(b_2)$. But we have already shown that $q_2^*(b_1) < q_2^*(b_2)$. Combining the two equations, we get $q_2^*(b_1) < t_1(b_2)$. Next, consider the case when $q^*(b_2) = q_2^*(b_2)$. In this case we have already shown that $q_2^*(b_2) > q_2^*(b_1)$. The result follows trivially when $q^*(b_2)$ is in region (iii). Thus in all the cases we have shown that optimal $q^*(b_2)$ is greater than $q^*(b_1)$. We can similarly show all other cases but avoid redoing them here for the sake of brevity.

Proof of Proposition D.2.1

Note that we have two kinds of consumers: those with and those without the subsidy. If the realized price is p_2 , then demand from non-subsidized population is $1 - \max(f, \frac{p_2}{1 + \delta})$. Similarly the demand from subsidized population is $[f - \max(\frac{p_2 - c}{1 + \delta}, 0)]^+$ since these consumers have can spend extra c . Thus the total demand is just the sum of the two demands. Finally, note that it is never optimal for government to transfer a cash subsidy of more than $1 + \delta$ because maximum price in the market without government intervention can not be higher than $1 + \delta$ (as highest budget allocation by consumers in period 2 is $1 + \delta$). Inversing the demand function to get price in terms of quantity under the above assumption gives us the desired result.

Proof of Proposition D.2.2

The revenue for the trader in the second period is given by $d(q)q$, where $d(q)$ is obtained from proposition D.2.1. In order to obtain the optimal q_2 for a given a and Δ , we have to consider different scenarios. Consider the case when $\frac{c}{1 + \delta} < \frac{3\sqrt{2}}{4} - 1$. In this case, note that $\frac{c}{1 + \delta} \leq \frac{1}{3} - \frac{c}{1 + \delta} \leq \frac{1}{3} - \frac{c}{3(1 + \delta)} \leq \frac{1}{2} - \frac{c}{2(1 + \delta)}$. Let us first consider the scenario when $a \leq \frac{1}{3} - \frac{\Delta}{1 + \delta}$. It is easy to check that the revenue function is always decreasing in region (ii) and (iii) of proposition D.2.1 and the optimal solution of the first region, $\frac{1}{2}$, is always less than $1 - f - \frac{c}{1 + \delta}$ here. Thus $q_2^* = \frac{1}{2}$ in this region. Next consider the case when $f \in (\frac{1}{3} - \frac{c}{1 + \delta}, \frac{1}{3} - \frac{c}{3(1 + \delta)})$. In this case we have a bimodal function with modes occurring at $\frac{1}{2}$ and at $\frac{1 + f + \frac{c}{1 + \delta}}{2}$. It is easy to compare the revenue at these two points and identify that revenue at $\frac{1}{2}$ is again optimal in this region. In the next region when $f \in (\frac{1}{3} - \frac{c}{3(1 + \delta)}, \frac{1}{2} - \frac{\delta}{1 + \delta})$, we again have a bimodal function with modes at $\frac{1}{2}$ and $1 - f + \frac{c}{1 + \delta}$. Comparing the revenue at two points, we have that if $f < \frac{1}{2} + \frac{c}{2} - \frac{\sqrt{2c(1 + \delta) + c^2}}{2(1 + \delta)}$ then the optimal solution is at $\frac{1}{2}$ and it is $1 - f + \frac{c}{1 + \delta}$ otherwise. Finally, if $f \geq \frac{1 + \frac{c}{1 + \delta}}{2}$, then the revenue function is always

increasing in the first two regions and the optimal solution is thus $\frac{1 + \frac{c}{1 + \delta}}{2}$. Combining the results gives us the first part of the proposition. The results for other parts follow similarly and we avoid redoing them here for the sake of brevity.

Proof of Lemma D.2.3

Note that the first mode of the revenue curve is the same as the no intervention case. Thus t_1 can be obtained by finding the point at which the revenue from the second mode becomes equal to the total revenue from the optimal solution in the no intervention case. Further, the second mode is just the point at which the revenue from the second part of the trader's revenue is maximized. The next part follows by simplifying the condition that $\delta^{-1}(4b - 1) \leq 0$ and using the fact that $\delta(q_1) = g(1 - \frac{q_1}{q_{max}})$.

Proof of Proposition D.2.4

Note that if the government decides to prioritize and sell q_d units at a price of p_g , then the demand for people with budget in the range of $(1 + \delta)(1 - (q_d - q_g))$ and $(1 + \delta)(1 - q_d)$ is served by the government. If the trader sells less than $q_d - q_g$ then there is no change in the demand function and the realized price is the same as original curve. If he sells more than $q_d - q_g$ units, then $q - (q_d - q_g)$ units are going to be sold to people with budget in the range of $(1 + \delta)(1 - (q_d))$ and $(1 + \delta)(1 - (q - q_d + q_g))$ since q_g units have already been sold by the government. This shows the second part of the proposition. We can similarly find the demand function under RA . If the trader charges p_2 then all the consumers who have the budget, $b_c(1 + \delta) \geq p_2$ and have not bought from the government will buy from the trader. This implies that $q_2 = (1 - \frac{p_2}{1 + \delta})(1 - \frac{q_g}{q_d})$ if $p_2 > p_g$ and $q_2 = (1 - q_g - \frac{p_w}{1 + \delta})$. Rewriting the equations in terms of p_2 gives us the desired result.

Proof of Proposition D.2.5

The optimal solution for the trader in the first period can be obtained by comparing the optimal solutions in the two regions. Note that maximum revenue in the first region is obtained at $\min(q_d - q_g, 1/2)$ and in the second region is obtained at $\max(q_d - q_g, (1 - q_g)/2)$. Comparing the revenue at each of these optimal points gives us the desired result for PA . We can similarly obtain the result in the proposition for RA .

Proof of Proposition D.2.6 Let us first consider PA and RA . Consumer surplus in the second period is more than that under no intervention by construction since the government maximizes consumer surplus in the second period. Further, we have already shown that the

total quantity released by the trader in the first period can only increase under PA or RA for any value of b . Since consumer surplus in the first period increases as quantity released in the first period increases, total consumer surplus in the two period also increases under both PA and RA .

Next consider CS . First note that if $b \leq \frac{8 + 4g + g^2}{32}$, $q_{CS}^1 \geq q_\phi^1$. As a result, in this region total surplus under CS is again more than that under no intervention. If $b \geq \frac{8 + 4g + g^2}{32}$, then $c^*(q_1) = \sqrt{b(1 + g(1 - q_1^*))}$ and $q_2^* = 1$. Note that the maximum loss in the first period due to CS is upper bounded by the difference in $CW(q_1 = 1/2 - g/8) - CW(q_1 = 1/2 - g/2) = \frac{3g(32\lambda - 5g + 8)}{128}$. However, note that the gain in surplus in the second period is lower bounded by λ since the $q_\phi^2 = 1/2$ and $q_{CS}^2 = 1$ in this range. Finally, since $\lambda > 1 + g$ by construction, it is easy to check that $\lambda > \frac{3g(32\lambda - 5g + 8)}{128}$. This proves the result for CS .

Proof of Proposition D.2.7 First consider the case of CS . It is easy to check that the trader's revenue in the second period under CS for a given value of q_1 is equal to (greater than) the second period revenue under no intervention if $b \leq (>)t_1$. The total revenue under CS is $R_{CS}^{t*} \geq R_{CS}^t = q_1(1 - q_1) + R^2(q_1, q_2^*) \geq q_1(1 - q_1) + (1 + g(1 - q_1))/4 = R_\phi^t$ where $R^2(q_1, q_2^*)$ is the second period revenue. Since the total revenue for any q_1 is more under CS , the optimal revenue under CS is also more than that under no intervention.

Next consider PA . If $b \leq p_b/2$ then the trader revenue is same as that under no intervention. If $b \geq p_b/2$, then the total revenue is $(1 - gq_g(1 - q_g))(1 + gq_g(1 - q_g))/4 + (1 + g(1 - (1 - gq_g(1 - q_g))/2))q_g(1 - q_g)$. It is easy to check that the revenue is indeed decreasing in b . The results for RA follow similarly since the demand function (and hence, revenue) of the trader is smaller than that under no intervention for all values of q_1 and q_2 . We avoid redoing them here for the sake of brevity.

Proof of Proposition D.3.1 First consider Period 2 revenue for a given value of q_1 . In this case, Period 2 revenue, $p_2(e^{-\frac{\lambda p_2}{1 + g(1 - q_1)}})$ is maximized at $p_2^* = \frac{1 + g(1 - q_1)}{\lambda}$. Replacing this value of p_2^* in total revenue across the two periods, we get: $\frac{1 + g(1 - e^{-\lambda p_1})}{\lambda}e^{-1} + e^{-\lambda p_1}p_1$. Differentiating w.r.t. p_1^* and using the foc gives us the desired result. Finally, it is easy to check that $q_1^* \leq q_2^*$ and $p_1^* \leq p_2^*$.

Proof of Lemma D.3.2 Replacing the values of $q^*(0)_{1,e}$ and $q^*(g)_{1,e}$ for both distributions, we get $S_U = g/4$ and $S_E = 1 - e^{-g/e}$. It is easy to check that $S_E \geq S_U$ when $g \in [0, 1]$.

Proof of Proposition D.3.3 Since $\mu = 0$, it is optimal for the government to buy as

many units as possible ($q_g = \frac{b}{p_b}$) and sell it at minimum price $q_d = 1$. First consider *PA*. If $b \leq p_b(1 - q_{2,e}^*)$, then $q_g \leq 1 - q_{2,e}^*$. Since the trader serves consumers with higher budget first, this implies that none of the consumers who buy from the trader in Period 2 are served by the government at equilibrium. Since the trader's total revenue is unaffected by the government intervention, it continues to sell $q_{1,e}^*$ in Period 1. In contrast, if $q_g > 1 - q_{2,e}^*$, then the trader sells $1 - q - g$ units. Since the revenue in the second period is now smaller than the second period revenue under no intervention, $q_{PA,e}^*$ increases. Next consider *RA*. In this case, all consumers who have not bought from the government will buy from the trader. This implies the $q_2(p_2) = (1 - q_g)e^{-\lambda p_2/(1+g(1-q_1))}$ and optimal revenue in the second period is again $(1 - q_g)e^{-1}(1 + g(1 - e^{-p_1\lambda}))/\lambda$. Replacing this as the second period revenue, and solving for optimal p_1 , we get that $p_{1,e}^* = 1/\lambda + (g(1 - q_g)/\lambda)e$. Since $p_{1,e}^*$ decreases as q_g increases, we have that $q_{1,RA}^* \geq q_{1,\phi}^*$ for all $b \geq 0$.

Proof of Proposition D.3.4 Total demand in period 1 when the wholesaler charges p is $\int_{p/n}^{\min(p,1)} \frac{b}{p} db + \int_p^1 db = 1 - \frac{(1+n^2)p}{2n^2}$ if $p \leq 1$ and it is $\frac{1}{2p} - \frac{p}{2n^2}$ if $p > 1$. Similarly, total demand in period 2 when wholesaler charges p_2 is $1 - \frac{(1+n^2)p}{2n^2(1+\delta)}$ if $p \leq 1 + \delta$ and it is $\frac{1+\delta}{2p} - \frac{p}{2n^2}$ if $p_2 > 1 + \delta$. It is easy to check that the revenue function is always decreasing when $p_2 > 1 + \delta$ and concave when $p_2 \leq 1 + \delta$. It is maximized at $p_2^* = \frac{(1+\delta)n^2}{1+n^2}$ for all values of p_1 . Replacing p_2^* and $\delta(p_1) = g(1 - q_1(p_1))$, we get that the revenue function $(1+g(\frac{(1+n^2)p_1}{2n^2}))\frac{n^2}{2(1+n^2)} + p_1(1 - \frac{(1+n^2)p_1}{2n^2})$ is concave if $p_1 \geq 1$ and it is always decreasing when $p_1 > 1$. The result in the proposition follows from the first order condition. Finally, it is easy to check that $q_1^* \leq q_2^*$ and $p_1^* < p_2^*$ in both the cases.

Proof of Proposition D.3.5 In order to prove the result, we start by characterizing the demand function of the trader in period 2 *after* the government's intervention. Let us consider *PA* first.

Recall that the government intervenes by selling q_g units and prioritizes low-income consumers. First, consider the case when $\frac{(n^2 - 1)p_g}{2(1+\delta)n^2} < q_g$. In this case, consumers with budget in the range $[\frac{p_g}{n(1+\delta)}, p_g]$ will buy $\frac{b(1+\delta)}{p_g}$ units and those with budgets in the range $[p_g, \frac{(1+n^2)p_g}{2(1+\delta)n^2} + q_g]$ will buy 1 unit from the government. We will now characterize the demand as a function of p_w , the price charged by the trader. There are 6 regions to consider. First, if $p_w \leq \frac{p_g}{n}$, then $q_w = \int_{p_w/n(1+\delta)}^{p_w/(1+\delta)} x \frac{(1+\delta)}{p_w} dx + \int_{p_w/(1+\delta)}^{p_g/n(1+\delta)} dx + \int_{(1+n^2)p_g/((2(1+\delta)n^2))+q_g}^1 dx =$

$1 - q_g - \frac{(-1+n)^2 p_g}{2(1+\delta)n^2} - \frac{(1+n^2)p_w}{2(1+\delta)n^2}$. If $p_w < p_g < np_w$, $q_w = 1 - q_g - \frac{(1+n^2)p_g}{2(1+\delta)n^2} - \frac{p_w}{2(1+\delta)n^2} + \frac{p_g^2}{2n^2 p_w (1+\delta)}$. If $\frac{(1+n^2)p_g}{2(1+\delta)n^2} + q_g < p_w < p_g$, then $q_w = 1 - \frac{1+2n^2}{2(1+\delta)n^2} p_w + \frac{p_g^2}{2(1+\delta)n^2 p_w}$. If $p_g < p_w < \frac{(1+n^2)p_g}{2(1+\delta)n^2} + q_g$, then $q_w = 1 - q_g - \frac{(1+n^2)p_g}{2(1+\delta)n^2}$. If $\frac{p_w}{n} < \frac{(1+n^2)p_g}{2(1+\delta)n^2} + q_g < p_w$, then $q_w = 1 - \frac{p_w}{2(1+\delta)} + \frac{((1+n^2)p_g + 2(1+\delta)n^2 q_g)^2}{8(1+\delta)n^4 p_w}$. Finally, if $\frac{(1+n^2)p_g}{2(1+\delta)n^2} + q_g < \frac{p_w}{n}$, then $q_w = 1 - \frac{(1+n^2)p_w}{2(1+\delta)n^2}$. Recall that the price-demand curve for the trader is again unaffected if p_w is larger than a threshold as none of the consumers who will buy from the government in this case can buy from the trader. This insight is similar to the insight obtained from the analysis of *PA*. Given how the policy affects trader's demand curve remains the same as in the original analysis, with similar algebra, we can show the result in Proposition 5.6.2. We avoid redoing the analysis for the sake of brevity here but highlight the key mechanism identified in Proposition 5.6.2. In particular, consumers who buy from the government when its budget is small would have not bought from the trader in equilibrium anyways. Since there is no revenue loss for the trader due to the government's intervention, it is optimal for the trader to create the same level of shortage in Period 1 as he would do under no intervention. Thus, when the government's budget is small relative to the import price (p_b), then the intervention is not effective in reducing artificial shortage in the market and $q_{PA}^1 = q_\phi^1$.

Next consider *RA*. We again assume that if the government has q_g units and the total number of consumers who can afford the subsidized price, referred as the potential demand, is $q_d \equiv 1 - \frac{(1+n^2)p_w}{2n^2(1+\delta)}$, then each consumer will be able to buy from the government with probability $\min(1, \frac{q_g}{q_d})$. If $p_w < \frac{p_g}{n}$, then $q_w = 1 - \frac{(1+n^2)}{2(1+\delta)n^2} p_w + \frac{2nq_g(p_g - (1+\delta)n)}{2(1+\delta)n^2}$. If $\frac{p_g}{n} < p_w < p_g$, then $q_w = 1 - \frac{q_g}{q_d} + \frac{p_g^2 q_g}{2n^2 q_d (1+\delta) p_w} - p_w \frac{q_d(1+n^2) + q_g n^2}{2(1+\delta)n^2 q_d}$. Finally, if $p_g < p_w$, $q_w = (1 - \frac{q_g}{q_d})(1 - \frac{(n^2+1)p_w}{2(1+\delta)n^2})$. In contrast to *PA*, note that the price-demand curve under *RA* always shifts downward regardless of the p_w . This insight is again consistent with the key mechanism from the analysis of *RA* in proposition 5.6.2 and confirms the result in this case as well. In particular, *RA* always affects the demand curve regardless of the government budget and ensures that the *RA* policy always decreases artificial shortage in the market.

Proof of Proposition D.3.6 The trader's revenue function is $\pi(q_1, q_2) = q_1(1 - q_1) + (1 + g(1 - q_1))q_2(1 - q_2) - h(Q - q_1)$. First consider the case when the inventory constraint is not tight and Q is larger than the optimal quantity. In this case, $q_2^* = 1/2$. Replacing this value in the objective function and solving for optimal q_1^* gives us the desired result in part (i). Next,

consider the case when the constraint is indeed tight. In this case, quantity sold in period 2 is just $Q - q_1^*$. Replacing this value in the objective function and solving for q_1^* gives us the desired result. Finally, if $h < \frac{g}{4}$, then it is easy to check that in both the cases, $q_2^* > q_1^*$ and $p_2^* > p_1^*$.

Proof of Proposition D.3.7 First consider PA . In this case, note that again government's intervention affects only lower income (higher and lower income) consumers if the budget is small (large). Similar to proposition D.2.4, this implies that either period 2 price is $(1 - q_g - q)(1 + \delta)$ or it is $(1 - q)(1 + \delta)$ for the trader. We will show that under both the cases, $q_{PA,c}^1 \geq q_{\phi,c}^1$. First, if price in period 2 is $(1 - q)(1 + \delta)$, then the total revenue of the trader is the same as that under no intervention. Thus the optimal solution should indeed be the same. Next, if price in period 2 is $(1 - q_g - q)(1 + \delta)$, then it is again easy to check that the revenue function of the trader is concave in q_1 . Thus we can use the first order condition to identify the optimal solution in this case, $q_{PA,c}^1 = \frac{2 + g(2Q + q_g) - \sqrt{4 + g(2Q + q_g) + g^2(3 + Q^2 + Q(-3 + q_g) - 3q_g + q_g^2)}}{3g}$. It is easy to check that $q_{PA,c}^1 \geq q_{\phi,c}^1$. Thus in both the cases, we have that quantity sold by the trader under PA is more than that under no intervention. The proof for RA follows similarly and we avoid redoing it here for the sake of brevity.

Proof of Proposition D.3.8 It is easy to check that the total revenue function for the trader, $q_1(1 - q_1) + q_2(1 - q_2)(1 + g(1 - q_1)(1 - q_2))$ is again concave in q_1 and q_2 . Thus, setting the first order conditions to 0 gives us the desired result. Next, let $f(q_2) \equiv 1 - 2q_2^* + \frac{g(1 + gq_2(1 - q_2)^2)(1 - 4q_2 + 3q_2^2)}{2} = 0$. Now $f(q_2)$ is monotonically decreasing in $q_2 \in (0, 1)$ and $f(0) \geq 0$. This implies that $f(q_2) \leq 0$ for all $q_2 > q_2^*$. Checking $f(1/2)$, we get that $f(1/2)$ is indeed less than 0. This proves that $q_2^* \leq 1/2$. Finally since $p(q_2) \geq 1/2$ for all $q_2 \leq \frac{1}{2}$, we have that $p_2(q_2^*) \geq \frac{1}{2}$.

Proof of Proposition D.3.9 First, we show that the result for the first part. In order to prove this result, we need to first characterize the government's optimal policy. Note that the consumer surplus here is exactly similar as in §5.5.2, except that the price in period 2 is now convex instead of being linear in q_2 . Thus,

$$\begin{aligned}
 CS^{PA}(q_g, q_d) &\equiv \int_0^{\min(q_w, q_d - q_g)} (1 - q) dq + \int_{q_d}^{\max(q_d, q + q_g)} (1 - q) dq - (1 - q_w)(1 + \delta(1 - q_w)q_w \\
 &+ \int_{q_d - q_g}^{q_d} (1 - q) dq - -(1 - q_d)(1 + \delta(1 - q_d)q_d) + \lambda(2(q_w + q_g) - 1)
 \end{aligned} \tag{D.2}$$

Note that the government's objective is again $\pi_g = CS^{PA}(q_g, q_d) + \mu q g p_g$. We will show

that under the assumption that $\mu \leq \frac{1}{2}$, π_g is always increasing in q_g for all values of q_d . Let us denote the optimal solution for the trader in the second period by q_w^* . Like in §5.5.2, we restrict q_d in a range such that the trader's price, q_w^* is always greater than q_d . Let us denote that lower threshold by q_d^l . Note that the trader can again have two optimal solutions, if $q_d - q_g \leq q_2^*$, then $q_w^* = q_d - q_g$ and if $q_2^* > q_d - q_g$, then $q_w^* = q_d$. We will show that in both these cases $\frac{\partial \pi_g}{\partial q_g} \geq 0$. If $q_d - q_g \leq q_2^*$, then $\frac{\partial \pi_g}{\partial q_g} = 2 + \delta(4 - 6(q_d - q_g)) > 2(1 - \delta) \geq 0$. The last inequality follows because we have assumed that $g < 1$. Next, consider the case when $q_d - q_g > q_2^*$. In this case, $\frac{\partial \pi_g}{\partial q_g} = 2\lambda + q_g - \delta(1 - q_d)^2(1 - \mu) + \mu(1 - q_d) \geq (1 - q_d)(\mu - (1 - \mu)(1 - q_d)\delta) \geq 0$. The last inequality follows from the assumption that $\mu < \frac{1}{2}$. Thus, $q_g^* = \frac{b}{p_b}$ in this case as well. Replacing this value and integrating w.r.t q_d , it is easy to check that the derivative is always increasing in q_d under the assumption that $\mu \leq \frac{1}{2}$. Thus, $q_d^* = 1$ and $q_g^* = \frac{b}{p_b}$. Now that we have characterized the optimal government solution, it is easy to see that if $1 - \frac{b}{p_b} \leq q_2^*$, then the first period optimal solution of the trader is indeed affected by the government's intervention. This is because similar to theorem 5.6.2, some of the revenue in the first period is again captured by the government and trader's incentive to create shortage in the first period decreases. In contrast if $1 - \frac{b}{p_b} > q_2^*$, then trader's optimal solution remains unchanged and we again get that the optimal shortage in the market is unchanged.

Note that under *RA*, the demand curve for the trader moves down at all points because of government's intervention. Thus, the incentive for the trader to create shortage decreases regardless of government's budget. We avoid redoing the proof here because it follows exactly as Theorem 5.6.2.

Proof of Proposition D.3.10 Let us first consider *PA*. We refer to q_g as Q throughout the proof. Let q_t^1 and q_t^2 be the optimal quantity released by the trader in the two periods respectively. Following the same logic as in Theorem 5.5.4, we can characterize the optimal trader strategy in the two periods. If $q_g^2 \leq \frac{1}{2}$, then $q_t^1 = \min(\frac{1}{2} - \frac{g}{8}, 1 - q_g^1)$ and $q_t^2 = 1/2$. If $q_g^2 > \frac{1}{2}$: $q_g^1 = \min(\frac{1 - gq_g^2(1 - q_g^2)}{2}, 1 - q_g^1)$ and $q_t^2 = 1 - q_g^2$. We need to consider different ranges of Q separately to identify the optimal solution for the government. First consider the case when $Q \leq \frac{1}{2}$. In this case, $q_g^2 = Q - q_g^1 \leq \frac{1}{2}$ and $q_g^1 \leq \frac{1}{2} + \frac{g}{8}$ for all $Q \leq 1/2$. Thus, $q_t^2 = 1/2$ and $q_t^1 = 1/2 - g/8$. Note that the consumption utility is constant and equal to $\lambda(2(Q + 1 - g/8) - 1)$

in this range. The surplus is,

$$\begin{aligned}
 CW(q_g^1, q_g^2) &= \int_0^{1/2} (1-x)dx + \int_{1-q_g^2}^1 (1-x)dx - \frac{1+g(1-(1/2-g/8+q_g^1))}{4} \\
 &+ \int_0^{1/2-g/8} (1-x)dx + \int_{1-q_g^1}^1 (1-x)dx - (1/2-g/8)(1/2+g/8) \quad (D.3) \\
 &= \frac{(32-3g(8+g)+32gq_g^1+64(q_g^1)^2+64(q_g^2)^2)}{128}
 \end{aligned}$$

Replacing $q_g^2 = Q - q_g^1$, it is easy to check that the surplus is always maximized when $q_g^1 = Q$. Next, consider the case when $Q \in [1/2, 1]$. If $q_g^1 \in [Q-1/2, Q]$, then the CW is the same as in the previous case and again maximized when $q_g^1 = Q$. If $q_g^1 \in [0, Q-1/2]$, then $q_t^1 = \frac{1-gq_g^2(1-q_g^2)}{2}$ and $q_t^2 = 1 - q_g^2$. In this region, the optimal value can be calculated by comparing the end points and the optimal value is obtained when $q_g^1 = Q - 1/2$. Finally comparing the welfare at $q_g^1 = Q - 1/2$ and $q_g^1 = Q$ gives us the desired result. Thus, we have so far shown that intervening preemptively, i.e., selling all the inventory in the first period when $Q \in [0, 1]$ is optimal for the government. Next consider the case when $Q \in [1, 1+g/8]$. First note that when $q_g^1 \in [Q-1, Q-1/2]$, $q_t^1 = \frac{1-gq_g^2(1-q_g^2)}{2}$ and $q_t^2 = 1 - q_g^2$; when $q_g^1 \in [Q-1/2, 1]$, $q_t^1 = \min\{1/2-g/8, 1-q_g^1\}$ and $q_t^2 = 1/2$. When $q_g^1 \leq 1/2+g/8$ and $q_g^2 \leq 1/2$, we have already shown that CW is increasing in q_g^1 . It is easy to show that if $q_g^1 > 1/2+g/8$ and $q_g^2 \leq 1/2$, then CW is again increasing in q_g^1 since it is,

$$\begin{aligned}
 CW(q_g^1, q_g^2) &= \int_0^{1/2} (1-x)dx + \int_{1-q_g^2}^1 (1-x)dx - \frac{1}{4} + \int_0^1 (1-x)dx - (1-q_g^1)(q_g^1) \\
 &= \frac{5}{8} + (q_g^1)^2 - q_g^1 + \frac{(q_g^2)^2}{2} \quad (D.4)
 \end{aligned}$$

In this range again, we can compare the CW at the end points and confirm that CW is maximized when inventory released in the first period is maximized. The results for the other range follows similarly, except that in these cases, the government can always satisfy *all* the consumers in the market by ensuring that $q_t^2 = 1 - q_g^2$ and $q_g^1 = 1 - q_g^1$ which is optimal.

Next consider RA . In this case, $q_t^1 = (\frac{1}{2} - \frac{g(1-q_g^2)}{8})(1-q_g^1)$ and $q_t^2 = \frac{1-q_g^2}{2}$. It is easy to check that the total units sold (and equivalently, the consumption utility) is maximized at the two end points i.e. either $q_g^1 = \min\{Q, 1\}$ or $q_g^2 = \min\{Q, 1\}$. Next we will show that the total savings are maximized when $q_g^1 = \min\{Q, 1\}$. Recall that the consumer surplus under RA is

defined in the following manner:

$$\begin{aligned}
 CW^{RA}(q_2, q_g) &= \int_0^{q_t^2/(1-q_g^2)} (1-q) dq + q_g^2 \int_{q_t^2/(1-q_g^2)}^1 (1-q) dq - \left(1 - \frac{q_t^2}{1-q_g^2}\right) (1 + g(1 - q_g^1 + q_t^1)) q_t^2 \\
 &+ \int_0^{q_t^1/(1-q_g^1)} (1-q) dq + q_g^1 \int_{q_t^1/(1-q_g^1)}^1 (1-q) dq - \left(1 - \frac{q_t^1}{1-q_g^1}\right) q_t^1 \\
 &= \int_0^{1/2} (1-q) dq + q_g^2 \int_{1/2}^1 (1-q) dq - \left(\frac{1-q_g^2}{4}\right) (1 + g(1 - (q_g^1 + (\frac{1}{2} - \frac{g(1-q_g^2)}{8}))(1 - q_g^1))) \\
 &+ \int_0^{(1/2 - g(1-q_g^2)/8)} (1-q) dq + q_g^1 \int_{(1/2 - g(1-q_g^2)/8)}^1 (1-q) dq \\
 &- \left(\frac{1}{2} + \frac{g(1-q_g^2)}{8}\right) \left(\frac{1}{2} + \frac{g(1-q_g^2)}{8}\right) (1 - q_g^1) \\
 &= \frac{3}{128} g(-1 + q_g^1)(-8 + g(-1 + q_g^2))(-1 + q_g^2) + \frac{(2 + 3q_g^1 + q_g^2)}{8}
 \end{aligned} \tag{D.5}$$

The function above is concave in the region and since we are maximizing the surplus, we only need to consider the end points. Comparing the two end points, when $q_g^2 = Q$, $CW^{RA} = \frac{32 - 3g(-8 + g(-1 + q_g^2))(-1 + q_g^2) + 48q_g^2}{128}$ and when $q_g^1 = Q$, $CW^{RA} = \frac{(32 + 3g(8 + g)(-1 + q_g^2) + 48q_g^2)}{128}$. Finally, comparing the the surplus is maximized when $q_g^2 = Q$.

Proof of Proposition D.3.11 We start by proving that q_1^* is decreasing in g under all government interventions. Let us consider CS first. We can replace g with g' in Theorem 5.5.2 in order to obtain q_1^* . In order to show the result, we will show that $\frac{\partial q_1^*}{\partial g} \leq 0$ in all the four regions. First, it is easy to check that the boundary points that separate the four regions are increasing in g . In the fourth region, $\frac{\partial q_1^*}{\partial g} = -1/2 \leq 0$. In the third region, $\frac{\partial^2 \pi}{\partial g \partial q_1} = \frac{-(2 + g(1 - q_1))\sqrt{b(1 + g - gq_1)}}{4(1 + g - gq_1)^2} \leq 0$. Since the revenue function is concave, by implicit function theorem, we have that $\frac{\partial q_1^*}{\partial g} \leq 0$ in this region as well. We can similarly show that in both the other regions as well q_1^* is decreasing in g . Under PA again, from Theorem 5.5.4, we have the optimal solutions. If $q_g^* \leq 1/2$, $\frac{\partial q_1^*}{\partial g} = -1/8 \leq 0$. (b) If $q_g^* > 1/2$, $\frac{\partial q_1^*}{\partial g} = \frac{q_g^*(1 - q_g^*)}{2} \leq 0$. Finally, consider the optimal solution under RA . If $\mu \leq (5 + 6\delta + 8\lambda)/(8(1 + \delta))$ then $q_d^* = 1$ and $\frac{\partial^2 \pi}{\partial g \partial q_1} = -q_g^*(1 - q_g^*) \leq 0$. We can similarly show the result for the other two regions in RA as well.

Next we prove part (i) of the proposition. Note that the minimum q_1^* under CS is $\frac{1 - g'}{2}$ from Theorem 5.5.2. Under no intervention q_ϕ^1 is $\frac{1}{2} - \frac{g}{8}$. If $g' < g/4$, $\frac{1 - g'}{2} \geq \frac{1}{2} - \frac{g}{8}$ and this proves the first part. Next, consider PA . Following, Theorem 5.5.4, we know that for any b ,

$q_1^{PA} \geq \frac{1}{2} - \frac{g'}{8} \geq \frac{1}{2} - \frac{g}{8} = q_1^\phi$ since $g' \leq g$. Further, replacing g with g' , and using the results shown in Theorem 5.5.4, we have that q_1^{PA} is constant if b is less than a threshold and increasing in b otherwise. The results for RA follow in a similar way and we avoid redoing them here for the sake of brevity.

Proof of Lemma D.3.12

In the first period, there is no reference price effect and all consumers with budget greater than period 1 price buy from the trader. Since budget is uniformly distributed, it is easy to see that $q_1 = 1 - p_1$. Next, let us consider the case when $p_2 > p_1$. In this case, consumers feel a loss because prices are higher than their reference point. Consumers buy if $b_c - \gamma_l(p_2 - p_1) \geq p_2$. This simplifies to $q_2 = 1 - (1 + \gamma_l)p_2 - \gamma_l p_1$. Rearranging this in terms of p_2 gives us the desired result. We can rewrite the condition that $p_2 > p_1$ in terms of q_1 and q_2 as $\frac{1 - q_2 + \gamma_l(1 - q_1)}{1 + \gamma_l} \geq (1 - q_1)$. This easily simplifies to $q_1 \geq q_2$. We can similarly consider the case when $p_2 \leq p_1$. Further, it is easy to check that the optimal solution in this case is $p_1^* = \frac{2 + 3\gamma_g}{4 + 4\gamma_g - \gamma_g^2}$ and $p_2^* = \frac{2 + 2\gamma_g}{4 + 4\gamma_g - \gamma_g^2}$. Thus, $p_1^* > p_2^*$.

Bibliography

- Ababa, Addis. 2017. Ethiopia's state-of-the-art commodity exchange. February 2, *The Economist*, available at <https://www.economist.com/finance-and-economics/2017/02/02/ethiopias-state-of-the-art-commodity-exchange>.
- Acharya, Shabd Swaroop. 2004. *Agricultural Marketing In India, 4/E*. Oxford and IBH Publishing.
- Adam, Marc TP, Jan Krämer, Marius B Müller. 2015. Auction fever! how time pressure and social competition affect bidders' arousal and bids in retail auctions. *Journal of Retailing* **91**(3) 468–485.
- Aggarwal, Nidhi, Sargam Jain, Sudha Narayanan, et al. 2017. The long road to transformation of agricultural markets in India: Lessons from Karnataka. *Economic and Political Weekly* **52**(41) 47–55.
- Aker, Jenny C. 2010. Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics* **2**(3) 46–59.
- Akkaya, Duygu, Kostas Bimpikis. 2016. Agricultural supply chains under government interventions. Working paper, Stanford University.
- Alizamir, S., F. Iravani, H. Mamani. 2018. An analysis of price vs. revenue protection: Government subsidies in the agriculture industry. *Management Science*, forthcoming.
- Alizamir, Saed, Foad Iravani, Hamed Mamani. 2019. An analysis of price vs. revenue protection: Government subsidies in the agriculture industry. *Management Science* **65**(1) 32–49.
- Angrist, Joshua D, Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton university press, Princeton, NJ.
- Antle, J. 1987. Econometric estimation of producers' risk attitudes. *American Journal of Agricultural Economics* **69**(3) 509–522.
- Ariely, Dan, Itamar Simonson. 2003. Buying, bidding, playing, or competing? value assessment and decision dynamics in online auctions. *Journal of Consumer psychology* **13**(1-2) 113–123.
- Atasu, A. (ed.). 2016. *Environmentally Responsible Supply Chains*. Springer, New York, NY.
- Athey, Susan, Jonathan Levin, Enrique Seira. 2011. Comparing open and sealed bid auctions: Evidence from timber auctions. *The Quarterly Journal of Economics* **126**(1) 207–257.

- Autor, David H. 2003. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics* **21**(1) 1–42.
- Babich, V. 2010. Independence of capacity ordering and financial subsidies to risky suppliers. *Manufacturing & Service Operations Management* **12**(4) 583–607.
- Babich, V., C. Tang. 2012. Managing opportunistic supplier product adulteration: Deferred payments, inspection, and combined mechanisms. *Manufacturing & Service Operations Management* **14**(2) 301–314.
- Barrett, Christopher B. 2002. Food security and food assistance programs. B. Gardner, G. Rausser, eds., *Handbook of Agricultural Economics*, vol. 2, chap. 40. Elsevier Science B.V., 2103–2190.
- Bennett, A., B. Dugdill, J.C. Lambert, T. Thapa. 2001. Report of the FAO e-mail conference on small-scale milk collection and processing in developing countries. Food & Agriculture Organization of the United Nations, Rome, Italy.
- Bergquist, Lauren Falcao. 2019. Pass-through, competition, and entry in agricultural markets: Experimental evidence from kenya .
- Bertrand, Marianne, Esther Duflo, Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* **119**(1) 249–275.
- Bhatnagar, Gaurav Vivek. 2011. Government subsidy for onions. January 8, *The Hindu*, available at <https://tinyurl.com/ybb2ae6a>.
- Bouchery, Y., C.J. Corbett, J.C. Fransoo, T. Tan (eds.). 2017. *Sustainable Supply Chains: A Research-Based Textbook on Operations and Strategy*. Springer, New York, NY.
- Boyabati, Onur, Javad Nasiry, Yangfang Zhou. 2019. Crop planning in sustainable agriculture: Dynamic farmland allocation in the presence of crop rotation benefits. *Management Science*, forthcoming.
- Business Line. 2019. Just 14% of farmers registered on eNAM platform. July 10, available at <https://tinyurl.com/y4ypbsks>.
- Business Standard. 2015. States start selling pulses at lower rates through their outlets. April 26, *Business Standard*, available at <https://tinyurl.com/ya98zyxz>.
- Caro, F., P. Chintapalli, K. Rajaram, C. Tang. 2018. Improving supplier compliance through joint and shared audits with collective penalty. *Manufacturing & Service Operations Management* **20**(2) 363–380.
- Casaburi, Lorenzo, Rachel Glennerster, Tavneet Suri. 2013. Rural roads and intermediated trade: Regression discontinuity evidence from sierra leone .
- CCTV. 2012. Report on “instant chickens”: Grow 5 pounds in 40 days with illegal drugs. December 18 (in Chinese), available at <http://finance.sina.com.cn/nongye/hyxx/20121218/112114038687.shtml>.

- CFDA. 2018. Implementation details on food safety supervision and food sampling. China Food and Drug Administration, notice no. 14, available at http://www.gov.cn/xinwen/2018-01/24/content_5260165.htm (accessed May 30, 2018).
- Chaki, Deborshi. 2013. Why onions bring tears to indian eyes. April 26, *Aljazeera*, available at <https://www.aljazeera.com/indepth/2013/10/why-onions-bring-tears-indian-eyes-201310301335083469.html>.
- Che, Y.K. 1993. Design competition through multidimensional auctions. *The Rand Journal of Economics* **24**(4) 668.
- Cheema, Amar, Dilip Soman. 2006. Malleable mental accounting: The effect of flexibility on the justification of attractive spending and consumption decisions. *Journal of Consumer Psychology* **16**(1) 33–44.
- Chen, C., J. Zhang, T. Delaurentis. 2014. Quality control in food supply chain management: An analytical model and case study of the adulterated milk incident in China. *International Journal of Production Economics* **152**(C) 188–199.
- Chen, L., H. Lee. 2017. Sourcing under supplier responsibility risk: The effects of certification, audit, and contingency payment. *Management Science* **63**(9) 2795–2812.
- Chintapalli, P., C. Tang. 2017. The impact of crop minimum support prices on crop selection and farmer welfare in the presence of strategic farmers. Working paper, UCLA Anderson School of Management, Los Angeles, CA.
- Cho, S.H., X. Fang, S. Tayur. 2015. Combating strategic counterfeiters in licit and illicit supply chains. *Manufacturing & Service Operations Management* **17**(3) 273–289.
- Cho, S.H., X. Fang, S.R. Tayur, Y. Xu. 2018. Combating child labor: Incentives and information disclosure in global supply chains. Working paper, Carnegie Mellon University, Pittsburgh, PA.
- Clemens, J., K. Inderfurth. 2015. Supply chain coordination by contracts under binomial production yield. *Business Research* **8**(2) 301–332.
- Cohen, Maxime C, Ruben Lobel, Georgia Perakis. 2015. The impact of demand uncertainty on consumer subsidies for green technology adoption. *Management Science* **62**(5) 1235–1258.
- Coricelli, Giorgio, Hugo D Critchley, Mateus Joffily, John P O’Doherty, Angela Sirigu, Raymond J Dolan. 2005. Regret and its avoidance: a neuroimaging study of choice behavior. *Nature neuroscience* **8**(9) 1255.
- Cui, Yao, A Yeşim Orhun, Izak Duenyas. 2018. How price dispersion changes when upgrades are introduced: Theory and empirical evidence from the airline industry. *Management Science* .
- Cunha, Jesse M, Giacomo De Giorgi, Seema Jayachandran. 2018. The price effects of cash versus in-kind transfers. *The Review of Economic Studies* **86**(1) 240–281.

- Dada, M., N.C. Petruzzi, L.B. Schwarz. 2007. A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing & Service Operations Management* **9**(1) 9–32.
- Daily Nation. 2017. Kenya: Traders Hoarding Maize to Force Rise in Flour Prices as Famine Persists. January 12, *Helgi Library*, available at <https://allafrica.com/stories/201701250179.html>.
- Dash, Dipak Kumar. 2013. Onion price rise may be due to hoarding: Centre . September 18, *The Times of India*, available at <https://tinyurl.com/ydafkbz9>.
- Dawande, Milind, Srinagesh Gavirneni, Mili Mehrotra, Vijay Mookerjee. 2013. Efficient distribution of water between head-reach and tail-end farms in developing countries. *Manufacturing & Service Operations Management* **15**(2) 221–238.
- Deichmann, Uwe, Aparajita Goyal, Deepak Mishra. 2016. *Will digital technologies transform agriculture in developing countries?*. The World Bank.
- Deshmukh, N. 2011. State to review milk pricing method to check adulteration. *Times of India*, available at <https://timesofindia.indiatimes.com/city/pune/State-to-review-milk-pricing-method-to-check-adulteration/articleshow/7323847.cms> (accessed on May 30, 2018).
- Doyle, M. P., G. H. Loneragan, H. M. Scott, R. S. Singer. 2013. Antimicrobial resistance: Challenges and perspectives. *Comprehensive Reviews in Food Science & Food Safety* **12**(2) 234–248.
- Draaiyer, J. 2002. Milk collection, processing and marketing. *Milk Producer Group Resource Book* 31–50.
- Elmaghraby, Wedad, Elena Katok. 2018. Behavioral research in competitive bidding and auction design. *The Handbook of Behavioral Operations* 525–556.
- Engelbrecht-Wiggans, Richard, Elena Katok. 2007. Regret in auctions: Theory and evidence. *Economic Theory* **33**(1) 81–101.
- Engelbrecht-Wiggans, Richard, Elena Katok. 2008. Regret and feedback information in first-price sealed-bid auctions. *Management Science* **54**(4) 808–819.
- Ethridge, Don E, Bob Davis. 1982. Hedonic price estimation for commodities: an application to cotton. *Western Journal of Agricultural Economics* **7**(1836-2016-150896) 293–300.
- Everstine, K., J. Spink, S. Kennedy. 2013. Economically motivated adulteration (EMA) of food: Common characteristics of EMA incidents. *Journal of Food Protection* **76**(4) 723–735.
- Fafchamps, Marcel, Bart Minten. 2012. Impact of sms-based agricultural information on indian farmers. *The World Bank Economic Review* **26**(3) 383–414.
- Fang, X., S.H. Cho. 2016. Cooperative approaches to managing social responsibility in supply chains: Joint auditing and information sharing. Working paper, Carnegie Mellon University, Pittsburgh, PA.
- FAO. 2009. Milk testing and quality control. Food and Agriculture Organization of the United Na-

- tions, FAO/TCP/KEN/6611 Project, available at <http://www.fao.org/ag/againfo/resources/documents/MPGuide/mpguide2.htm> (accessed May 30, 2018).
- FAO. 2017. Maximum residue limits (MRLs) and risk management recommendations (RMRs) for residues of veterinary drugs in foods. Food and Agriculture Organization of the United Nations, World Health Organization, CAC/MRL 2-2017, available at <http://www.fao.org/fao-who-codexalimentarius/sh-proxy/en/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252Fstandards%252FCAC%2BMRL%2B2%252FMRL2e.pdf> (accessed May 30, 2018).
- Federgruen, A., U. Lall, A.S. Şimşek. 2015. Supply chain analysis of contract farming. Working paper, Columbia University, New York, NY.
- Federgruen, A., N. Yang. 2008. Selecting a portfolio of suppliers under demand and supply risks. *Operations Research* **56**(4) 916–936.
- Federgruen, Awi, Upmanu Lall, A Serdar Şimşek. 2019. Supply chain analysis of contract farming. *Manufacturing & Service Operations Management* **21**(2) 361–378.
- Filiz-Ozbay, Emel, Erkut Y Ozbay. 2007. Auctions with anticipated regret: Theory and experiment. *American Economic Review* **97**(4) 1407–1418.
- Flynn, B.B., X. Zhao. 2014. *Global Supply Chain Quality Management: Product Recalls and Their Impact*. CRC Press, Boca Raton, FL.
- Fudenberg, D., J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA.
- Gadzikwa, L., M. Lyne, S. Hendriks. 2007. Horizontal coordination and free-riding in a group of certified organic crop growers: An empirical study of the Ezemvelo Farmers' Organization in KwaZulu-Natal, South Africa. *African Journal of Agricultural & Resource Economics* **1**(2) 129–144.
- Gale, F., C. Arnade. 2015. Effects of rising feed and labor costs on China's chicken price. *International Food & Agribusiness Management Review* **18** 137–150.
- Gale., F., D. Hu. 2009. Supply chain issues in China's milk adulteration incident. *Agricultural Economists' Conference*. International Association of Agricultural Economists, Beijing, China. Contributed paper no. 342.
- Gallego, Guillermo, Robert Phillips, Özge Şahin. 2008. Strategic management of distressed inventory. *Production and Operations Management* **17**(4) 402–415.
- Ganesh, Maya, Sarang Deo, Sripad K Devalkar. 2019. Leveraging digital technology to improve monitoring and planning in public sector supply chains: Evidence from india's food security program. Available at SSRN 3419350 .
- Golosov, M., G. Lorenzoni, A. Tsyvinski. 2014. Decentralized trading with private information. *Econometrica* **82**(3) 1055–1091.

- Golrezaei, Negin, Hamid Nazerzadeh. 2016. Auctions with dynamic costly information acquisition. *Operations Research* **65**(1) 130–144.
- Gosh, Subir. 2017. Surge in 2015 Pulse Prices Was a Result of Cartellisation. July 6, *The Wire*, available at <https://thewire.in/economy/surge-in-2015-pulse-prices-was-a-result-of-cartellisation>.
- Gouri, Geeta. 2017. How e-payments can prevent cartelization. January 19, *livemint*, available at <https://www.livemint.com/Opinion/lHPYlaKTpUUFETRNLpt4BM/How-epayments-can-prevent-cartelization.html>.
- Goyal, Aparajita. 2010. Information, direct access to farmers, and rural market performance in central india. *American Economic Journal: Applied Economics* **2**(3) 22–45.
- Guda, Harish, Tharanga Rajapakshe, Milind Dawande, Ganesh Janakiraman. 2016. Agricultural support prices in developing economies: Operational analysis and its use in policy making. Working paper, The University of Texas at Dallas.
- Gummagolmath, K. C. 2013. Trends in Marketing and Export of Onion in India. Research Report 2012-13.
- Hamodia. 2018. Ministry: Farmers, wholesalers artificially pumping up onion prices. <https://hamodia.com/2018/11/21/ministry-farmers-wholesalers-artificially-pumping-onion-prices/>.
- Handford, C. E., K. Campbell, C.T. Elliott. 2016. Impacts of milk fraud on food safety and nutrition with special emphasis on developing countries. *Comprehensive Reviews in Food Science & Food Safety* **15**(1) 130–142.
- Helgi Library. 2018. Onion Consumption Per Capita in India. September 18, *Helgi Library*, available at <https://www.helgilibrary.com/indicators/onion-consumption-per-capita/india/>.
- Herald, Deccan. 2010. India imports onion from Pakistan. September 12, *Reuters*, available at <https://www.deccanherald.com/content/122172/india-imports-onion-pakistan.html>.
- Hernandez, Manuel A, Shahidur Rashid, Solomon Lemma, Tadesse Kuma. 2017. Market institutions and price relationships: the case of coffee in the ethiopian commodity exchange. *American Journal of Agricultural Economics* **99**(3) 683–704.
- Hindustan Times. 2019. Onion stocks: 500 quintal cap for dealers, 100 quintal for retailers in chandigarh. October 05, *Hindustan Times*, available at <https://tinyurl.com/y7vftw77>.
- Holst, Gesa Sophie, Daniel Hermann, Oliver Musshoff. 2015. Anchoring effects in an experimental auction—are farmers anchored? *Journal of Economic Psychology* **48** 106–117.
- Hu, M., Y. Liu, W. Wang. 2017. Socially beneficial rationality: The value of strategic farmers, social entrepreneurs and for-profit firms in crop planting decisions. Working paper, University of Toronto, Toronto, Canada.

- Hu, Ming, Yan Liu, Wenbin Wang. 2019. Socially beneficial rationality: The value of strategic farmers, social entrepreneurs, and for-profit firms in crop planting decisions. *Management Science* **65**(8) 3654–3672.
- Hu, Shanshan, Gilvan C Souza, Mark E Ferguson, Wenbin Wang. 2015. Capacity investment in renewable energy technology with supply intermittency: Data granularity matters! *Manufacturing & Service Operations Management* **17**(4) 480–494.
- Huang, L., J.S. Song, R. Swinney. 2017a. Managing social responsibility in multitier supply chains. Working paper, Duke University, Durham, NC.
- Huang, Y., R. Levi, S. Springs, S. Wang, Y. Zheng. 2017b. Risk drivers for economically motivated food adulteration in China’s farming supply chains. Working paper, Massachusetts Institute of Technology, Cambridge, MA.
- IEG. 2011. Capturing technology for development: An evaluation of world bank group activities in information and communication technologies. *Independent Evaluation Group, The World Bank Group* .
- Imbens, Guido W, Jeffrey M Wooldridge. 2009. Recent developments in the econometrics of program evaluation. *Journal of economic literature* **47**(1) 5–86.
- Inouye, A. 2017. China: Poultry and products semi-annual. Tech. rep., USDA Foreign Agricultural Service. GAIN report no. CH17005.
- Inter Press Service. 2012. ‘indian farmers hostage to middlemen. March 9, *Inter Press Service*, available at <http://www.ipsnews.net/2012/03/indian-farmers-hostage-to-middlemen/>.
- Jain, Meetu. 2017. How ‘artificial’ onion crisis created by Indian traders helped Pakistan. September 15, *India Today*, available at <https://tinyurl.com/y9w82bek>.
- Jamieson, Alastair. 2018. Turkey’s erdogan targets inflation by hunting stockpiled onions. April 26, *NBC News*, available at <https://tinyurl.com/y9zdlwxu>.
- Jensen, Robert. 2007. The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The quarterly journal of economics* **122**(3) 879–924.
- Jess, Davis. 2017. Investigation into Indian pulse price surge reveals cartels are to blame. July 12, *ABC Rural*, available at <https://tinyurl.com/ybb7rq68>.
- Johnson, R. 2014. Food fraud and “economically motivated adulteration” of food and food ingredients. Tech. rep., Congressional Research Service. Available at <https://fas.org/sgp/crs/misc/R43358.pdf>.
- Kahneman, Daniel, Amos Tversky. 2013. Prospect theory: An analysis of decision under risk. *Handbook of the fundamentals of financial decision making: Part I*. World Scientific, 99–127.
- Kalkanci, Basak, Morvarid Rahmani, L Beril Toktay. 2018. Social sustainability in emerging economies: The role of ‘inclusive innovation’ .

- Kapur, M., C. Riley. 2013. Rising onion prices bring India woes home. September 13, *CNN Money*, available at <https://money.cnn.com/2013/09/12/news/economy/india-onion/index.html>.
- Karplus, Valerie J, Shuang Zhang, Douglas Almond. 2018. Quantifying coal power plant responses to tighter so2 emissions standards in china. *Proceedings of the National Academy of Sciences* **115**(27) 7004–7009.
- Kaul, Tara. 2018. Household responses to food subsidies: evidence from India. Working paper, University of Chicago.
- Kőszegi, Botond, Matthew Rabin. 2006. A model of reference-dependent preferences. *The Quarterly Journal of Economics* **121**(4) 1133–1165.
- Krishna, Vijay. 2009. *Auction theory*. Academic press.
- Ku, Gillian, Adam D Galinsky, J Keith Murnighan. 2006. Starting low but ending high: A reversal of the anchoring effect in auctions. *Journal of Personality and social Psychology* **90**(6) 975.
- Lebrun, Bernard. 1999. First price auctions in the asymmetric n bidder case. *International Economic Review* **40**(1) 125–142.
- Lee, H.L., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Science* **43**(4) 546–558.
- Letizia, P., G. Hendrikse. 2016. Supply chain structure incentives for corporate social responsibility: An incomplete contracting analysis. *Production and Operations Management* **25**(11) 1919–1941.
- Levi, Retsef, Georgia Perakis, Gonzalo Romero. 2016. On the effectiveness of uniform subsidies in increasing market consumption. *Management Science* **63**(1) 40–57.
- Liao, Chen-Nan, Ying-Ju Chen, Christopher S Tang. 2019. Information provision policies for improving farmer welfare in developing countries: Heterogeneous farmers and market selection. *Manufacturing & Service Operations Management* **21**(2) 254–270.
- Lim, Michael K, Ho-Yin Mak, Seung Jae Park. 2019. Money well spent? operations, mainstreaming, and fairness of fair trade. *Production and Operations Management* **28**(12) 3023–3041.
- Liu, Qian, Garrett Van Ryzin. 2011. Strategic capacity rationing when customers learn. *Manufacturing & Service Operations Management* **13**(1) 89–107.
- livemint. 2016. Govt to increase rice subsidy given via DBT scheme from October. September 12, *Reuters*, available at <https://tinyurl.com/y9h5gata>.
- Lune, Howard, Bruce L Berg. 2016. *Qualitative research methods for the social sciences*. Pearson Higher Ed.
- Lynn, Michael. 1991. Scarcity effects on value: A quantitative review of the commodity theory literature. *Psychology & Marketing* **8**(1) 43–57.
- Lynn, Michael, Paulette Bogert. 1996. The effect of scarcity on anticipated price appreciation. *Journal of Applied Social Psychology* **26**(22) 1978–1984.

- Malhotra, Deepak. 2010. The desire to win: The effects of competitive arousal on motivation and behavior. *Organizational behavior and human decision processes* **111**(2) 139–146.
- Mankiw, N Gregory. 2014. *Principles of macroeconomics*. Cengage Learning.
- McDonald, John H. 2009. *Handbook of Biological Statistics*. 2nd ed. Sparky House Publishing, Baltimore, MD.
- Meena, P.L., S.P. Sarmah, A. Sarkar. 2011. Sourcing decisions under risks of catastrophic event disruptions. *Transportation Research Part E: Logistics & Transportation Review* **47**(6) 1058–1074.
- Milgrom, Paul R, Robert J Weber. 1982. A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society* 1089–1122.
- Ministry of Agriculture. 2015. Operational guidelines for price stabilization fund. March, available at http://agricoop.nic.in/sites/default/files/opr_gud_303_0.pdf.
- Mishra, Vinit Kumar, Karthik Natarajan, Dhanesh Padmanabhan, Chung-Piaw Teo, Xiaobo Li. 2014. On theoretical and empirical aspects of marginal distribution choice models. *Management Science* **60**(6) 1511–1531.
- MOAFW. 2017. Monthly Report Onion (September, 2017). Report, Horticultural Statistics Division, Ministry of Agriculture & Farmers' Welfare.
- Mondiale, Banque. 2008. World development report: Agriculture for development. *The International Bank for Reconstruction and Development/The World Bank* .
- Mu, L., M. Dawande, X. Geng, V. Mookerjee. 2016. Milking the quality test: Improving the milk supply chain under competing collection intermediaries. *Management Science* **62**(5) 1259–1277.
- Mu, L., M. Dawande, V. Mookerjee. 2014. Improving the milk supply chain in developing countries: Analysis, insights, and recommendations. *Production & Operations Management* **23**(7) 1098–1112.
- Mungroo, N., S. Neethirajan. 2014. Biosensors for the detection of antibiotics in poultry industry – A review. *Biosensors* **4**(4) 472–493.
- Musalem, Andrés, Marcelo Olivares, Eric T Bradlow, Christian Terwiesch, Daniel Corsten. 2010. Structural estimation of the effect of out-of-stocks. *Management Science* **56**(7) 1180–1197.
- Nagurney, A., D. Li. 2016. *Competing on Supply Chain Quality, Springer Series in Supply Chain Management*, vol. 2. Springer International Publishing, Basel, Switzerland.
- Narrood, C., M. Tiongco, A. Costales. 2008. Global poultry sector trends and external drivers for structural change. *Proceedings of the International Poultry Conference on Poultry in the 21st century: Avian influenza & beyond*. Food and Agricultural Organization of the United Nations, Bangkok, Thailand, 1–28.
- Nasiry, Javad. 2010. *Pricing and revenue management with behavioral considerations*. INSEAD.
- Nautiyal, PC. 2002. Groundnut: post-harvest operations. *Food and Agriculture Organization of the United Nations* .

- Nestlé. 2015. Nestlé in society: Creating shared value. Available at http://www.nestle.com/asset-library/documents/library/documents/corporate_social_responsibility/nestle-csv-full-report-2015-en.pdf.
- Nirmal, Rajlakshmi. 2017. Why the eNAM platform hasn't taken off despite all the fanfare. July 18, *BusinessLine*, available at <https://www.thehindubusinessline.com/economy/agri-business/why-the-enam-platform-hasnt-taken-off-despite-all-the-fanfare/article9776034.ece>.
- Novemsky, Nathan, Daniel Kahneman. 2005. The boundaries of loss aversion. *Journal of Marketing research* **42**(2) 119–128.
- Orsdemir, A., B. Hu, V. Deshpande. 2016. Ensuring corporate social and environmental responsibility through vertical integration and horizontal sourcing. Working paper, UC Riverside, Riverside, CA.
- Ovchinnikov, A. 2018. Incorporating customer behavior into operational decisions. K. Donohue, E. Katok, S. Leider, eds., *The Handbook of Behavioral Operations*, chap. 17. John Wiley & Sons, Hoboken, NJ, 587–617.
- Özer, Ö., Y. Zheng. 2012. Behavioral issues in pricing management. Ö. Özer, R. Phillips, eds., *The Oxford Handbook of Pricing Management*, chap. 20. Oxford University Press, Oxford, United Kingdom, 415–461.
- Özer, Ö., Y. Zheng. 2015. Markdown or everyday low price? The role of behavioral motives. *Management Science* **62**(2) 326–346.
- Pal, Sanchari. 2017. We indians love our dal. but do you know the fascinating history of our favourite comfort food? March 23, *The Better India*, available at <https://tinyurl.com/yafovd2y>.
- Parker, Chris, Kamalini Ramdas, Nicos Savva. 2016. Is IT enough? Evidence from a natural experiment in India's agriculture markets. *Management Science* **62**(9) 2481–2503.
- Pi, C., Z. Rou, S. Horowitz. 2014. Fair or fowl? Industrialization of poultry production in China. Tech. rep., Institute for Agriculture & Trade Policy.
- Pikkemaat, M. 2009. Microbial screening methods for detection of antibiotic residues in slaughter animals. *Analytical & Bioanalytical Chemistry* **395**(4) 893–905.
- Pindyck, Robert, Daniel Rubinfeld. 2013. *Microeconomics GE.*. Pearson Education Inc.
- Plambeck, E.L., T.A. Taylor. 2016. Supplier evasion of a buyer's audit: Implications for motivating supplier social and environmental responsibility. *Manufacturing & Service Operations Management* **18**(2) 184–197.
- Plum, Michael. 1992. Characterization and computation of nash-equilibria for auctions with incomplete information. *International Journal of Game Theory* **20**(4) 393–418.
- PRC-MIIT. 2011. Notice of maximum residue limit of melamine in food products (in Chinese). Ministry of Industry and Information Technology of the People's Republic of China, notice no.

- 10, available at <http://www.miit.gov.cn/n1146295/n1652858/n1652930/n4509607/c4513875/content.html> (accessed May 30, 2018).
- PRC-NPC. 2015. Food safety law of the People's Republic of China (in Chinese). The National People's Congress of the People's Republic of China, available at http://www.npc.gov.cn/npc/cwhhy/12jcw/2015-04/25/content_1934591.htm (accessed May 30, 2018).
- Prelec, Drazen, George Loewenstein. 1998. The red and black: Mental accounting of savings and debt. *Marketing science* .
- Quint, Daniel, Kenneth Hendricks. 2013. Indicative bidding in auctions with costly entry .
- Rhoades, Stephen A. 1993. The herfindahl-hirschman index. *Fed. Res. Bull.* **79** 188.
- Robinson, Peter John, WJ Wouter Botzen. 2019. Determinants of probability neglect and risk attitudes for disaster risk: An online experimental study of flood insurance demand among homeowners. *Risk Analysis* **39**(11) 2514–2527.
- Roese, Neal J, KAI Epstude, Florian Fessel, Mike Morrison, Rachel Smallman, Amy Summerville, Adam D Galinsky, Suzanne Segerstrom. 2009. Repetitive regret, depression, and anxiety: Findings from a nationally representative survey. *Journal of Social and Clinical Psychology* **28**(6) 671–688.
- Ross, S.M. 2005. *Introductory Statistics*. Academic Press, Cambridge, MA.
- Sarkar, Md. Sujahangir Kabir, Hiranman Sikdar, M Alam, K Rahmana. 2010. Impact of price hike on the consumers livelihood. *International Journal of BioResearch* **2** 23–27.
- Sharma, Aman. 2018. DBT payments cross Rs 1 lakh crore in '17-18; govt may declare savings of Rs 75,000 crore . *The Economic Times*, available at <https://tinyurl.com/ybfu4cny>.
- Sharma, S., Z. Rou. 2014. China's dairy dilemma. Tech. rep., Institute for Agriculture & Trade Policy.
- Sharma, Vijay Paul. 2012. Food Subsidy in India: Trends, Causes and Policy Reform Options. Working Paper.
- Sheffi, Y. 2005. *The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage*. MIT Press, Cambridge, MA.
- Shi, Xiaohui, Feng Li, Pattarin Chumnumpan. 2020. The use of product scarcity in marketing. *European Journal of Marketing* **ahead-of-print**. doi:10.1108/EJM-04-2018-0285.
- Sidhartha. 2016. 500 central schemes to go DBT way next fiscal year . October 19, *The Times Of India*, available at <http://timesofindia.indiatimes.com/articleshow/54928603.cms>.
- Simchi-Levi, D., W. Schmidt, Y. Wei. 2014. From superstorms to factory fires: Managing unpredictable supply chain disruptions. *Harvard Business Review* **January–February** 96–101.
- Singh, AD. 2014. Maize in India. Tech. rep., Federation of Indian Chambers of Commerce and Industry, New Delhi, India. India Maize Summit.

- Stirzaker, D. 1999. *Probability and Random Variables: A Beginner's Guide*. Cambridge University Press, Cambridge, England.
- Strzebicki, Dariusz. 2015. The electronic marketplace as the element of the agricultural market infrastructure. *Problems of Agricultural Economics* (1).
- Tanzina, A., A. Shoeb. 2016. Common milk adulteration and their detection techniques. *International Journal of Food Contamination* **3**(1) 22–30.
- Taylor, Terry A, Wenqiang Xiao. 2014. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Management Science* **60**(10) 2461–2477.
- Telangana Today. 2017. Onion prices in telangana bringing tears to buyers. November 14, *Telangana Today*, available at <https://telanganatoday.com/onion-prices-telangana-bringing-tears-buyers>.
- Thaler, Richard. 1985. Mental accounting and consumer choice. *Marketing Science* **4**(3) 199–214.
- The Economic Times. 2016. Centre asks states to implement direct benefit transfer for food subsidy. April 26, *The Economic Times*, available at <https://tinyurl.com/y3on8xb4>.
- The Economic Times. 2018. Govt estimates 4.5% drop in onion output for 2017-18. <https://tinyurl.com/y6t6s916>.
- The Economic Times. 2019. Farmers in poll-bound maharashtra fume at centre over onion export ban. September 30, *The Economic Times*, available at <https://tinyurl.com/y9u5k6o2>.
- The New Indian Express. 2019. Notices slapped on 47 traders for hoarding onions. November 8, *The New Indian Express*, available at <https://tinyurl.com/ydcq5e3n>.
- The Quint. 2019. 'killing our agriculture': Analysts on govt banning onion exports. September 30, *The Quint*, available at <https://www.thequint.com/voices/opinion/analysts-on-govt-export-ban-onions>.
- Tomlin, B. 2009. Disruption-management strategies for short life-cycle products. *Naval Research Logistics* **56**(4) 318–347.
- Tversky, Amos, Daniel Kahneman. 1991. Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics* **106**(4) 1039–1061.
- Van Mieghem, J. 2007. Risk mitigation in newsvendor networks: Resource diversification, flexibility, sharing, and hedging. *Management Science* **53**(8) 1269–1288.
- Vissa, Kirankumar. 2017. How long can india's farmers subsidise the nation? November 24, *BusinessLine*, available at <https://thewire.in/agriculture/farmers-protests-kisan-ki-loot-msp-fair-price>.
- Wang, J.J.D., J.F. Zender. 2002. Auctioning divisible goods. *Economic Theory* **19**(4) 673–705.
- Wang, S., P. Sun, F. Véricourt. 2016. Inducing environmental disclosures: A dynamic mechanism design approach. *Operations Research* **64**(2) 371–389.

-
- Wang, Y., W. Gilland, B. Tomlin. 2010. Mitigating supply risk: Dual sourcing or process improvement? *Manufacturing & Service Operations Management* **12**(3) 489–510.
- Wodon, Quentin, Hassan Zaman. 2008. *Rising food prices in sub-Saharan Africa: Poverty impact and policy responses*. The World Bank.
- Wongpom, B., S. Koonawootrittriron, M.A. Elzo, T. Suwanasopee. 2017. Milk yield, fat yield and fat percentage associations in a thai multibreed dairy population. *Agriculture and Natural Resources* **51**(3) 218–222.
- Yan, A. 2017. Suspects in China’s ‘biggest baby milk scandal in decade’ go on trial. March 27, *Southern China Morning Post*, available at <http://www.scmp.com/news/china/policies-politics/article/2082409/suspects-chinas-biggest-baby-milk-scandal-decade-go>.
- Yayla-Kullu, Muge, Omkar Palsule-Desai, Srinagesh Gavirneni. 2019. Reining in onion prices by introducing a vertically differentiated substitute: Models, analysis, and insights. Working paper, July 8.
- Ye, Lixin. 2007. Indicative bidding and a theory of two-stage auctions. *Games and Economic Behavior* **58**(1) 181–207.
- Yousaf, Muhammad, Jifu Li, Jianwei Lu, Tao Ren, Rihuan Cong, Shah Fahad, Xiaokun Li. 2017. Effects of fertilization on crop production and nutrient-supplying capacity under rice-oilseed rape rotation system. *Scientific reports* **7**(1) 1270.
- Zhang, Dennis, Yuval Salant, Jan A Van Mieghem. 2018. Where did the time go? on the increase in airline schedule padding over 21 years. *On the Increase in Airline Schedule Padding Over* **21**.
- Zhang, H., G. Aydin, H.S. Heese. 2017. Curbing the use of conflict minerals: A supply network perspective. Working paper, Indiana University, Bloomington, IN.
- Zhang, J., T. Bhatt. 2014. A guidance document on the best practices in food traceability. *Comprehensive Reviews in Food Science & Food Safety* **13**(5) 1074–1103.

