Emerging Questions in Agricultural Finance

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Youwei Yang

August 2021

Advisors: Calum Turvey (Chair), Yongmiao Hong, Jennifer Ifft, Lin William Cong, Scott Yonker

Emerging Questions in Agricultural Finance

Youwei Yang, Ph.D.

Cornell University 2021

Observing a fast-evolving world economy and agricultural financial system, this dissertation studies three emerging questions in agricultural finance: agricultural financial technology (FinTech), blockchain and cryptocurrencies; persistence and growth of religious farming communities that limit use of technology in the U.S.; and livestock insurance willingness to offer in China. These emerging questions and developments fit closely into the top priority of the G20 Global Partnership for Financial Inclusion (GPFI), and the vital tools recognized to meet the Sustainable Development Goals (SDGs) set by the United Nations. These emerging agricultural finance strategies are crucial to improve agricultural value chain, reduce inequality, and spur economic growth, especially for rural areas with intensive agricultural activities. These topics are sparsely studied empirically due to limited data, evolving policies, and rapidly changing technologies. In this thesis I take the initiative to understand these both qualitatively and quantitatively.

The first chapter discusses the overall importance of these emerging areas of agricultural finance strategies to support the development of the agricultural economy. We provide background domain knowledge on Agricultural FinTech with an emphasis to blockchain and cryptocurrencies, elaborating the applications of blockchain in food and agriculture value chain and supply chain, that improves the transparency and efficiency of tracing and trading. Understanding the relationship between blockchain and cryptocurrencies is crucial to applications of this technology in agriculture, thus we relay quantitative research to the more practically operated cryptocurrencies to understand its market property.

Chapter 2 examines the financial property of 3,351 cryptocurrency price series, particularly looking at the long or short memory within its price series and evaluate if they follow a fractional Brownian motion through Hurst estimators and Stepwise autoregression. We investigate the releasing mechanism of Bitcoin and use top 105 coins' supply structure to explain the memory in their price series, for which we find the % of total coins issued is

explanatory. Due to concerns of supply related built-in memory within the price series, we propose a de-mean method to decompose the price series with the time varying mean and the variations, and we found that some previously found long memory was actually spurious. The gradually increasing (time-varying) mean can explain this spuriousness, and we use the market structure such as the 4-year bull circle and the deflationary releasing mechanism to explain this time-varying phenomenon. The findings of this chapter confirm the fractionality of crypto markets generally, and provide clarity of true or spurious long memory using both traditional and advanced methods. This enhances our understanding of these emerging financial assets that could potentially be applied to agricultural value chains. One possibility is a blockchain token system that is linked to warehouse receipts and enables the holding and trading of grains electronically.

The third chapter looks at Amish population growth and how that could affect farmland prices. The Amish culture promotes strong family ties and human-nature interactions; thus, they generally use less modern agricultural production technology. We conceptualize the co-existence of different farming styles between Amish farmers and conventional farmers, to explain how the less productive Amish farmers can stay competitive. The main rationale is that the effectively cheaper labor costs of Amish due to their larger families can compensate the disadvantage from less output due to their limited use of farming technologies. Simply speaking, Amish may have lower revenue, though their costs are low as well, so they could maintain the similar level of profits to their non-Amish neighbors. Because farmland is the most important farm asset class, we hypothesize that any differences in profitability of farming systems would be reflected in farmland prices and analyze whether farmland prices are influenced by Amish population growth. We use a standard hedonic approach and unique shift-share like instrumental variable in this empirical model and find no statistically significant relationship between Amish population growth and farmland prices. Therefore, we infer that Amish compete on the farmland markets similar as the conventional farmers, which aligns with our conceptual framework scenario of similar profitability. This chapter does not only offer insights into the coexistence of multiple farming systems in one market, represented in

the farmland market; but also, methodologically showcases how an identification strategy from the labor literature can be applied to agriculture finance issues.

The fourth chapter presents our work in investigating agricultural insurance agents' willingness to offer (WTO) livestock insurance in China, through an in-the-field discrete choice experiment (DCE). We include eight main attributes of livestock insurance and contain various combinations of different levels of them on the choice cards. We implemented the analysis in 6 blocks, with 35 insurance agents in each block. We ask each of them to respond to 15 cards, on each card choose from one of the two hypothetically designed livestock insurance. The card choice combinations of various attributes and levels are designed using JMP software through a D-optimal approach, which use limited sets of choices to maximize the choice exposures to participants and reveal their utility changes when they decide the trade-offs between two choices on the cards based on the different attribute levels included. Premium subsidy and insurance types are two out of eight attributes we primarily study and also find strong clear evidence on. We find that a one level (10%) increase of subsidy lead agents' probability to offer be 3.166 times higher. We also find the more traditional type, mortality insurance, is still strongly preferred than the newer introduced insurance, with weather-based index insurance being the least preferred, because of farmers' difficulty in understanding and conflicts when basis risk occurs. Through using survey question to generate interaction term model, we also find knowledge to the newer type of the insurance products improves the WTO on that particular insurance type. This chapter is among the very first to study the supply side of agriculture insurance, and the DCE method we use is a recently popularized method in evaluating insurance products and participants preferences with designs of attributes variations. Our research provides important policy implications as the government and insurance companies work out the details in enlarging the take-up of insurance via interventions of subsidy, education, and innovative /diverse product offering.

BIOGRAPHICAL SKETCH

Born in Gansu, Rural Northwestern China; fruits, farming, animal husbandry, and related finance development have all played a role in Youwei Yang's life growing up. Youwei is a PhD candidate studying Agricultural Finance in Dyson School of Applied Economics and Management at Cornell University. His recent research involves (1) cryptocurrency time series, blockchain, and agricultural financial technology, (2) farmland markets and Amish population in NY, (3) willingness to offer of livestock insurance agents in China via Choice Experiment.

At Cornell, Youwei has instructed sections as the lead TA for core courses in Economics, Business, Finance, and Statistics. Including: Statistical Methods II: Multiple Regression Analysis (PhD level); Econometrics-II (upper undergraduate level); Futures, Options and Financial Derivatives (upper undergrad / master's level); Big Data, FinTech and Finalytics (MBA course); and two introductory level courses in Business Management and Development Economics.

Youwei came to the United States in 2013, attended Kansas State University for 3 years, and earned a bachelor's degree with honors in Agricultural Economics and Minor in Statistics. His undergraduate honor's thesis uses a partial equilibrium model to test beef market shocks, is titled "The Impact of U.S. Beef Exports on U.S. Domestic Beef Price". The thesis won the 3rd place in the 2016 SS-AAEA national outstanding paper competition. He was an undergraduate research assistant for three reputable professors at K-State and mainly worked on two USDA funded projects: rural grocery stores and healthy food choices; and the impact of policy interventions on farmers' producing incentives in multiple countries.

Youwei had extensive practical market research experience in the private sector. He worked at StoneX (NASDAQ: SNEX) full time in 2016-2017 and part-time in 2018-2021, as a Market Intelligence Analyst. He mainly conducted commodity market analysis, primarily agricultural. He led several projects: (1) Traded commodity index and momentum indicator; (2) Economic impact of China's African Swine Fever; (3) Global meat & meal demand forecast. During summer 2015, Youwei interned in U.S. 5th largest cattle feeding operation – Cattle Empire LLC as Risk Management Intern. Besides, he supported a startup to analyze business opportunities in the international trade of pet food and nutrients, also pet diagnostic services.

ACKNOWLEDGMENTS

I would like to first and foremost express my deepest gratitude to the Chair of my dissertation committee, Prof. Calum Turvey, for his guidance and encouragement throughout my graduate studies. I was lucky to have your input during our many meetings and to have your generous support through the enlightening, joyful, and especially the difficult times.

I would also like to thank my advisory committee members, Prof. Yongmiao Hong, Prof. Lin William Cong, and Prof. Scott Yonker. You are great mentors who inspire my intellectual pursuit in econometrics, finance, and financial economics. Your advice enhances my skills in storytelling and conceptualization of market events using financial language. Your motivation and suggestion deepen my exposure to Tokenomics and financial technology research.

My minor member in agricultural finance and research assistantship supervisor, Prof. Jennifer Ifft, provided me with the most precious guidance on academic writing and the identification strategy with advanced econometrics. The writing and methods would not have been as rigorous and in-depth without her expertise.

My sincere thanks are due to Dr. Wenjun Long (Ministry of Agriculture of China) and his research team, graduate students at China Agriculture University. I am very grateful for all the efforts they spent on the arrangement, and the supports to conduct the field experiment with the extensive participants of agriculture insurance agents.

Moreover, I would like to thank the support from W.I. Myers Endowment Fund and USDA funding. I am also grateful for Arlan Suderman and Eileen Stein, managers of my part time job at StoneX Group Inc., who offer me substantial support and keeps me stay in the forefront of the real-world research and application of financial economics and FinTech.

I was blessed to be surrounded by intelligent and dedicated classmates, colleagues, and professors in the Dyson School of AEM, Department of Economics, and SC Johnson School of Business. Thank you for your support, suggestions, and feedback on my research and coursework. I am grateful for you being part of this challenging but rewarding graduate studies.

Lastly, and most importantly, I would like to express my love and gratitude to my family. Thank you for making my academic journey well-balanced and extraordinarily supported.

Table of Contents

ABSTRACT **2**

BIOGRAPHICAL SKETCH **5**

ACKNOWLEDGMENTS **6**

CHAPTER AND SECTIONS **7**

LIST OF TABLES **9**

LIST OF FIGURES **9**

1. CHAPTER 1 11

Introduction - Emerging Agricultural Finance Strategies 11

1.1. Agricultural Financial Technology, Blockchain, and Cryptocurrencies 11

1.2. Farmland Markets in Minority Communities - Amish 18

1.3. China's Livestock Insurance Supply via Field Choice Experiment 19

2. CHAPTER 2 22

2.1. Introduction 22

2.1.1. Motivation 22

2.1.2. Background 24

2.1.3. Purpose 26

2.2. Literature Review 28

2.3. Cryptocurrency Releasing Mechanism and Supply/Demand Equilibrium 32

2.3.1. Blockchain and Cryptocurrency 32

2.3.2. The Role of Miners and Mining in Blockchain Development 33

2.4. Model and Theory of Random Walk Process 37

2.4.1. Geometric / Fractional Brownian Motion & Efficient Market Hypothesis 37

2.4.2. Traditional Hurst Coefficient Estimation & Implication in gBm / fBm 39

2.4.3. Time-varying Mean Estimation, Structural Breaks, and De-Mean Hurst Effect Enhanced Method 41

2.4.4. Stepwise Autoregression and Quasi-Fractional Brownian Motion 43

2.5. Empirical Evidence of Cryptos' Fractionality and Reasons 44

2.5.1. Data Illustration 44

2.5.2. Traditional Method Hurst Coefficient Estimations and Distribution 45

2.5.3. Hurst Coefficient of Traditional Method vs. De-Mean Enhanced Method 47

2.5.4. Summation of Stepwise Autoregression Estimates and Quasi - fBm 51

2.5.5. The Relationship between the Hurst Measure and Supply Structure: OLS Regression 54

2.6. Conclusion and Discussion 56

2.7. References: 58
3. CHAPTER 3 64
Horses vs. Tractors? Old Order Amish Population Growth and New York Farmland Markets 64
3.1. Introduction 64
3.2. Conceptual Model 68
3.2.1. Theoretical Background: Ricardian Theory of Rent 71
3.2.2. Capitalization Model and Production Profit Maximization 72
3.2.3. Market Clearing Conditions and Equilibrium 74
3.2.4. Intuitive Discussion of Equilibrium 75
3.3. Data Source and Descriptive Statistics 76
3.4. Empirical Model and Identification Strategy 79
3.4.1. Standard OLS Model: Farmland Price and Amish Population 80
3.4.2. Potential Endogeneity 81
3.4.3. Instrumental Variable Identification Strategy 82
3.5. Empirical Results 83
3.5.1. Standard OLS Model Results 84
3.5.2. Enclave IV Two-stage Least Squares (2SLS) Model Results 84
3.5.3. Robustness Check: Alternative IV 87
3.6. Conclusion 87
3.7. References 89
4. CHAPTER 4 95
The Willingness to Offer Livestock Insurance in Rural China: A Discrete Choice Experiment Among Chinese Insurance Agents 95
4.1. Introduction 95
4.2. Empirical Methodology 100
4.2.1. Discrete Choice Experiment Method 100
4.2.2. Choice Cards Design and Survey Questions 101
4.2.3. Conditional Logit and Mixed Logit Regression Models 107
4.2.3.1. Conditional Logit Model 108
4.2.3.2. Mixed Logit Model 110
4.2.3.3. Willingness to Offer (WTO) 111
4.3. Data and Results 111
4.3.1. Demographic Characteristics and Selected Survey Questions Summary 112
4.3.2. Conditional and Mixed Logit Model Results 114
4.3.3. Odds Ratio 118
4.3.4. Willingness to Offer (WTO) 120

4.4. Conclusion 122
4.5. Reference 124
5. CHAPTER 5 130
Conclusion – Investigating Emerging Issues in Agricultural Finance 130
APPENDIX 134

LIST OF TABLES

Table 2.1: Bitcoin Hurst Coefficients in Different Time Windows: Traditional vs. De-Mean Enhanced Method 50
Table 2.2: Ethereum Hurst Coef. in Different Time Windows: Traditional vs. De-Mean Enhanced Method 50
Table 2.3: Summation of Stepwise Autoregression Estimates for Top 20 Crypto Currencies Price History 51
Table 2.4: Top 105 Cryptocurrencies influencing factors data summary. 54
Table 2.5: Top 105 Cryptocurrencies Hurst Coefficients OLS Fit with Age, Supply Issued, and Type 55
Table 3.1: Farm Credit East dairy farm 2018 survey data: Conventional vs. Plain Farms 76
Table 3.2: Key variables summary statistics 77
Table 3.3: Regular OLS regression results 84
Table 3.4: Enclave IV two-stage least squares (2SLS) regression results 85
Table 3.5: Enclave IV (dropping small acres) 86
Table 3.6: Robustness check: Alternative IV value yield 87
Table 4.1: Attributes and Levels used in Discrete Choice Experiment. 102
Table 4.2: Some Demographic Characteristics of Insurance Agents Respondents (N = 211) 113
Table 4.3: Conditional and Mixed Logit Results With/Without Type Knowledge Interaction Terms ... 115
Table 4.4: Conditional Logit Results by 4 Regions 117
Table 4.5: Mixed Logit Results by 4 Regions 118
Table 4.6: Odds Ratios for Conditional Logit (CL) and Mixed Logit (ML) Models 119
Table 4.7: Willingness to Offer (WTO) in Conditional Logit (CL) and Mixed Logit (ML) 122

LIST OF FIGURES

Figure 2.1: Top 4 Crypto Coins Daily Price Series and Natural Log Change 26
Figure 2.2: Theoretically Simulated Bitcoin Total Supply 35
Figure 2.3: New Addition to Supply following the Bitcoin Model 35
Figure 2.4: Demand and Supply of Cryptocurrency under the Bitcoin Model 36

Figure 2.5: Distribution of Cryptos' Hurst Coefficients Measured in Lag of Observation's Square Root 47

Figure 2.6: Natural Log Nonparametrically Estimated Mean of Bitcoin Prices 49

Figure 2.7: Natural Log Nonparametrically Estimated Mean of Ethereum Prices 49

Figure 2.8: Three-Time Range Scatterplot Cryptos Stepwise Autoregression Coefficients Summation . 53

Figure 3.1: Amish Production and Equilibrium Conceptual 70

Figure 3.2: Average number of Amish districts located with 10 miles of an arms-length NY farmland transaction (Ifft and Gao, 2019) 78

Figure 3.3: New York state total number of parcels sold trend 78

Figure 3.4: New York farmland sales with nearby Amish districts, 1999-2015 (Ifft and Gao, 2019) 79

Figure 3.5: Amish Settle in Lower Farmland Price Areas in NY 82

Figure 4.1: Choice cards in block design 107

1. CHAPTER 1

Introduction - Emerging Agricultural Finance Strategies

Agricultural finance is important to support the growth, strength, and long-term development of agricultural production and economy. Financial technology in the agriculture and value chain, land ownership and financing for minorities and special communities in the agricultural sector, and agricultural insurance in developing countries are three of the key areas of agricultural finance innovation that have been emerging. These are among the top priorities of the G20 Global Partnership for Financial Inclusion (GPFI)[1], and the vital tools recognized to meet the Sustainable Development Goals (SDGs) set by the United Nations[2]. These emerging agricultural finance strategies are crucial to improve health and education, reduce inequality, and spur economic growth, especially for rural areas with intensive agricultural activities. We study three areas of the emerging agricultural finance issues, including agricultural financial technology (FinTech) and cryptocurrencies, Amish population growth and farmland markets, and livestock insurance supply in China. These topics are sparsely studied due to limited data and the fast-changing structure as they are still being piloted, we take the initiative to understand them both qualitatively and quantitatively. We provide background discussion in this chapter especially for the sector of Agri FinTech, blockchain and cryptocurrencies; as it is relatively the newest emerging topic that requires substantial domain knowledge to start comprehending how it may impact agricultural finance.

1.1. Agricultural Financial Technology, Blockchain, and Cryptocurrencies

Understanding the needs, opportunities, and challenges for small farmers and the approaches to enhance their inclusion into the financial system has been a major focus in agriculture. The rapid pace of technological transformation in developing and underdeveloped countries has important implications for the financial inclusion of the world's poor. Amid this change, new technologies that enable inclusive agricultural finance have garnered heightened interest.

[1] "G20 Roundtable on Innovations in Agricultural Finance" convened on September 9, 2015, in Antalya, Turkey by the Small and Medium Enterprise (SME) Finance Sub-Group.

[2] United Nations, Department of Economic and Social Affairs. https://sdgs.un.org/goals

From mobile phones, drones to financial technologies; digital tools may one day overcome longstanding barriers to reaching the world's 500 million smallholder farming households (Christen and Anderson 2013). Starting in 2009 and during the past decade, distributed ledger technologies (DLT), more regularly referred to as Blockchain has first hyped as driven by the Bitcoin craze, because the best-known Blockchain is the Bitcoin peer-to-peer electronic cash system. In more recent years, some potentially materializable Blockchain applications have begun piloting. Among the sectors of blockchain applications that are developing on the frontier, financial services are the most practical and promising[3], then followed by the supply chain and value chain in the food and agriculture sector[4].

Blockchain is built around the concept of distributed ledger, or shared record of transactions. Unlike traditional ledgers that are maintained in a centralized server by a trusted third party or government authorities, Blockchain theoretically provides a cryptographic mechanism for creating a shared record of transactions among several institutions or many individuals in the absence of a trusted arbiter. Participants who usually are programmers can become individual validators of blockchain, as verification nodes, who verify the encrypted hash of transactions and contracts then record it on a block after consensus has reached among other validators on this transaction. Likewise, when editing or deleting a transaction, it also needs consensus, which would prevent malicious manipulation by a single person or small group of people who has authority in the centralized system, also refers to as the single point failure. Validators will receive a reward by conducting these verification jobs, usually the platform tokens generated by the activities on this blockchain. If validators keep a good record of being an honest node, they will eventually gain governance or voting rights in this ecosystem. If the mechanism of incentive, governance, and consensus is designed properly; the community will grow sustainably, thereafter the blockchain and its tokens will become more valuable. Tokens can be traded on crypto exchanges, for people who believe and invest

[3] Deloitte's 2020 GLOBAL BLOCKCHAIN Survey - From Promise to Reality.

[4] One major example: a consortium of major food companies (Walmart, Dole, Driscoll, Golden State Foods, Kroger, McCormick & Company, Nestlé, and Tyson Foods) are collaborating with IBM and AWS to use distributed ledger solutions to make their food supply chains more transparent, more traceable and to streamline payments.

in the particular blockchain's potential, there of course are many speculators involved as well. The money raised from the sale of tokens to the public, can be used to reward founding teams and validators. The token sale and reward system provides a more open and transparent way of incentivizing and governing to develop a blockchain that will eventually be practically beneficial for a specific community or purpose, such as lending, tracing, cloud computing, auditing, and more recently, art galleries. The economics of blockchain token system design is recently termed "Tokenomics".

Blockchain claims to have the benefits of transparency, shared control, disintermediation, immutability, reduced transaction costs, resilience, and interoperability[5]. It seems like a panacea that tackles many long-enduring problems facing by the current system in many businesses including agricultural finance, it may not! There is an already discovered challenge: "Blockchain Trilemma", termed by Vitalik Buterin[6], which refers to the difficulties that developers face in creating a blockchain that is scalable, decentralized, and secure — without compromising on any facet. Set aside how successful these promises are met in practice over the years of exploration, as the industry is still in more of a piloting phase overall which takes patience and time to prove. It is not the focus of this dissertation to examine the technological successes or failures of blockchain and its applications in agriculture. Another basic but practical challenge is that implementation and use of this technology need electronic devices and computers, which are usually lacking in developing areas. CGAP found that an average of just 5.5 percent of smallholders across six markets surveyed owned a smartphone (Anderson 2015, 2016). We briefly introduce some applications in agricultural finance that may potentially develop value: value chain, collateralization, supply chain management, credit bureaus, and contracts.

The high administrative cost of serving sporadically located smallholders; lack of collateral, financial records, and formal identity; and challenges in contract enforcement are just a few of the numerous barriers between small farms and conventional financial

[5] Reference: "Exploring Blockchain Applications to Agricultural Finance", by CGAP, July 2018.

[6] Vitalik is a Russian-Canadian programmer and is best known as one of the co-founders of Ethereum, the most actively used blockchain platform, with its token being the 2nd largest cryptocurrency (ETH) by market cap.

institutions. The underdeveloped financial inclusion disturbs the efficient operating of agricultural value chains. Because farmers and ranchers may be unable to invest in machinery, fertilizers, pesticides, and barns to maximize their production and profits; thus, cannot eventually to re-invest some of their profits to expand. It is generally a vicious circle for financially excluded smallholders to get out of the "Poverty Trap". Distributors and consumers may struggle to access an adequate, consistent, and safe supply of agricultural commodities and foods. A recent survey from sprout social shows that 86% of Americans believe transparency from businesses is more important than ever before; and 73% of consumers are willing to pay more for products that guarantee total transparency[7]. Blockchain can exactly provide transparency for consumers and this premium pay can reward hardworking farmers, also empower the overall agricultural value chain. Without accessible financial services, cash-strapped smallholders may sometimes have to sell their crops or animals at lower prices in exchange for quicker payment[8].

The major benefits of Blockchain - disintermediation and shared control can help to tackle issues in the value chain like the above mentioned. Because business activities can be more active and thus emergent cash need liquidity can be smoothed when having a broader network that is developed in a transparent way. Also, when there are no third-party, peer-to-peer transactions can save intermediation costs, at the same time under the supervision of other participants in the ecosystem. The collateralization of agricultural assets such as land, crop, livestock, and equipment might support smallholders to access financing for inputs and post-harvest liquidity. Digitalization and encrypted publicization of collateral activities can enhance transparency and contract liquidation through shared control. Likewise, distributed ledger in supply chain system can also boost peer-to-peer trustless networks, and reduce intermediation, at the same time offer transparency and traceability of food distributing and processing records. This setup will be particularly helpful when a food outbreak occurs, then people can easily track down and block the problematic origins of the infection. However, the

[7] https://sproutsocial.com/insights/data/social-media-transparency/

[8] How technology and financial services can improve the efficient functioning of agricultural value chains. Mattern and Ramirez (2017)

original data input needs to be accurate and may require careful inspection. Credit activities and contracts histories in lending and trading can be stored on a blockchain, these records cannot be easily manipulated. This blockchain encrypted credit and contract setup can also eliminate intermediation in enforcement without the involvement of a trusted third party.

Especially during the early evolvement and publicization of blockchain technology[9], Bitcoin and crypto mania played a critical role, as the exponential price increase (and decrease) has attracted many speculators' participation, media coverage, and regulators' attention. Cryptocurrencies and tokens[10] are blockchain-based digital assets that are means of payment, fundraising, and reward. In other words, cryptos are the primary currency and investment tools in the blockchain ecosystem. People may wonder why not use US Dollar, Euro, Chinese Yuan, or any major Fiat currencies; it's primarily because blockchain is set up using digital cryptography that has the technology advantage in building and pairing up a virtual crypto-cash system that works for its own ecosystem. The distributed nature of blockchain requires a wide adoption and network effects to reach a level of efficiency, which takes time to develop and onboard, we don't yet see reliable data that can support empirical studies on blockchain adoption in agriculture, either in payment or supply chain. However, cryptocurrencies have reached a 1.696 trillion[11] US dollar market cap and are traded widely across the globe, in which we can obtain data from the major crypto exchanges. We thus take the approach to analyze the medium of exchange and fundraising tool -cryptocurrencies in the blockchain ecosystem first, to understand the current status, and evaluate potential use and policy implications in this particular angle of financial technology adoption in agriculture. Recent initiatives[12] include developing and piloting a blockchain based token system that is backed by warehouse receipts, in order to smooth the marketing liquidity upon harvest and provide a risk management tool for farmers in the relatively less efficient commodity trading market.

---

[9] For a history of the Bitcoin Blockchain and DLTs, see Iansiti and Lakhani (2017).

[10] We use cryptocurrencies and crypto tokens interchangeably, and this is common practice. Even though there's slight difference more on the technical side, it does not imply much differentiation on the financial studies in this dissertation context.
https://www.gemini.com/cryptopedia/cryptocurrencies-vs-tokens-difference

[11] As of March 11, 2021, Source: https://coinmarketcap.com/

[12] With Apurba Shee, PhD at University of Greenwich in UK, and STFC Food Network+

Underdeveloped commodity markets result in asymmetric information and distorted price signals that leads to inefficiencies in the market functioning role of transparent price discovery. Although warehouses are scattered throughout the region, supply-chain linkages are weak so that price transmission between one market and another are weakly correlated. This results in large swings in the price basis between markets. If information were symmetric price signals would indicate arbitrage opportunities moving commodity from one market to another to balance supply and demand.

To improve the supply chain between local markets and international markets, we propose to use blockchain technologies through the traditional warehouse receipt system as farmers deliver bags of (for example maize) to the warehouse the farmer is issued a receipt. Simultaneously a commodity token can be issued to the farmers and encoded into the blockchain. The transaction is then verified by users to store on the blockchain and cannot be easily changed / destroyed / stolen. The token has a unique hash and is tethered to the market price of the corresponding grain (e.g. refer to CME, AfMX, or DCE commodity futures price).

Within the blockchain, third party marketers can immediately identify the amount of grain available by the number of tokens reported. Local and global marketers can access the blockchain through a blockchain commodity exchange portal. The marketer can purchase tokens through any crypto exchange that listed this traded token (through initial coin offering, i.e. ICO), with each token representing a physical quantity at a specific location timely. The marketer in essence takes ownership of the warehouse receipt. The marketer can then arrange shipment to their location, or can hold on to the token and resell it at a later date through a crypto exchange. If resold, the token passes to a new marketer who now has possession of the commodity.

The blockchain token system serves as a clearinghouse. At first sale the marketer deposits cash into the blockchain through a global exchange. The hash attached to the token immediately passes from the digital wallet of the farmer to the digital wallet of the marketer. Cash, equal to the value of the commodity token is then converted out of crypto exchange and transferred to the farmer's hand. Through a web app the farmer's short cash position is transferred from the clearing house to his/her bank account. The shipping process can also be

operated through blockchain, for the ease of scalable integration and benefit of timely traceability. IBM and Maersk are already together piloting such use.

The blockchain does not only serve as a timely and secure database and supply chain management tool that provides transparency and traceability, but also connect the local agriculture community to the international financial market through the publicly listed commodity token on the crypto exchanges. This act will improve the accuracy and efficiency of database management and financial market accessibility. Through this foundational infrastructure, we can then integrate agriculture insurance and commodity futures, to better conduct risk management that eliminates market failure in geographic basis risk and temporal volatility (harvesting season crowd sell and off season hoarding).

In order to develop or understand how a blockchain token system works for agricultural value chains, commodity trading and risk management, we need to evaluate how the current tokens are functioning that may enlighten / warn us to better development of it. Therefore we studied over 3,000 cryptocurrencies that are currently traded on the exchanges. These crypto tokens are various projects applied in various purposes including finance, art, gaming, data, gamble, and a few in food industry. Through comprehensive measures of Hurst coefficients and Autoregressions, we find due primarily to the market structure of some coins especially their releasing mechanism, most of coins' market properties appear to be inefficient. This finding does not only support potential tokenomic design and regulation in food and agriculture business, but also assist broader areas of applications that utilize blockchain token system. In a related project[13] of ours that is not completely included in this dissertation, we have conducted extensive literature review (over 80 pages) on all economic related studies about Crypto Currencies and their applications in different industries, including 200 academic articles and many other industry reports, newsletters, and tweets[14]. Surprisingly we find very

[13] We have translated this material to Chinese, published on Financial Times Chinese Web.

[14] https://www.nytimes.com/2018/01/29/opinion/bitcoin-bubble-fraud.html Web. The NY Times https://twitter.com/paulkrugman/status/1014112633602101251?s=20 Twitter. @paulkrugman https://www.coindesk.com/bitcoin-outlawed-economist-joseph-stiglitz-says Web. CoinDesk https://www.yahoo.com/entertainment/bitcoin-bubble-perfect-example-faddish-145000767.html https://www.cnbc.com/2021/05/04/jpmorgan-ceo-jamie-dimon-im-not-a-bitcoin-supporter.html

little of them explain empirically why the tokens are traded inefficiently. Our work in this dissertation contributes to the understanding and advancement of cryptocurrencies and their applications in agriculture.

1.2. Farmland Markets in Minority Communities - Amish

Minorities and some religious groups rely heavily on agriculture as source of income and way of living. The discussion and initiative on improving financial inclusion for minority groups and reducing the inequality in financing resources especially rural agricultural areas have been one of the focuses among NGOs[15]. We study a different angle of minority groups, by looking at how they compete with the majorities and survive in farming. Even in the developed economy of the U.S., we still see Amish communities are mostly involved in farming, probably not only because of less access to other vocations, but due to their strong religious beliefs to highlight human-nature ties. Due to these beliefs, they use less modern agricultural technology to farm, such as they use horses and mules for draught power, rather than tractors or machineries. We assume the less technology use led to lower productivity and we find basic evidence showing they have lower yields for crop or less milk production per cow. Thus, we are interested in studying how Amish farmers survive16 with the consistently lower output and still actively growing and prospering in several states including New York.

Farm real estate makes up over 80 % of U.S. farm sector asset. About 25 percent of CaseIH.com survey respondents rank availability and price of land for expansion as having the most impact on their business[17]. Therefore, we use this most important and expensive input of farming– farmland, to study if and how Amish farmers can afford the land and compete with conventional farmers on farmland market. We conceptually evaluate the competition between Amish and conventional farmers and explain they may reach similar levels of profits.

---

[15] "G20 Roundtable on Innovations in Agricultural Finance" convened on September 9, 2015 in Antalya, Turkey by the Small and Medium Enterprise (SME) Finance Sub-Group.

[16] Chris Wayne, the FARMroots director at GrowNYC, has helped create close to 100 farm businesses, many of them owned by beginning farmers who are at an increased disadvantage in the marketplace, including immigrant and Indigenous farmers.

[17] https://www.farmprogress.com/management/farmers-list-top-issues-impacting-agriculture

Because Amish has larger family member and thus cheaper labor costs, also their higher saving habits can compensate the disadvantage of less output due to lower technology use. Because we are originally motivated by anecdotes that “Amish drive up land prices”, we empirically evaluate how Amish population growth impact the farmland prices. We construct the empirical model using a standard hedonic approach, with considering characteristics that may influence land prices. We include a shift-share like instrumental variable, an enclaving variable to address possible endogeneity from reverse causality and correlated unobservable. We find Amish population growth does not have statistically significant relationship with farmland prices. This in turn helps us align the conceptual with the empirical that Amish is competing with the conventional farmers on farmland in a similar manner, as they have similar profits despite the differences in revenue and costs.

This chapter also contributes to the literature of relationship between farm size and productivity. The earlier literature mostly concludes there’s an inverse farm size-productivity and thus small farms can survive without expanding, as they find small farms are more intensively managed and thus obtain higher yield, that are even more productive than the large farms that use advanced mechanical tools in farming with economies of scale. So, the less productive small Amish farmers who can survive do not fit in this earlier strand of literature. Instead, our conceptual argument and empirical findings align with the more recent literature, that claims using profit rather than yield to measure productivity is more realistic in evaluating size, technology adoption and survivability.

1.3. China's Livestock Insurance Supply via Field Choice Experiment

Agriculture insurance as the key tool to support the risk management of farmers, has recently been playing an even more important role in China’s agriculture development. Because the national goal of lifting all citizens to come out of extreme poverty and boosting village economy through agricultural finance innovations has led efforts to reform agriculture insurance products and expand insurance take-ups. Livestock insurance is less developed than crop insurance in China thus needed improvement urgently and comprehensively. Also, as African Swine Fever (ASF) outbreak since 2018 has damaged the small pig farmers and the

overall supply chain significantly, the livestock feeding industry has shifting to become more enterprise based, therefore more knowledgeable and demanding of livestock insurance. Because of the nature of agricultural risks that are systematically correlated for example hail damage or disease spread in a certain location, that are different from auto or home insurance that are much more random, insurance companies are thus generally not willing to conduct agriculture insurance without heavy government subsidy. However surprisingly we see very little research on the supply side of agricultural insurance, and especially for livestock insurance we seen none. Therefore, we study this very important matter as it emerges in China and other countries such as Mongolia as well.

China's livestock insurance has mostly been mortality insurance and just recently started to pilot newer type of insurance such as index and revenue insurance, whereas the U.S. and other developed countries have been offering revenue and even profit insurance for a long time. We wanted to study the attitude of China agriculture insurance agents on these various newer types of insurance and to evaluate their willingness to offer (WTO). We were motivated to conduct this study and quantitatively evaluate agents' WTO as hearing agents decreased / capped product offering under low or delayed subsidy, and also as we hear agents find index insurance very difficult to explain to the farmers and hard to conduct, despite the many benefits the literature is showing. In evaluating the effects of different attributes on utility or WTO, such as attributes like subsidy and insurance type, Discrete Choice Experiment with D-Optimal design combining and altering different attributes and levels in a set of choices has been proved to be a valid method.

We designed the insurance choices cards, having over 200 agents each to answer 15 cards that contain 2 choices each. The cards and choices include different combinations of various levels of eight main attributes of livestock insurance, that we selected based on extensive consultation with livestock insurance experts. Through conditional and mixed logit regression analysis on the in-the-field data, we find that 10% increase of subsidy lead agents 3.166 times higher probability to offer. We also find mortality insurance is still strongly preferred than the newer insurance, with index insurance being the least preferred, because agents have had very difficult time communicating with farmers about how it works and

especially challenged by farmers' complainants facing basis risk. Following the choice card experiment, we have asked the insurance agents with a series of questions in a following survey. That confirms the dominant preference on mortality insurance. Also, one of the survey questions asking about agents' knowledge regarding the newer introduced types of livestock insurance was used to generate the interaction term with the insurance type attribute in the choice card. The interaction term model reveals that more knowledge they have regarding a particular newer type of insurance, the more likely they are willing to offer that product. Therefore, we can see that training helps to understand and offer a new product. This chapter provides policy suggestions with quantitative measure especially on subsidies, as it is important to optimize the insurance take-up with the limited budget.

In this dissertation, we study and contribute to the understanding of these emerging issues in agricultural finance including Agri FinTech, Amish growth, and livestock insurance supply in China. We used diverse set of tools to analyze these issues: primarily time series econometrics in chapter 2 evaluating cryptocurrency prices properties; standard hedonic modelling for farmland values with using instrumental variable to identify the effect of Amish population growth in chapter 3; and discrete choice experiment with conditional and logit models to study China insurance agents' WTO livestock insurance products. These methods and approaches also contribute methodologically to the research in agricultural finance. Our exposure to these frontline topics uncovers fresh, debatable while also limited view of the role they play in further developing agricultural finance. We hope to initiate the proactive thinking, awareness, and analysis in the emerging issues that commences the innovative development in agricultural finance.

2. CHAPTER 2

Fractional Processes and Economics of Cryptocurrencies

Cryptocurrencies have seen extraordinary volatility and mixed growth over the past years, though there is a lack of empirical evidence on their price time-series properties and underlying mechanisms. This paper studies 3,351 cryptocurrencies' daily price data empirically to fill this gap. Assuming Brownian motion we use the variance ratio approach to estimating the Hurst coefficient for each crypto coin and find a wide range of fractional processes displaying both long and short memory. Despite these finding there is a possibility that the conventional Hurst measure does not take into account structural changes on time series, which may result in a misleading "long-memory" phenomena. To address this issue, we further use the proposed three-stage method to estimate and remove the time-varying mean caused by structural changes and re-estimate the Hurst effect of the de-mean process. Empirical results suggest strong evidence that long-memory does not exist in cryptocurrencies, because the existing literature mistakes structural change on the mean process for long-memory. The mean process is changing smoothly over time, because of capped supply and increasing demand structure. Another interesting finding is that most cryptocurrencies follow fractional Brownian motion, instead of a continuous random walk (e.g., geometric Brownian motion) that was assumed in some recent crypto studies, which violates the Efficient Market Hypothesis financial theory.

2.1. Introduction

2.1.1. Motivation

The increased preference for cryptocurrency as a unique asset class raises some significant economic questions. The irregular market structure of crypto currencies, coins and tokens (in this study we use cryptocurrency or cryptos for all) defies economic interpretation in several dimensions. Perhaps most peculiar is the unresponsiveness of

cryptocurrency supply in reflection to either price signals or demand. Despite this well-stated characteristic of the crypto market, there is an emerging trend among some economists to treat cryptos as a conventional asset class for which conventional utility, arbitrage, and equilibria apply.

More recently, theoretical and empirical studies on crypto markets are on the rise as cryptocurrency attracted enormous interest. To name a few: a recent article by Trimborn and Härdle (2018) creates an index CRIX for cryptocurrencies as an attempt to develop an investable and accurate benchmark. Borri (2019) studies conditional tail-risk exposure, the capability of investing and hedging for top cryptocurrencies. In our paper, we show that most of the 3,351 cryptocurrencies we examine do not satisfy the Markov property of geometric Brownian motion, but do satisfy the properties of a fractional Brownian motion.

Why (and how) systemic forms of memory can persist in any asset class is a perplexing economic problem. In conventional currency markets, it is normally assumed that supply and demand are endogenous to each other. We argue that one of the potential sources of memory in cryptocurrencies is because in most instances price is not an endogenous determinant of supply. Our reasoning is that the vast majority of cryptocurrencies - and in particular Bitcoin - are released according to a predetermined schedule so that at any moment in time the supply is perfectly inelastic. Consequently, the price swings and excursion patterns observed are due to non-fundamental (e.g. speculation, hype, emotion, bandwagon) demand forces, independent of supply. We argue that it is difficult to bind cryptocurrency demand to any meaningful economic force. For a currency to exist in any form it must satisfy the law of one price of which there is no evidence. Consequently, we offer a cautionary warning that volatility and valuations in crypto markets do not follow conventional or foundational economic approaches to fundamental value.

2.1.2. Background

The emergence of Financial Technology (FinTech) this past decade has inspired a series of technology inventions in the financial sector, to provide financial services and product delivery processes cheaper, faster, and more efficient than traditional systems. Blockchain is considered one of the promising and popular technologies among the FinTech sector and has led to many applications. Among them, cryptocurrency is by far the largest and most important application. There are over 10,000 cryptocurrencies actively traded globally (Coinmarketcap 2021a), with the market expected to continue to develop and grow.

Bitcoin, the first and leading cryptocurrency, has a 47.78% market share of the total cryptocurrency market as of late-June of 2021 (Tradingview 2021). Bitcoin was developed in 2009, it has attracted an abnormal amount of attention especially since 2017, because of the exponentially rising price. Bitcoin is a digital, decentralized, partially anonymous currency, originally not backed by any government[18] or other legal entity, and not redeemable for gold or other commodities. It relies on peer-to-peer networking and cryptography to maintain its integrity (Grinberg 2011). Bitcoin along with other cryptocurrencies has established its firm standing in the market, with a global market cap all time high of $2.54 Trillion on May 12, 2021, but this has since depreciated by almost 50% to $1.3 Trillion as of June 27, 2021 (Coinmarketcap 2021b). Nonetheless, there is speculation that cryptos could continue to rise to an 8 trillion-dollar market[19]. At the same time, some people question if it could totally collapse; no doubt cryptocurrencies remain debatable.

Bitcoin price movements in Figure 2.1 provide a rough sense of how volatile the market is. For it is what appears to be the first time in history individuals and institutions are actively trading an 'asset' that has no discernible fundamental value when looking at

[18] El Salvador will start adopting bitcoin as legal tender in Sept. 2021 after passing law in June 2021. Web: https://www.cnbc.com/2021/06/09/el-salvador-proposes-law-to-make-bitcoin-legal-tender.html
[19] https://www.forbes.com/sites/billybambrough/2020/02/04/wealth-manager-makes-massive-8-trillion-bitcoin-betbut-adds-a-serious-warning/

it from the traditional means. It is, so far as one can tell, the first existential asset to be traded in the sense that Bitcoin exists for no other reason than its existence, unlike stocks which are backed by the company or commodity futures which are tethered to an underlying physical asset or good. Though some later developed cryptos have features similar to early stage equity markets, categorized as security tokens. These and other asset-tethered coins and tokens are the exception and not the rule, yet are distinctively important since by Ito's Lemma they are an arbitrageable class that will follow a geometric Brownian motion (gBm) if the tethered asset is a gBm.

Seeing the hype, the argument of questioning Bitcoin as a bubble has been raised and discussed by some scholars. Dr. Robert Shiller has referred to Bitcoin as a bubble, on a CNBC interview in April 2018 he said: "To me, it's interesting as another example of faddish human behavior. It's glamorous. It reminds me of the Tulip mania in Holland in the 1640s" (Lahiff 2018). Dr. Paul Krugman posted on his Twitter saying "[Bitcoin] looks like a pure bubble" (Krugman 2018). Krugman was referring specifically to Bitcoin, and we would tend to agree given our finding that Bitcoin exhibits a bubble feature that includes persistent time-varying mean structure and short memory which indicates a fractional Brownian motion process. But we caution against 'bubble' analogies in general. We find that persistence - Hurst coefficients greater than 0.5 - is more of an exception than a rule. Indeed, we find that the vast majority of cryptocurrencies are ergodic, or mean-reverting with a Hurst coefficient less than 0.5. These currencies will generally have more high-frequency excursion patterns (lower stopping times with frequent reversals) than what might be considered a 'bubble'. An extreme example is with Tether - a coin that is tied to the US dollar also called stablecoin- which has (and should have) a Hurst coefficient near zero. This is a pure white noise ergodic process.

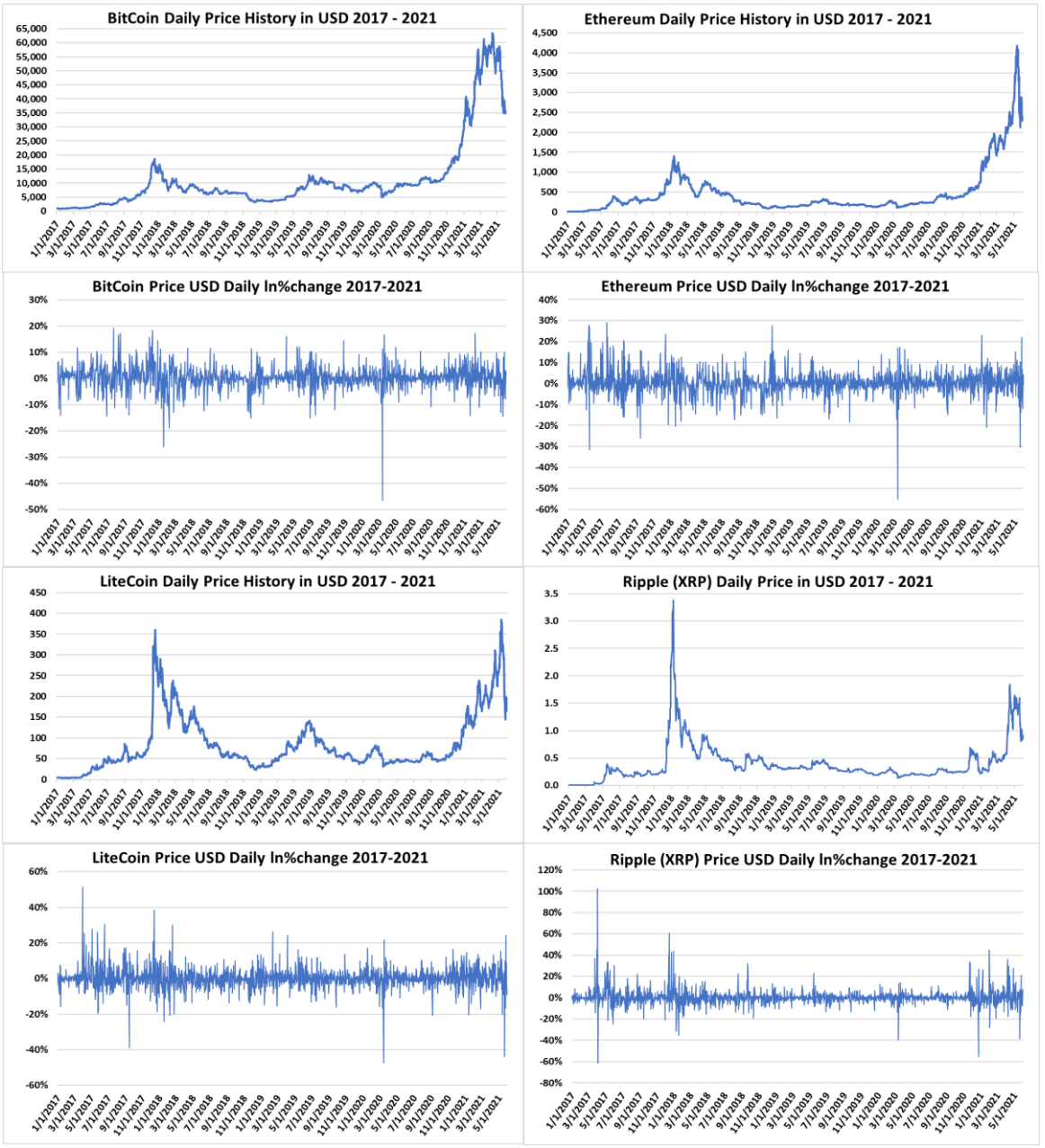


Figure 2.1: Top 4 Crypto Coins Daily Price Series and Natural Log Change

2.1.3. Purpose

This paper provides a foundational understanding of the background, market structure, and price series properties of cryptocurrencies. Examples of these seemingly erratic time series for Bitcoin, Ethereum, Litecoin and Ripple are provided in Figure 2.1. More specifically, we test for Brownian motion using the scaled variance ratio Hurst measure and its enhanced method with adjustment to time-varying mean. A fractional Brownian

motion (fBm) is a memorized stochastic process with multiple roots, one of which is a unit root. Only a geometric Brownian motion (fBm) satisfies the Markov property. A gBm has only one root, a unit root, with a Hurst coefficient of 0.5. An fBm has multiple roots and a Hurst coefficient greater than or less than (and perhaps equal to) 0.5. Our analysis reveals that 1,140 of 3,351 cryptocurrencies fall within the 95% confidence intervals required of a Brownian motion. The numbers increase to 1,396 when the confidence intervals are expanded to 99%. Cryptocurrencies outside of these ranges do not have a unit root and are wildly non-stationary.

Using the scaled variance ratio Hurst measure, we first show a wide range of memory systems including long (persistent) and short (mean-reversion) memory, which is rarely found in the securities, commodities, and many other financial markets. Our analysis provides a more comprehensive view than existing crypto literature described in the next section that mostly concludes long memory for top cryptocurrencies. Digging deeper, we apply the nonparametric estimation method to detect time-varying mean, and then again measure the Hurst effect after de-meaning. We reveal that some previously found long memory was inaccurately interpreted because of a time-varying mean resulting from built in or external structural changes. The actual causes are subject to future research. Nonetheless, our results raise serious questions about the efficiency of cryptocurrencies as a conventional asset class, considering the predictiveness and underlying structural changes. Because Brownian motion and efficiency are crucial in asset pricing and practice, we aim to provide a coherent analysis to empirically verify the issue and communicate the importance of awareness on it.

Besides extensive analysis in fractional properties and time series models, this study will briefly explore the existing literature to shed light on bubbles, price discovery, categorization, risk management, possible influence on economics and financial theories, impact on monetary policy and financial technology development, which can inspire further research. Our approach here is to study more in-depth the properties of fractional Brownian motion to reveal unique findings including the possibility of ‘bubbles’.

This paper first investigates the market structure of cryptocurrency, that have deflated supply releasing mechanism (majority), and we use Bitcoin as an example. Then we examine certain stochastic relationships and geometric/ fractional Brownian motion in cryptos time series, using Hurst measure, time-varying mean estimation, and stepwise autoregression. We use primary real-time data of cryptocurrency time series from publicly available and subscribed sources. Deploying an autoregressive process, we find that the vast majority of cryptocurrencies satisfy the conditions for a fractional Brownian motion. Using the Hurst measure as a dependent variable we use regression to examine the relationship between various attributes of coins/tokens and show that no fundamental value beyond the volume of coins outstanding explains market risk.

In the remaining sections, section 2 reviews related literature on cryptocurrencies and random walk methods. Section 3 explains the primary releasing mechanism, market structure, and supply-demand relationship of cryptocurrencies with a specific focus on Bitcoin as an example. Section 4 lays out the fundamental theory, implications, and underlying mathematics of geometric and fractional Brownian motion. Section 5 illustrates the data, methodology, and empirical findings that imply the fractionality of cryptocurrencies. Finally, section 6 discusses future work and concludes the study.

2.2. Literature Review

The origins of cryptocurrency date to the financial crisis in 2008, which raised questions about the stability of a centralized currency system and the integrity of financial institutions. A pseudonymous brain trust called Satoshi Nakamoto published a white paper that explained a solution to the currency problem, with Bitcoin using a blockchain-based peer-to-peer electronic transaction system (Nakamoto, 2008). The system is based on cryptography and takes computer programming to conduct verification and expansion for what is referred to as Proof-of-Work and Proof-of-Stake. There have been many technological studies about how it functions and what benefits the system brings to technical efficiency. These include Saleh's (2017) and Biais et al.'s (2017) analysis on mining and minting games, Juels' (2018) work on secure dynamic access structures for

cryptocurrencies, and Eyal and Sirer's (2014) claim of Bitcoin mining's vulnerability. We focus on summarizing the economics and finance studies that are more related to this study.

There are several strings of literature, one being in the market microstructure area especially in the transaction fees domain: Basu, Easley, O'Hara, and Sirer (2018) and Lavi, Sattah, and Zohar (2017) observe the inefficiency, instability, and unpredictability of cryptocurrency functional fee structure and propose mechanisms that could improve the efficiency. Basu et al applied a generalized second-price auction focusing on maximizing social welfare, while Lavi et al. assumes a single monopolistic miner who maximizes fee revenues at the cost of social welfare. Huberman, Leshno, and Moallemi (2017), Cong, He, and Li (2018), and Easley, O'Hara, and Basu (2017) study miner revenue, transaction fees, mining efficiency, and market structure.

Another string of literature tries to categorize bitcoin or cryptocurrency as a specific asset, either it's commodity, currency, or securities. Kaplanov (2012) studies Bitcoin's nature more in thinking of a medium of exchange and argues its legitimacy of serving as a community currency, also proposes regulations the U.S. federal reserve may take instead of widely prohibiting it. Dyhrberg (2016) analyzes the relationship between Bitcoin, gold, and the US dollar, the study implies that Bitcoin can be classified as something in between gold and the US dollar. Though a replication and extension study done by Baur et al (2018) has applied similar methods with more timely data, which states Bitcoin displays distinctively different return, volatility, and correlation characteristics compared to other assets including gold and the US dollar. Klein et al (2018) further enhanced the position of not too reasonable in categorizing Bitcoin as gold or other commodities. Baur et al (2017) and Burniske and White (2017) study returns on bitcoin comparing to other types of assets including currencies, equity, bond, precious metals, and energy products, with the result showing they have consistently low correlations among asset groups. Grinberg (2011) questions the categorization of Bitcoin among stocks, investment contracts, commodities, and currency, thus has imposed the question of its legalization path as well as the market volatility associated with it.

Studies on fundamental value, bubble, volatility, and pricing have also largely emerged in the literature. Cheah and Fry (2015) state Bitcoin market is highly speculative and volatile, which is subject to speculative bubbles with fundamental value hardly definable or near zero. Though on the opposite view, Blau (2017) states that the high volatility of Bitcoin is not related to speculative trading when looking at the first large price wave in 2013. Buchholz et al. (2012) find that Bitcoin price movements to a large extend can be explained by interactions between its supply and demand. Koutmos (2018) focuses on the transaction demand for bitcoin by utilizing bivariate vector autoregression (VAR) models to show the strong linkages between Bitcoin returns and transaction activity. Ciaian et al (2016) integrated previous methods of studying the formation of bitcoin prices, using 2009-2015's bitcoin daily time-serial data to assess the determinants of bitcoin price formation. Kristoufek (2015) utilizes continuous wavelet analysis, specifically wavelet coherence, to identify the possible sources of bitcoin price movement ranging from fundamental economic sources to speculative and technical sources. Hayes (2016) suggests that Bitcoin does indeed have a quantifiable intrinsic value and formalizes a pricing model based on its marginal cost of production mainly considering mining supply and market demand. Wheatley et al (2018)'s research on the predictability of bitcoin uses generalized Metcalfe's law and the Log-periodic Power-law Singularity (LPPLS) model to analyze the bubbles, crashes, and pricing of bitcoin. Cheung et al (2015) use a well-established methodology Phillips-Shi–Yu (2013) with Mt. Gox bitcoin prices to identify and evaluate bubbles. Thiesa and Molnár (2018) applied Bayesian change point analysis to investigate the average return and volatility of the Bitcoin price.

Literature on risk and hedging focus on comparisons with existing financial assets. Dyhrberg (2015a) recommends that bitcoin has a place in the financial markets and portfolio management. In turn, Katsiampa (2017) suggests examining its volatility is crucial given the emerging market capitalization and its enhancing role in the financial market. Dyhrberg (2016) analyzes the volatility of Bitcoin using GARCH and finds that it may have a positive time trend and shows non-stationarity. Enders (2010), on Bitcoin, concludes that the most noticeable stylized fact is that volatility variability, contains

periods of very high volatility and also relative tranquility, which is confirmed using Engle's Lagrange multiplier showing a strong ARCH effect in the first differenced logged bitcoin price residuals.

Further, Katsiampa (2017) uses various GARCH models to compare and see which one has the best goodness-of-fit to Bitcoin price data; and concludes with evidence from AR- CGARCH[20] that it is crucial to consider both short-run and long-run components of conditional variance. Klein et al. (2018) find asymmetric responses to market shocks in Bitcoin returns and finds FIAPARCH[21] to be the best fitting model for Bitcoin conditional volatility measuring. Gkillas and Katsiampa (2018) use a timely dataset of 5 major cryptocurrencies Bitcoin, Ethereum, Ripple, Bitcoin Cash, and Litecoin, and employed extreme value theory to investigate the tail behavior of the returns of these cryptocurrencies. Bouri et. al. (2017) uses a dynamic conditional correlation (DCC) model to examine Bitcoin across major world stock indices, bonds, oil, gold, the general commodity index, and the US dollar index to see if it acts as a hedge, diversifier or safe haven. Urquhart et. al. (2018) investigates whether Bitcoin can act as a hedge or safe-haven against major currencies at the hourly frequency level, capturing intraday large volatility that Bitcoin experiences. Dyhrberg (2015a) claims that Bitcoin's hedging capability is somewhere between gold and the US dollar, though later arguing Dyhrberg (2015b) that Bitcoin can act as a hedge against UK equities and the US dollar.

The main string of studies most closely related to our approach are those that focus on long memory, structural breaks, and Hurst estimation. Mensi et al (2018) conducted a study on Bitcoin and Ethereum prices to identify the impact of structural breaks on dual long memory, using four different generalized autoregressive conditional heteroskedasticity models (GARCH, FIGARCH, FIAPARCH, and HYGARCH). Several studies have computed Hurst coefficients for Bitcoin and discuss geometric fractional Brownian motion. These include Tarnopolski (2017), Jiang, He, and Ruan (2018), Al-Yahyaee, Mensi

[20] AR-CGARCH is Autoregressive-Component generalized autoregressive conditional heteroskedasticity which captures both short-run and a long-run component of the conditional variance.
[21] FIAPARCH is Fractionally Integrated Asymmetric Power autoregressive conditional heteroskedasticity.

and Yoon (2018), Caporale, Gil-Alana, and Plastun (2018), and Bariviera et al. (2017). They conclude as we do that fractional properties exist, but these studies are rather short mechanical tests for certain cryptocurrencies such as the top 3. We explore the generality of these findings to include 3,351 traded cryptocurrencies, reveal the general properties in their price series, and then link these properties to underlying market structure to study how certain influencing factors affect supply and demand.

2.3. Cryptocurrency Releasing Mechanism and Supply/Demand Equilibrium

2.3.1. Blockchain and Cryptocurrency

The electronic coin is a chain of digital signatures. According to Wikipedia, a digital signature is a mathematical scheme for demonstrating the authenticity of digital messages or documents. A valid digital signature gives a recipient reason to believe that the message was created by a known sender (authentication), that the sender cannot deny having sent the message (non-repudiation), and that the message was not altered in transit (integrity) (Digital 2018). One of the major contributions this system has is to solve this problem of double-spending, in which the recipient of a bitcoin has to be guaranteed that the coin was not simultaneously sent to another recipient at the same time. This mechanism relies on the development of a timestamp that keeps track of each coin at each instant in time. The encryption includes the current timestamp and all previous timestamps linked in what is referred to as a blockchain. The more that a coin is represented in a blockchain the more difficult it would be to double-count, sabotage, or steal the coin because, in order to alter its history, including ownership, the attacker would have to recreate a block and all blocks that follow it, requiring immense computing power to overtake the natural block construction. The security is solid because new transactions are broadcast to all nodes, building new blocks need to go through proof-of-work then accepted by all nodes, then this process is conducted through granting digital signature from among consensus. The detailed algorithm of this network can be found in Nakamoto's white paper. Later smart contracting and proof-of-stake were developed to

improve the networks, and the logic of how blockchain and cryptocurrency generating process is still similar. Take Bitcoin as an example, once created, enter a labyrinth of linked computers and networks which allow for anonymous and irreversible transactions on a peer-to-peer basis then entered in the blockchain.

2.3.2. The Role of Miners and Mining in Blockchain Development

Cryptocurrency enters the market through a network of miners i.e., computer programmers, who ensure the authenticity of information and update the blockchain with the transaction under the decentralized consensus of blockchain nodes that are represented by participants around the globe, either individuals or institutions. These miners get rewarded with the cryptocurrencies by providing such cryptography verification to help expand and develop blockchain. Their rewards' worth may vary especially during recent years of high volatility. The reward may seem tempting and the worry of unlimited attempts to mine may occur, so the blockchain community agreed to set certain rules to release coins at a constant or decreasing rate from one period to the next with a fixed amount of the release designed to maintain the value of the coins and make mining more competitive. The Bitcoin creator community consensually decided to set a limited total supply of 21 million Bitcoins and cumulatively release Bitcoins at a decreasing rate from time to time: The amount released in one period (e.g. 1 year) is ½ of the amounts released in the previous period. Once the 21 million mark is hit the supply will be perfectly inelastic at a fixed point so that any increase in demand via a demand shift will clear the market at a higher price and any decrease in demand will clear the market at a lower price. Compare this to open market operations of a central bank which can increase and decrease money supply as needed in order to maintain exchange. That is an increase in the demand for money, signaled by an increase in interest rates, for example, can be offset by an increase in supply. Credit facilities will become more inelastic as market risks increase dampening the demand for credit and money, with elasticity returning as risk diminishes. The central bank will moderate inflation and economic growth through these credit markets by reducing supply at the federal reserve discount

window, reducing or increasing the money supply, or through moral suasion and other methods to get money into circulation or out of circulation.

Cryptocurrencies have no such moderating impact on supply, interest rates, and inflation, etc. When bitcoin hits 21 million coins, the supply of bitcoin will neither increase nor decrease. If the cumulative demand for bitcoin increases at a rate greater than its declining entry into the cybermarket, the price will naturally increase at an increasing rate providing the first ingredient for a bubble. But this is a structural characteristic, not an emotional one. Emotion simply accelerates the process and adds a second ingredient to the bubble. Figures 2 and 3 show the never decreasing price in nature from the current supply structure and demand pattern. This has given rise to a flurry of new alternative currencies that compete with bitcoin, so in effect, the supply of cryptocurrency will actually be increasing. Bitcoin, Ripple, Dogecoin, Ethereum, Litecoin, and thousands more have been issued. So, while the overall supply of a particular coin such as bitcoin may be limited, the total supply and market capitalization of cryptocurrency will forever be increasing but never decreasing.

Cryptocurrency is an asset in its own classes. Currencies are invented by individuals, computers, and the CPU, but it is the irreversibility of supply that distinguishes bitcoin from all others. Commodities respond in elasticity to supply and demand, conventional currencies can expand, or contract as needed to support government policy objectives, stocks can be issued or repurchased as the situation warrants. But there is no mechanism to reduce the supply of cryptocurrencies as economic conditions might warrant. On top of this, commodities, stocks, and conventional currency are ground to some tangible fundamental value whereas it is not clear at all what the fundamental basis of cryptocurrency is since the vast majority of purchases beyond the deep web can be purchased in cash, credit card, debit card, PayPal, or mobile transfers. Without any fundamental grounding in the real economy, cryptocurrency price fluctuations are ubiquitously determined by human emotion, or madness depending on the point of view, with a design that exacerbates volatility and guarantees a structural bubble with any increase in demand. In fact, as will be illustrated here the only circumstance in which the price of Bitcoin or other cryptocurrencies can sustain a stable price is if the rate at which

demand increases, actually decreases with increased scarcity, an economic outcome that defies direct observation and any axiom of economics known to man.

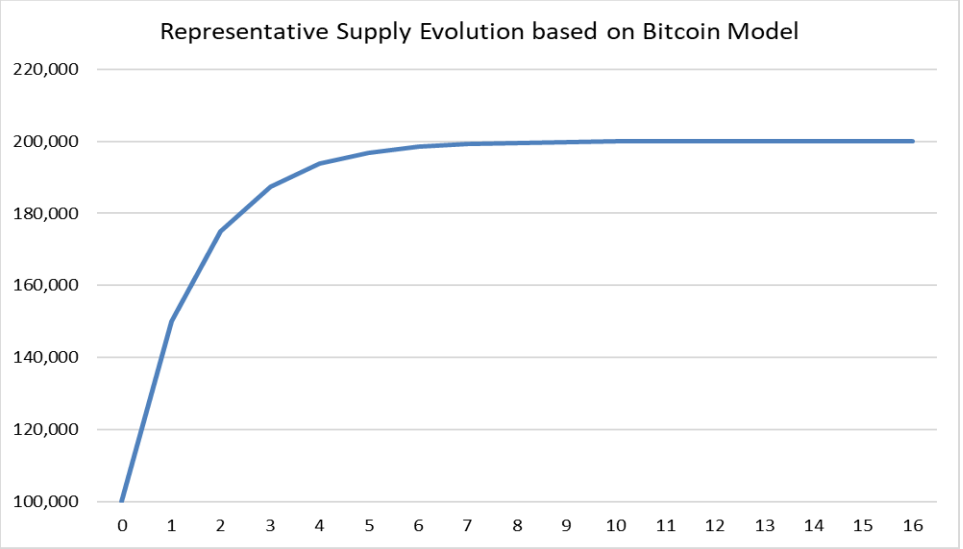


Figure 2.2: Theoretically Simulated Bitcoin Total Supply

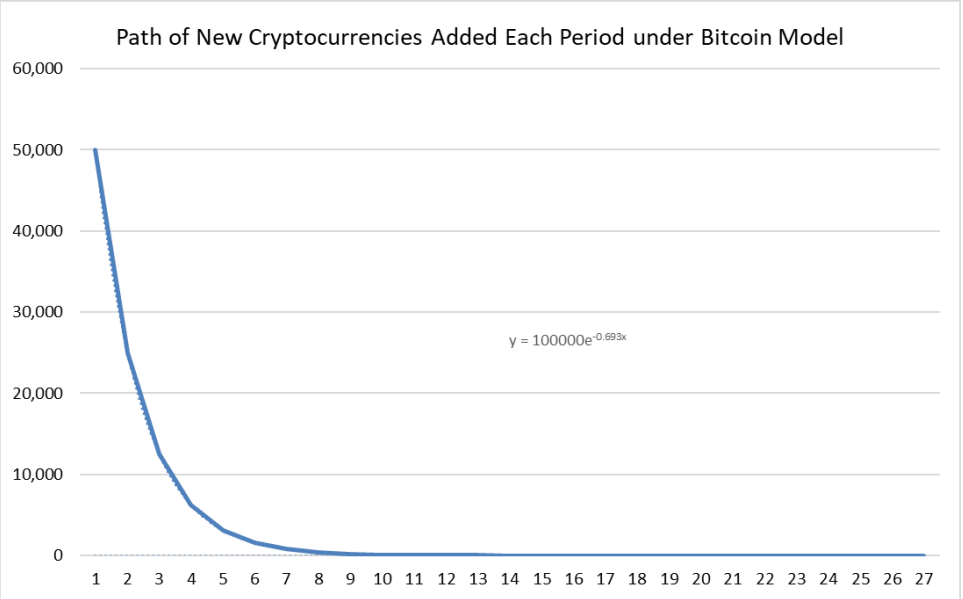


Figure 2.3: New Addition to Supply following the Bitcoin Model

Thus, understanding the supply side of the story is a crucial starting point. As previously mentioned, Bitcoin is released according to a schedule in which the amount received in one period (say a year) is one-half of the amount released in the previous period. Based on the public knowledge about the supply release of Bitcoin, we derive the characteristic equation in (2.1) to describe its time path.

$$Y^{S}_{T} = Y_0 \left(1 + \Sigma^{T}_{t=1} \; e^{-0.693t}\right). \qquad (2.1)$$

The price formula can be used to determine the projected supply at any moment in time T. This is illustrated in Figure 2.2 and notated year by year. Figure 2.3 demonstrates (not actual) how the Bitcoin model evolves by schedule. The total supply increases but at a decreasing rate, Bitcoin shown as an example, the majority of crypto coins follow a similar capped total supply structure. Bitcoin suggests that its 21 million coins will max out around 2040, but with the exponential decay of new coins being added, as shown in Figure 2.3, at a rate of 0.693 it is self-evident that the bitcoin max will be close to its limit well before that date.

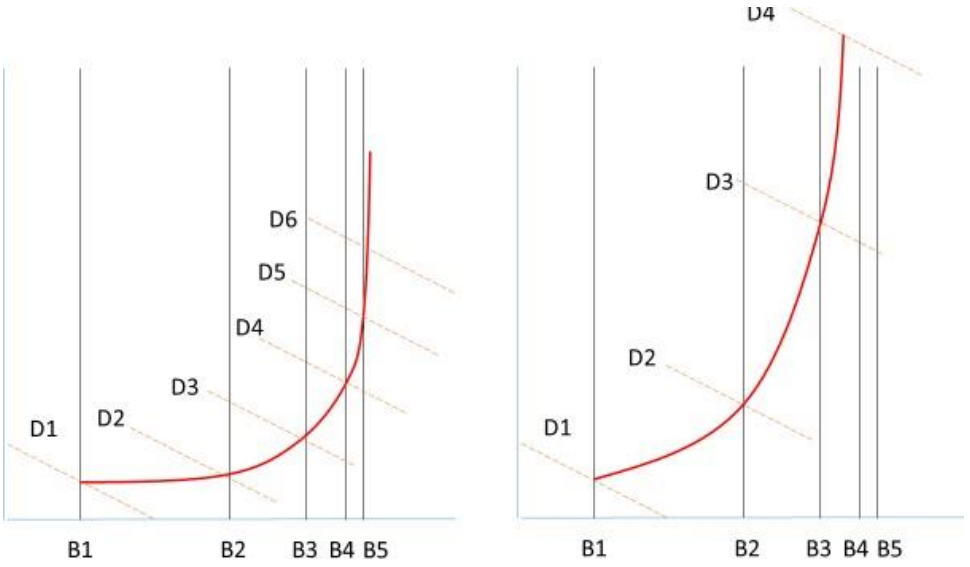


Figure 2.4: Demand and Supply of Cryptocurrency under the Bitcoin Model

The peculiar disposition of Bitcoin leads to peculiar economics which is depicted in Figure 2.4. Since supply cannot be reversed and does not increase with demand or any other signal other than its scheduled release it must be perfectly inelastic at each moment in time. Supply is thus increasing along the x-axis and is represented by the vertical line. Note that the space between supply curves decreases by half the previous period.

Price is therefore determined by demand. In the left panel demand (D1-D6) increases at a constant rate while on the right panel it increases at an increasing rate (D1-D4). The result is that with hyper-demand the price has no choice but to increase, and this it will do at an increasing rate. So, there are two drivers of price. The first is the structural

supply and the second is the emotion of demand. What is observed in the Bitcoin and crypto coin market is a rapidly increasing price. So long as more and more people want to enter the market, the structural characteristics are guaranteeing a bubble-like circus. Only when people tire of the market and start to exit, taking trade volume and liquidity with them, will prices fall systematically. And when they do, it will be a classic crash. When will this happen? Nobody knows, but inevitably it will happen.

Specifically, we propose demand equation as

$$Y^{D} = A * P^{-B}, \qquad (2.2)$$

then set the $Y^{S} = Y^{D}$ *to* solve for price and get

$$P = (Y_0 / A) * (1 + \sum_{t=1}^{T} e^{-0.693t})^{-1/B}. \qquad (2.3)$$

Looking at the derived equation, we can see that supply doesn't respond to price, which means the supply/demand price equilibria relationship here for Bitcoin is not endogenous. Empirically, we used semi-weekly supply and price data for Bitcoin of last 10 years run simple OLS linear form, find:

$$\widehat{P} = 0.001699 * \widehat{Y} - 19646.66. \qquad (2.4)$$

In this regression, 27.5% of the variation is explained and statistical significance is found. Through this very basic regression result, we observe an abnormal relationship, where higher supply results in a higher price; which is opposite from the classical economic theory. With non-price endogenous supply market structure as part of the crypto design and the abnormal empirical findings on supply to price relationship, one would suspect irregular price properties, including long-term memory.

2.4. Model and Theory of Random Walk Process

2.4.1. Geometric / Fractional Brownian Motion & Efficient Market Hypothesis

Generally speaking, commodity derivatives, stocks, currencies, and other tradable assets follow a random walk process. This is generally consistent with a stochastic differential

equation and geometric Brownian motion (gBm). This is a memoryless system consistent with a Markov process.

If the independence axiom is violated the process is referred to as a fractional Brownian motion (fBm). By fractional, it means there is some memory in the system. Mean reversion, ostensibly captured by predictable rises and falls in prices in a somewhat predictable way is a fractional process with a negative correlation between one or more price changes. Thus, as prices are rising it is increasingly likely that the prices will reverse themselves with some measurable probability in the future. The predictable behavior of day traders, for example, might cause frequent rises and falls, or reversals in asset prices. In such mean-reverting fBm, price excursions are short-lived thus have a rather short excursion length than a random walk, i.e. gBm. A persistent fractional Brownian motion is said to have positive memory. This is a self-reinforcing memory in which a rise or fall in prices is, in probability, going to correlate with a rise or fall in prices at some future date. Insider trading about the release of common stock from treasury or some other structural element of the market that is in all appearances random, but actually deterministic and predictable, can result in a positive intertemporal correlation. The fractional market is not efficient because it is (to some extent) predictable and thus violates the Efficient Market Hypothesis (EMH) generally assumed in financial theories.

The financial economics theory of efficient-market hypothesis (EMH) claims that traded financial products' prices fully reflect all available information. EMH is usually defined in three forms based on the categories of information known by each group of participants: Weak- form, Semi-strong-form, and strong form efficiency. An intuitive implication of EMH is that participants cannot "beat the market" consistently on a risk-adjusted basis, because market prices should only respond to new information or market shocks (Efficient 2019). Considering the incentive and ability to profit from private information, self-interested participants are motivated to obtain and act with private information. With many traders and participants attempting to earn profit from various ways of reflecting their information, these conducts contribute to reaching more and more efficient market prices. In a competitive market setting, traded prices reflect all

available information, and prices can only react to new market shocks. Thus, there is a very close link between EMH and the random walk hypothesis which states the new price is the latest price with a deviation caused by a random shock (Kirman 2009).

2.4.2. Traditional Hurst Coefficient Estimation & Implication in gBm / fBm

Random walk and Brownian motion process can be expressed in form of stochastic differential equation:

$$dx = \mu x dt + x\sigma dZ, \tag{2.5}$$

where $dZ = \varepsilon\sqrt{t^{2H}}$ is a Weiner process following a power law in H. Then x is the asset price; μ is the instantaneous change in prices, and σ is the standard deviation of the percentage change in prices. The Hurst coefficient is used as a measure of long or short memory in time series. It relates to the autocorrelation of the time series. Therefore, the Hurst coefficient, H can be used to imply random walk and its deviations, as a random walk is considered to have no memory in time series and a non-pure random walk can be regarded as having memory. H is commonly estimated by rescaled range (R/S) analysis, to examine the dependence of the rescaled range of N observations, divided into shorter periods.

When evaluating the dependence among period, we look at variance and covariance between time periods:

$$E[x(t_2) - x(t_1)]^2 = \sigma^2(t_2 - t_1)^{2H}, \tag{2.6}$$

$$E\{[x(t) - x(0)][x(t + \Delta t) - x(t)]\} = \frac{1}{2}\sigma^2[(t + \Delta t)^{2H} - t^{2H} - \Delta t^{2H}]. \tag{2.7}$$

We can write the Brownian motion variance in general form as:

$$Var[x_{\mathrm{T}} - x_1] = Var[x_t - x_{t-1}]T^{2H}. \tag{2.8}$$

With the supply structure and the emotional hype in demand, we suspect that the Bitcoin market is fractional and persistent, and other cryptocurrencies may appear to have memory too, either persistent or mean-reverting. The supply characteristics are

known in advance, so as long as prices are rising, and new investors are entering the market the indefiniteness with which prices rise. A convenient means of estimating system memory is the Hurst coefficient that's described above. For $H = \frac{1}{2}$ the system is geometric Brownian motion, for $H \neq \frac{1}{2}$ the system is fractional Brownian motion where the process contains memory. Specifically, when $H < \frac{1}{2}$ the system is mean-reverting, and for $H > \frac{1}{2}$ the system has a long memory and is said to be persistent. The boundaries at 0 and 1 are rarely met in nature with 0 being pure white noise and 1 being almost perpetually reinforcing. Here we will skip the details of proof except to say that for a random variable x time series, the Hurst coefficient can be estimated using:

$$H = \frac{1}{2}\frac{Log\left(\frac{Var[x_T - x_1]}{Var[x_t - x_{t-1}]}\right)}{Log(T)}. \tag{2.9}$$

The variance ratio in the numerator will collapse to T if there is no correlation or covariance. In this case, the numerator and denominator cancel out and $H = \frac{1}{2}$, we will have the relationship:

$$Var[x_T - x_1] = T * Var[x_t - x_{t-1}], \tag{2.10}$$

is a geometric Brownian motion. But in another case if:

$$Var[x_T - x_1] > T * Var[x_t - x_{t-1}], \tag{2.11}$$

then there must be some positive correlation between daily price changes and the numerator will have some value $T + n$ so that the variance ratio will be greater than 1 and $H > \frac{1}{2}$. Likewise, if

$$Var[x_T - x_1] < T * Var[x_t - x_{t-1}], \tag{2.12}$$

there must be some negative correlation between daily price changes so that the variance ratio in the numerator will have some value $T - n$, and $H < \frac{1}{2}$. We estimate the Hurst coefficients of cryptocurrency price series in the next section use the method illustrated above.

2.4.3. Time-varying Mean Estimation, Structural Breaks, and De-Mean Hurst Effect Enhanced Method

It is widely acknowledged that the financial time series always suffer from some structural breaks or smooth structural changes due to technological progress, preference change and policy switch, etc. Indeed, as Hansen (2001) pointed out "it may seem unlikely that a structural break could be immediate and might seem more reasonable to allow a structural change to take a period of time to take effect". For example, financial institutions may react to shocks, including monetary policy switch and technology changes, in a gradual manner. Even though the change of a financial time series may be abrupt at the individual level at discontinuity-time points, it may exhibit evolutionary changes after aggregation.

Structural changes always arise in various and crucial difficulties in empirical finance. For example, if a time series process includes a time-varying intercept, which is changing smoothly over time, this nonstationary phenomenon is likely to behave a characteristic similar to long-memory. Thus, this process may be mistaken for a long-memory process, resulting in false inference, unreliable forecasting, and misleading policy recommendations. Indeed, the long-memory process in bitcoin has drawn increasing attention in the existing literature: Mensi et al (2018), Tarnopolski (2017), Jiang, He, and Ruan (2018), Al-Yahyaee, Mensi and Yoon (2018), Caporale, Gil-Alana and Plastun (2018) and Bariviera et al. (2017) are examples previously discussed in section 2. However, it has been frequently noticed that using the original Bitcoin time series in a given trading period to examine the Hurst coefficient may ignore the importance of structural changes and lead to misleading results. Therefore, it is important to remove the structural changes of the original time series before Hurst effect estimation.

In this paper, we propose a three-stage procedure to examine whether long memory exists in cryptocurrencies. In the first stage, we use a test proposed by Hong et al. (2017) to examine whether the distribution of financial time series is changing over time.

In the second stage, we propose a two-step nonparametric method to remove the smooth structural changes in time series and obtain a new time series. Smooth structural changes can be estimated by parametric models, including smooth transition regression

developed by Lin and Terasvirta (1994). However, there is no economic and financial theory to guarantee the concrete functional form for time-varying mean, which is likely to result in misspecification. We follow the spirit of Robinson (1989) and Chen and Hong (2012) to use a nonparametric model with time-varying mean as follows: $x_t = \mu(t/T) + \varepsilon_t$, where μ: [0,1] → R is unknown smooth function except for a finite number of points on [0,1] with structural breaks.

First, for every time point t, we develop a local cross-validation criterion to select the optimal bandwidth, which is changing over time. Define a "leave-b-out" estimator:

$$\widehat{\mu}_{-t} = e_1'(\sum_{s=t-[Th_t],s\neq t-b+1,\cdots,t+b-1}^{t+Th_t} K_{st}Z_{st}Z_{st}')^{-1}(\sum_{s=t-[Th_t],s\neq t-b+1,\cdots,t+b-1}^{t+Th_t} K_{st}Z_{st}x_s) \tag{2.13}$$

Where $Z_{st} = (1, \frac{s-t}{T})'$, $K_{st} = K(\frac{s-t}{Th_t})$ with the prespecified symmetric kernel K (·): $[-1,1] \to \mathbb{R}^+$, and $e_1 = (1,0)'$. Then a data-driven bandwidth is selected from minimizing the local cross-validation criterion as follows:

$$h_{t,CV} = \arg \min_{c_1T^{-1/5}\leq h_t\leq c_2T^{-1/5}} CV(h_t), \tag{2.14}$$

where $CV(h_t) = \sum_{t=1}^{T} K_{st}(x_s - \widehat{\mu}_{-t})^2$, and c1 and c2 are two specified constants.

Second, the time-varying mean estimator based on the optimal bandwidth $h_{t,CV}$ is obtained by minimizing the local sum squared errors

$$\begin{aligned}
\widehat{\mu}_t &= \arg\min_{\mu_t} \sum_{s=t-[Th_{t,CV}]}^{t+[Th_{t,CV}]} K_{st}[x_s - \mu_0 - \mu_1(\frac{s-t}{T})]^2 \\
&= \arg\min_{\mu_t} \sum_{s=t-[Th]}^{t+[Th]} K_{st}(x_s - (\mu_0, \mu_1)'Z_{st})^2 \\
&= e_1'(\sum_{s=t-[Th_{t,CV}]}^{t+Th_{t,CV}} K_{st}Z_{st}Z_{st}')^{-1}(\sum_{s=t-[Th_{t,CV}]}^{t+Th_{t,CV}} K_{st}Z_{st}x_s),
\end{aligned} \tag{2.15}$$

where μ_j is the coefficient for $(\frac{s-t}{T})^j x_s,\ j = 0, 1.$

In the third stage, we use Hurst coefficient estimation developed in Section 4.2 to examine whether the long memory exists in the de-mean time series $\{\widehat{x}_t = x_t - \widehat{\mu}_t\}_{t=1}^{T}$.

2.4.4. Stepwise Autoregression and Quasi-Fractional Brownian Motion

In this section, we discuss the relationship between a standard autoregressive process and fractional Brownian motion. Following Turvey and Wongsasutthikul (2016), a quasi-fractional Brownian motion can be modeled using an autoregressive approach. The notation AR(p) indicates an autoregressive model of order p. The AR(p) model is defined as:

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p}, \qquad (2.16)$$

where a_1, a_2, …, a_p are the parameters (AR coefficient) of the model, and ε_t is white noise. An autoregressive process of order p>1 can have multiple complex roots, but if $\sum_{i=1}^{p} a_i = 1$, then one of those roots will be a unit root, satisfying the stationarity property in differences. With finite variance and multiple roots, the system is non-Markov and indicative of a fractional Brownian motion. The detailed proof of the statement and simulation process can be found in Turvey and Wongsasutthikul (2016)'s paper, the core part of the proof procedure is as follows:

Stationarity of above defined AR(p) process implies:

$$E[Y_t - Y_{t-1}] = E[Y_{t-1} - Y_{t-2}] = \ldots = E\left[Y_{t-q} - Y_{t-q-1}\right] \quad , \qquad (2.17)$$

Moving the terms of two sides, then we have:

$$E[Y_t - Y_{t-1}] = (a_1 + a_2 + \ldots + a_q) E[Y_t - Y_{t-1}] \, , \qquad (2.18)$$

$$\sum_{i=1}^{q} a_i = 1. \qquad (2.19)$$

Based on the above setup, finding the one time-step increment variance and derive the variance ratio formula can get:

$$\frac{Var[Y_t - Y_{t-k}]}{Var[Y_t - Y_{t-1}]} = k + \frac{2\sum_{i=1,j=1,i<j}^{k} Cov[Z_i, Z_j]}{Var[Y_t - Y_{t-1}]} \quad , \qquad (2.20)$$

They also have the variance ratio shown in the form of:

$$\frac{E[x_{t+k}-x_t]^2}{E[x_{t+1}-x_t]^2} = k^{\frac{2}{\alpha}} = k^{2H}, \qquad (2.21)$$

Using the equations above they solve for H and obtained:

$$H = \frac{\ln\left(k + \frac{2\sum_{i=1,j=1,i<j}^{k} Cov[Z_i,Z_j]}{\sum_{i=1}^{q} a_i^2\left(E[Z_{i+1}^2]-E[Z_{i+1}]^2\right)+2\sum_{i=1,j=1,i<j}^{q} a_i a_j\left(E[Z_{i+1}Z_{j+1}]-E[Z_{i+1}]E[Z_{j+1}]\right)+2\sigma^2-2Cov[\varepsilon_t,\varepsilon_{t-1}]}\right)}{2\ln(k)} \qquad (2.22)$$

The above included short but core proof of linkages between AR process and fBm show time series that have multiple AR estimates sum to 1 follow a fractional Brownian motion. The stepwise procedure can be done by using a programming tool to automatically find the autoregressive (AR) lag coefficients that best fit the time series in the setting from previous to current data points.

2.5. Empirical Evidence of Cryptos' Fractionality and Reasons

2.5.1. Data Illustration

This analysis uses primary real-time data of cryptocurrency time series from publicly available and subscribed sources[22]. We mainly use daily data on prices, volume, supply, and other categorical influencing factors from 2013[23] or their launching date (mostly started in 2017) to June 2021. There are many countries and exchanges participate in trading cryptocurrencies, the data source of this project is by default the world's average

[22] Primarily https://coinmarketcap.com/, https://www.coingecko.com/en, and https://messari.io/

[23] Even though Bitcoin is launched in 2009, it was very quiet for about 4 years until 2013. There were not many alternative cryptocurrencies strated trading before 2013. Actually, most of our crypto data start from 2017 when the ICO boom and the first real "bull cycle" of the overall crypto began. Bitcoin price rised greatly in 2013 but lasted short period of time, many other coins are not "born" yet at that time, so I tend to see 2017 is the real start. The 4 year "bull cycle" do appear to be on track from the short 12 year history we've seen thus far including the current 2021 "bull".

unless otherwise noted. Among over 10,000 cryptocurrencies that are traded, we select the cryptos based on two criteria: (1) initially released longer than 180 days (started trading before 1/1/2021) to make sure the quality and maturity of the data, and (2) accumulated trading volume over the last year is greater than $500,000 to assure the crypto projects not bankrupted yet nor too thinly traded. Our dataset contains 3,351 cryptos after this filtering process.

2.5.2. Traditional Method Hurst Coefficient Estimations and Distribution

Using the Hurst measure approach stated in section 4.2, we calculate the Hurst coefficient using the scaled variance ratio for 3,351 cryptocurrencies. For the major top cryptos, we find the Hurst for Bitcoin as H=0.571, for Ripple H=0.554, for Ethereum it is 0.576, and for Litecoin H=0.556. Figure 2.5 displays Hurst coefficients for all selected crypto assets in 3 different times that we collected the data, allowing us to capture the time varying properties of cryptos and further confirms the validity of our findings. We find most of the cryptos potentially follow fractional Brownian motion (fBm), with both H > 0.5 (persistent fBm) and H < 0.5 (mean-reverting fBm). Very few have H=0.5 (geometric Brownian motion). Surprisingly, some extreme cases have Hurst coefficients reaching near 0 and some hitting near 0.7, both of which are rarely seen in the markets of securities, stocks, indexes, and commodities. Zero- Hurst coefficients are found for 'stable' coins that play the role of central reserve type currencies (e.g. US dollar) within the crypto world. Stable coins are used as stores of value or units of account to minimize the effects of price volatility. Tether, for example, a stable coin[24] that is forced to conform to a 1:1 ratio with

[24] Since June 2019, Facebook has been developing a low-volatility cryptocurrency, Libra. It is tied to underlying basket of assets, have collaboration with various FinTech leading firms and academia. This design helps the coin values be stable but is fundamentally different from the existing stable coins (Tether, USDC, and Dai) that fix the exchange rates with fiat currencies. This way of tethering Libra coin to underlying assets is in the same logic of tethering to gold, silver, wheat, or corn, which extends to our interest in future research. Our raw observation regards this Libra coin a way of linking crypto currency to real world assets, allowing the arbitrage to always maintain the realization of value fluctuations, which hints toward a mutual fund approach, with faster transaction through digital blockchain platform. We expect to see greater development of stable coins and value in understanding Tokenomics.

the U.S. dollar regardless of any market fluctuation above or below this ratio in the fastest reaction possible. The correction process is an ergodic pure white-noise process with immediate reversibility.

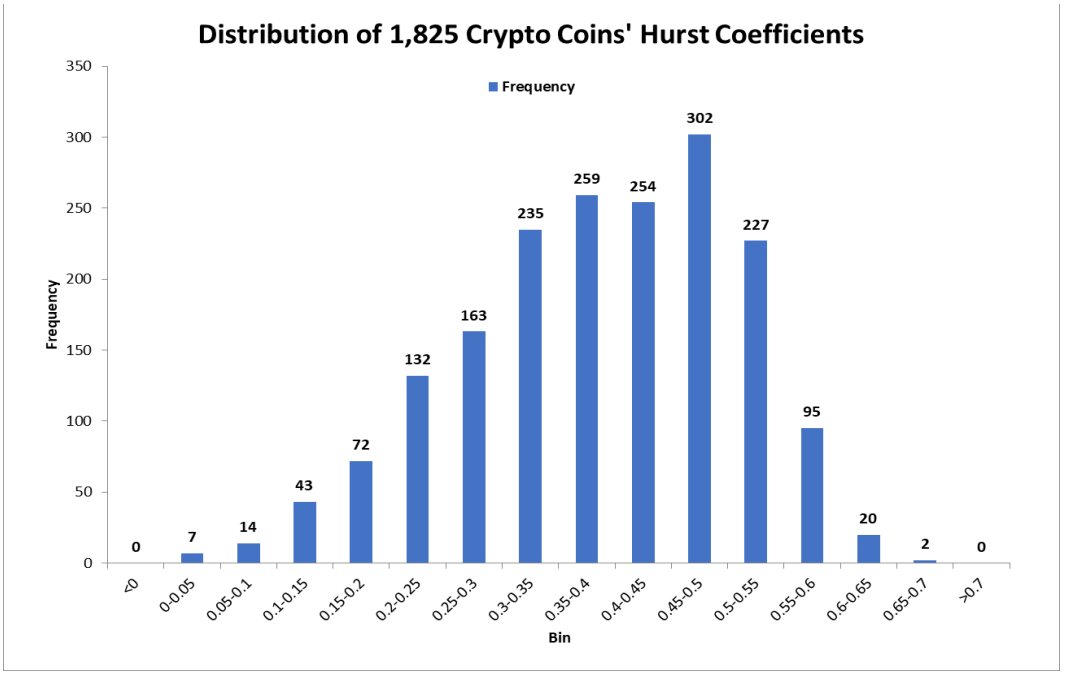


Figure 2.5 (a): Data up to Aug. 2019 for 1,825 Cryptocurrencies Hurst Coefficients Distribution

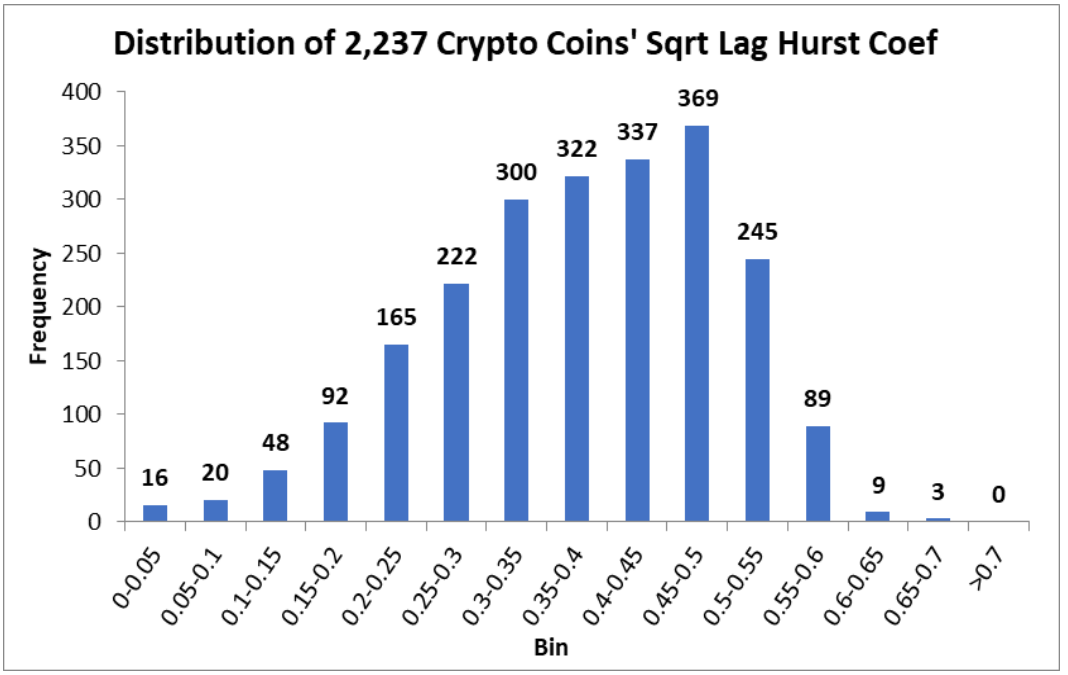


Figure 2.5 (b): Data up to Aug. 2020 for 2,237 Cryptocurrencies Hurst Coefficients Distribution

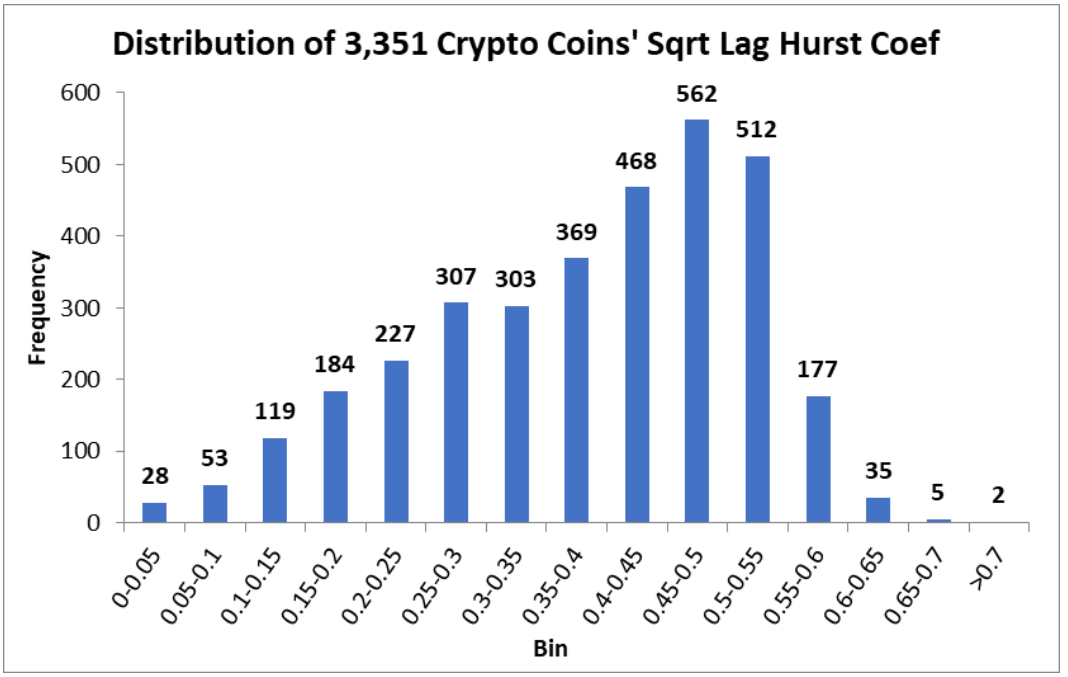


Figure 2.5 (c): Data up to June 2021 for 3,351 Cryptocurrencies Hurst Coefficients Distribution

Figure 2.5: Distribution of Cryptos' Hurst Coefficients Measured in Lag of Observation's Square Root

2.5.3. Hurst Coefficient of Traditional Method vs. De-Mean Enhanced Method

As we mentioned in Section 4.3, the traditional Hurst coefficient estimation stated in Section 4.2 may ignore the influence of structural changes on the mean that are likely to be time-varying, which may lead to a misleading or spurious "long-memory" phenomena. In this section, we use the proposed three-stage methodology to examine whether the long-memory still exists in the de-mean process. Further, various time windows are used to check the robustness of our findings.

Figures 2.6 and 2.7 display the estimated mean process of Bitcoin and Ethereum prices, respectively. First, it is shown that the estimated means change smoothly over time with an overall increasing time trend. This is primarily due to the capped supply vs. increasing demand structure that we detailly illustrated in Section 3.2 and figures 2.2-2.4, also due to the increasing interest and publicity from network effects and media introductions.

Second, there are some slight but obvious turning points and bull / bear cycles: two major humps around 2013 and 2017 in figure 2.6 refer to the miners' reward "halving" from Bitcoin mining, which means mining becomes harder and the supply releasing is less. The dramatical demand increase of Bitcoin continuing all along due to excessive media coverage and hot money entering the crypto market, as well as the deflationary supply structure. The hot money may come from other sectors that are relatively saturated in these years such as High-Tech stocks and real estate sectors. The financial industry has first really sensed the potential development of Bitcoin in 2014, especially with Dell, Microsoft, and PayPal these big firms started accepting Bitcoin, which may have provided some positive credential confidence to the primary financial industry and some more sensitive public. Such momentum in 2014 delivered the first wave and then suffered many hacking events which dampened market confidence that caused a downward trend and thus established the first hump.

2017 is the year that Bitcoin became broadly known by the public, and at the same time regulation was not (and currently is not) yet fully implemented, so all sorts of public and private investments entered which built up the belief even further. Towards the year-end of 2017 when Bitcoin reached an all-time high near $20,000, regulation forces come in such as China and South Korea governments to shut down cryptocurrency exchanges, the same time CBOE Bitcoin Futures contract launched that provides the tool to short Bitcoin, that together significantly changed the dynamic of the institutional structure and turned the market to crash in early 2018. Such a roller-coaster year delivered the second hump to the estimated mean of Bitcoin prices. 2021 is another bull cycle we are currently experiencing, again due to the halving event on the supply side, and also the money flow and demand from quantitative easing occurring around the globe due to Covid. Crypto assets tend to fluctuate in very correlated manner following the leading indicator, Bitcoin. As we can see the Ethereum smoothed mean in figure 7 also displays a similar pattern of the bull in 2017 and 2021. Ethereum was not yet created in 2013, thus no comparison to be drawn for the first "Bull" cycle that appeared for Bitcoin.

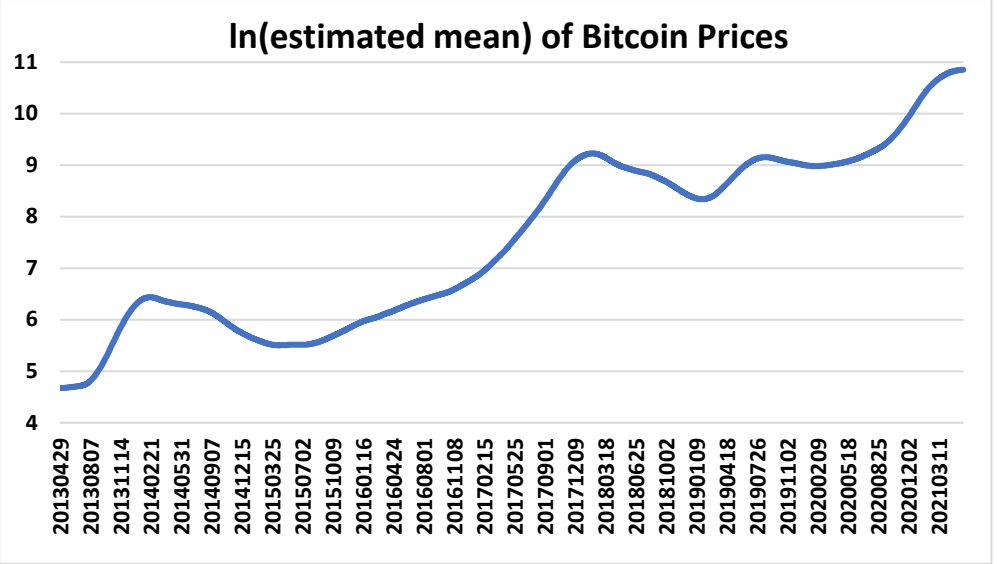


Figure 2.6: Natural Log Nonparametrically Estimated Mean of Bitcoin Prices

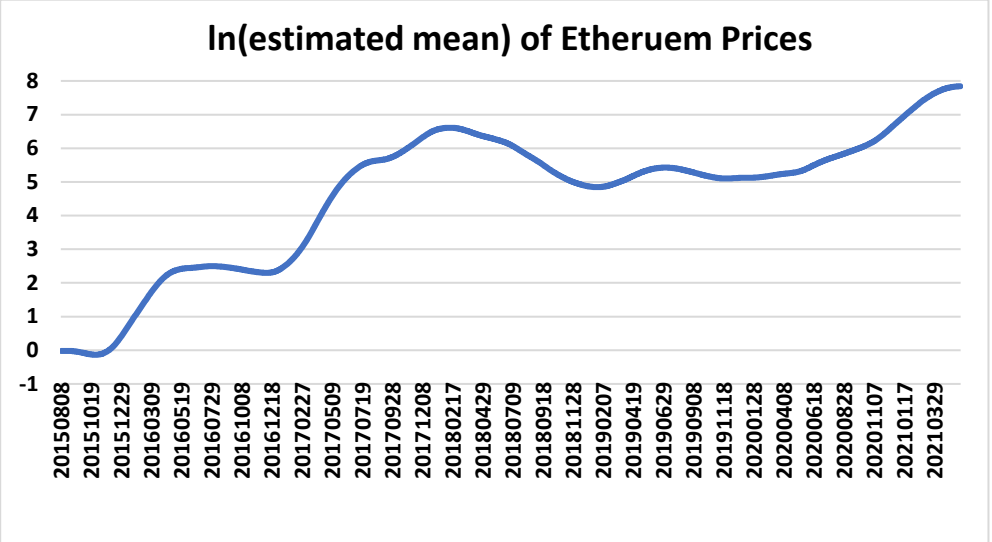


Figure 2.7: Natural Log Nonparametrically Estimated Mean of Ethereum Prices

The crypto structural events, investment environment changes, macro conditions, and regulation evolvement had certainly contributed to the structural changes and cycles of the crypto markets. Hereby we only intuitively relate the background and concurrent events behind the structural changes rather than empirically verify certain event impacts, as it's not the focus of this study and left for future research. Due to the high standard of computational power for this particular method, limited time and space, here we only presented the top two coins' results. The rest cryptos should have similar properties and can be verified in the future.

Table 2.1 and 2.2 report all Hurst coefficient estimation results of Bitcoin and Ethereum with traditional and enhanced method and with side-by-side comparisons. Some interesting findings emerge: first, all the Hurst coefficient estimators of the de-mean process are smaller than 0.5, while those of the original process is larger than 0.5. This implies that existing literature's commonly concluded long-memory phenomenon does not exist in Bitcoin and Ethereum, that are generally conducted by the traditional alike method stated in this paper's section 4.2 and 5.2. Because the traditional method does not remove the influence of structural changes on these currencies that resulted in the increasing or say persistent time-varying mean. It is widely acknowledged that a time series with a smoothly time-varying intercept is likely to behave as a process with a long memory. Indeed, the false "long-memory" observed in the existing literature is caused by structural changes, particularly related to the deflated supply structure.

Table 2.1: Bitcoin Hurst Coefficients in Different Time Windows: Traditional vs. De-Mean Enhanced Method

Bitcoin Hurst Coef. Comparison	**Traditional**	**De-Mean**
Begin_20130428_end_20210530	0.571135	0.239652
Begin_20130428_end_20190729	0.573941	0.284480
Begin_20150101_end_20181231	0.535033	0.286274
Begin_20150601_end_20181231	0.551618	0.289514
Begin_20160101_end_20181231	0.55414	0.311321
Begin_20160601_end_20181231	0.551987	0.307983
Begin_20170101_end_20181231	0.557567	0.325664
Begin_20170601_end_20181231	0.55277	0.329637

Table 2.2: Ethereum Hurst Coefficients in Different Time Windows: Traditional vs. De-Mean Enhanced Method

Ethereum Hurst Coef. Comparison	**Traditional**	**De-Mean**
Begin_20130428_end_20210530	0.576252	0.259957
Begin_20150807_end_20190729	0.582762	0.277532
Begin_20150807_end_20180731	0.571285	0.274978
Begin_20160101_end_20181231	0.605019	0.291626
Begin_20160601_end_20181231	0.593426	0.258042
Begin_20170101_end_20181231	0.610558	0.284591
Begin_20170101_end_2019729	0.608076	0.282562

2.5.4. Summation of Stepwise Autoregression Estimates and Quasi - fBm

Without any prior knowledge or expectation about the lag structure of the fractional process, we use stepwise autoregression (AR) of the first 50 lags to determine the order and sequence of lags that are statistically significant and non-unique. If the coins follow a fractional process, the sum of the lags must sum to 1.0 exactly. Table 2.3 shows the natural summation of the selected lag coefficients for the top 20 coins (a sample). Figure 8 displays scatterplot of stepwise AR estimates summation for all selected crypto assets in 3 different times that we collected the data. We found the majority cryptos' coefficients summation are close to 1 indicating their price series follow fractional Brownian motion, as seen in the Figure 2.8 scatterplots. However, there are still some crypto' AR coefficient summations that are below 0.9 or above 1.1 (95% confidence interval) which indicates processes that are not able to be explained by known random process time series. Those type of processes can possibly go to infinity or converge to 0, some of these may be stable coins that subject to a forced movement towards a fixed range, some may have experienced pump and dump then became inactive. There are some others may pertain to abnormal properties that require further studies.

We also used ANOVA test to empirically compare the initial Stepwise AR model with the restricted least square model which forces the coefficients of the original stepwise lag structure to sum to 1.0. The restricted model expects to be not statistically different from the stepwise model, and from the results we do see 2,779 out of the 3,351 cryptos fit the expectation. Thus, we conclude the model satisfies the unit root condition for a random non-stationary process and that the price movements are consistent with fractional Brownian motion.

Table 2.3: Summation of Stepwise Autoregression Estimates for Top 20 Cryptocurrencies Price History

Crypto Name	Sum Autoregressive Estimates
Bitcoin	0.998117413
Ethereum	1.005429382
Binance Coin	1.014689795
Tether (Stable Coin)	0.999990749

Cardano	1.003671276
Dogecoin	1.107829812
Ripple (XRP)	0.997038177
Polkadot	0.98706603
USD Coin (Stable Coin)	0.999896944
Uniswap	0.998812693
Chainlink	1.007212056
Bitcoin Cash	0.996019635
Litecoin	0.999543216
Polygon	1.140983559
Stellar	0.998222776
Solana	1.045936701
Binance USD	0.999999178
Ethereum Classic	1.000812374
VeChain	1.014362255
Theta Network	0.976255414

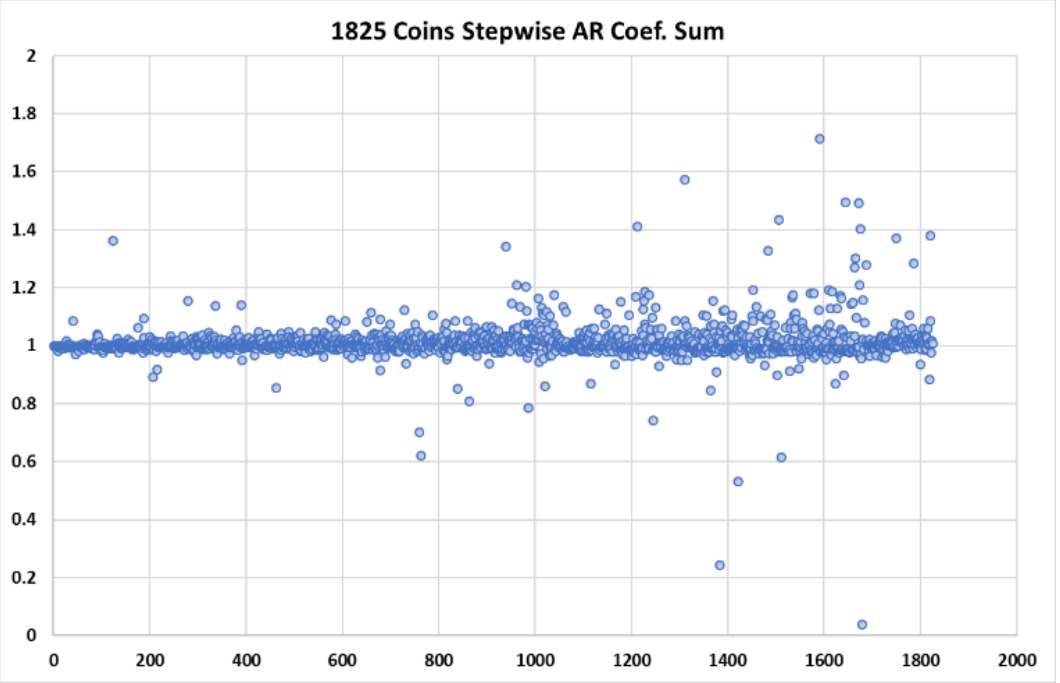


Figure 2.8 (a): Data up to Aug. 2019 for 1,825 Cryptos Stepwise Autoregression Coefficients Summation

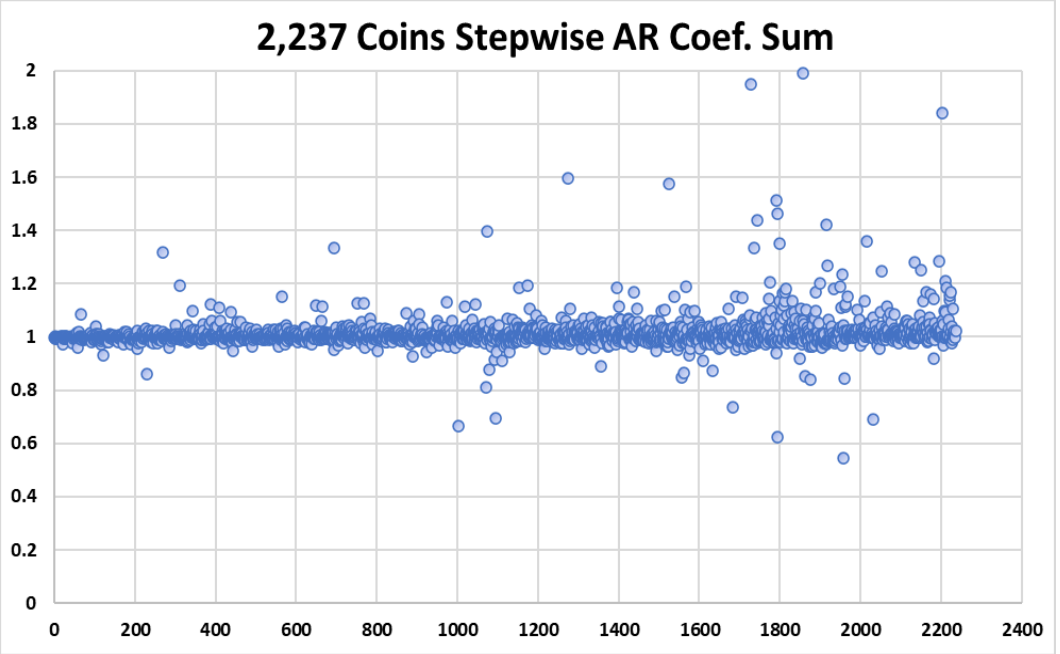


Figure 2.8 (b): Data up to Aug. 2020 for 2,237 Cryptos Stepwise Autoregression Coefficients Summation

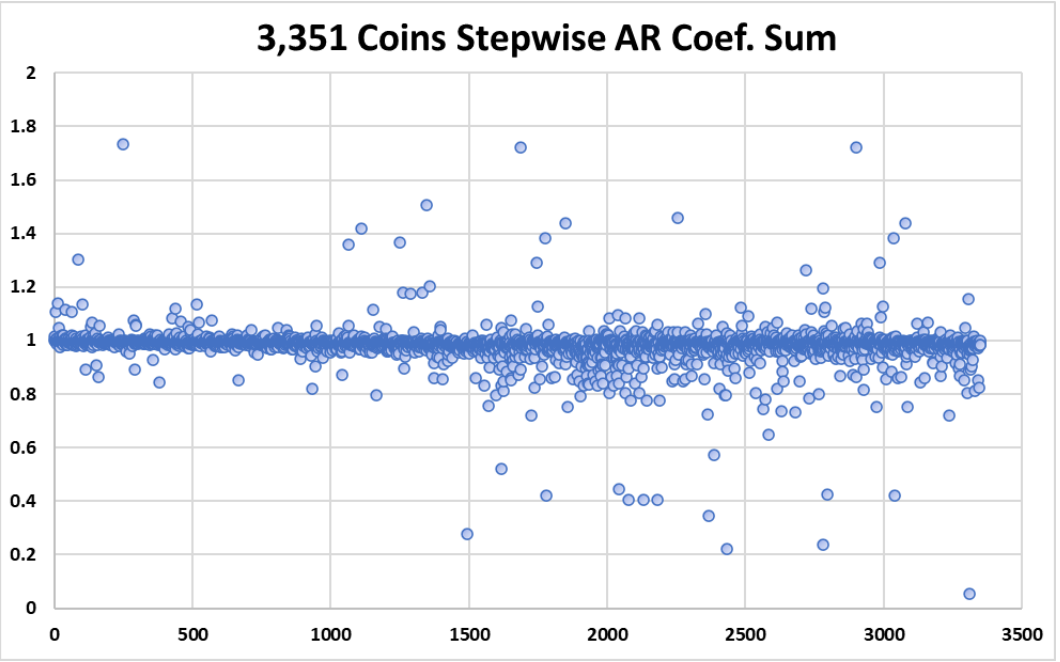


Figure 2.8 (c): Data up to June 2021 for 3,351 Cryptos[25] Stepwise Autoregression Coefficients Summation

Figure 2.8: Three-Time Range Scatterplot Cryptos Stepwise Autoregression Coefficients Summation

[25] There are 12 extreme outliers we identified as their summation above 2 or below 0.3, due to bankruptcy or stablecoin.

2.5.5. The Relationship between the Hurst Measure and Supply Structure: OLS Regression

To understand the factors influencing crypto prices' fractional property, we collected information on the top 105 cryptocurrencies for the price, volume, age (period since its release), estimated the year 2050's supply, total possible supply, current supply issued ratio, type of supply, ongoing emission type. We ran simple OLS regression of these variables against Hurst coefficients estimated earlier with various specifications of these variables, the best fit specification as the following stated. The sole significant factor is the current supply issued ratio, and the various major specifications are as shown below. Control is supply type of currencies, 3 main types: Capped total, Perpetual inflation, and Decaying inflation; where most coins are capped, similar to Bitcoin, and perpetual means adding the same amount of coins each year, whereas decaying means adding fewer coins each year without the total capped amount of coins. We provide Table 2.4, the data summary statistics for influencing factors and control variables that we use below:

Table 2.4: Top 105 Cryptocurrencies influencing factors data summary.

Factors	*Hurst*	*Age (yrs)*	*Supply Issued%*	*Control*	Supply Type
Mean	0.5153	2.0810	0.5878	Capped total	90
Standard Error	0.0063	0.1533	0.0272	Decaying inflation	8
Minimum	0.3110	0.5	0.0407	Perpetual inflation	7
Maximum	0.6829	9.7	1.4444	Total Count	105

The regression results in Table 2.5 show supply issued % decrease the Hurst at a decreasing rate (convex shape), meaning more portions of coins issued, less persistent the price series' memories are. Decaying inflation supply type contributes to more persistency, and Perpetual inflation supply type is opposite. Age can be somewhat statistically significant, though adding age square did not improve the model, indicating age doesn't contribute much to the persistency in the price series compare to supply structure for both the amount issued and the type of issuing.

$$Hurst_i = \beta_{0i} + \beta_1 * Age_i + \beta_2 * SupplyIssued_i + \beta_3 * S_i^2 + \delta * Control + \mathcal{E} \quad (2.23)$$

Intuitively speaking, fewer coins left to be mined indicates a less persistent price trend going to be for this coin, which has been the case for major coins. While for the alternative coins, there are wild relationships in Hurst with supply and other factors, even within these top 105 coins, we suspect an even noisier fitting for the remaining 3,000+ coins which we only have price data but no other factors at this time. The overall R^2 is ranging from 0.0728-0.0888 and the p-values mostly show no statistical significance, such overall poor fitness again shows the lack of price endogeneity in supply and demand, leaving open the question as to how to better model equilibrium. This structure reveals the supply releasing mechanism causing the fractionality in the price series of cryptocurrencies.

Table 2.5: Top 105 Cryptocurrencies Hurst Coefficients OLS Fit with Age, Supply Issued, and Type

Hurst OLS Fit with Market Structure Influencing Factors			
Hurst (Sqrt n lag)	(1)	(2)	(3)
Intercept	0.524851***	0.527834***	0.535094***
	(0.024870)	(0.0287689)	(0.026443)
Age	0.006628 .	0.0044473	0.006644
	(0.003935)	(0.0111458)	(0.004094)
Age^2		0.0003073	
		(0.0014685)	
SupplyIssued%	-0.124696	-0.1267831	-0.157494 .
	(0.079850)	(0.0808486)	(0.084275)
SupplyIssued%^2	0.118355 .	0.1200127 .	0.140730*
	(0.061408)	(0.0622071)	(0.063894)
Supply Type (reference category = Capped total)			
Decaying inflation			0.017784
			(0.023957)
Perpetual inflation			-0.026986
			(0.026529)
Observations	105	105	105
R-squared	0.0728	0.0732	0.0888

. P < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Standard errors in parentheses.

2.6. Conclusion and Discussion

The development of cryptocurrencies has just reached its 10-year milestone; the public, various related industries, and the government have had extensive participation and discussion about it, so does academia. Even though the existing literature and scholars' approaches in studying cryptocurrencies are among diverse fields and angles, they lean heavily toward using traditional economics frameworks to explain the phenomena with cryptocurrencies. However, our empirical findings first show clear evidence that cryptocurrency processes violate the financial economics theory of random walk and efficient market hypothesis, with dominating proportions of cryptos following a fractional Brownian motion process. This finding challenges the use of assumption geometric Brownian motion in conventional economics analysis. Second, the equivalently important argument and empirical evidence of this study is the non-existing price endogeneity in the supply and demand relationship of cryptos. Our results confirm our conjecture that price movement in cryptocurrencies is demand-driven along with a prescribed deterministic supply, e.g., supply in each moment in time is perfectly inelastic. Our study shows the major supply indicators: supply issued % and inflation structures show effects on Hurst estimates (an indicator of the memory in the price series). This is because supply structure with pre-determined and independent of price, along with the emotional hype, implies potential market inefficiency and fractional Brownian motion. Third, we also discovered flaws with using traditional Hurst coefficient estimation to indicate long memory, with using time-varying mean estimation and de-mean procedure to re-conduct Hurst estimation, we find what previously concluded long memory in some literatures are actually short memory, as to that they might have neglected structural changes along with the time series that blurred the true processes we need to look after. Furthermore, we use the stepwise autoregression method to more rigorously verify that most cryptocurrencies have at least one unit root and follow fractional Brownian motion.

With the fractional property and non-endogenous price relationship of cryptocurrencies empirical evidence, we show here, some current studies' assumptions and results may be challenged and therefore the implications can be re-evaluated. This

missed economics relevance in supply-demand structure, financial properties that do not align with classical assumptions, and possibly mistaken conduct of method can completely alter the results and thus implications. Hence we believe our findings provide value in enlightening considerations to include fractionality and structural changes in studying such unique and new asset class.

This study on cryptocurrency prices provides insights for supporting future work in related fields, such as economic conditions, policies, and market shocks' effect on the price dynamics; hedging efficiency among coins and Bitcoin futures; market power of crypto miners and price manipulation; energy cost and environmental impact of mining; and social welfare change due to the existence and emergence of cryptos. A recent study was done by Saleh (2018) which was presented in the AEA ASSA 2019 annual meeting shows that cryptos cause welfare loss. This study along with other recent studies show energy and human capital cost associated with cryptos, together with the question on the true value and benefit of cryptocurrencies. Nonetheless, we believe there are values and great potential in the closely tied Blockchain technology, which we hold interest in studying more. However, clearly evaluating the relationship between cryptocurrencies and blockchain especially their underlying economic linkages requires additional analysis. All these issues briefly or intuitively discussed are left for further investigation in future work.

2.7. References:

Al-Yahyaee, K. H., Mensi, W., & Yoon, S. M. (2018). “Efficiency, multifractality, and the long-memory property of the Bitcoin market: A comparative analysis with stock, currency, and gold markets.” *Finance Research Letters.* *https://doi.org/10.1016/j.frl.2018.03.017*

Bariviera, Aurelio F. & Basgall, María José & Hasperué, Waldo & Naiouf, Marcelo (2017). Some stylized facts of the Bitcoin market, *Physica A: Statistical Mechanics and its Applications,* Volume 484, 2017, Pages 82-90, *https://doi.org/10.1016/j.physa.2017.04.159*

Basu, Soumya and Easley, David and O’Hara, Maureen and Sirer, Emin, Towards a Functional Fee Market for Cryptocurrencies (January 18, 2019). Available at SSRN: *https://ssrn.com/abstract=3318327*

Baur, D. G. & Dimpfl, T. & Kuck, K. (2018). Bitcoin, gold and the US dollar – A replication and extension, *Finance Research Letters*, Volume 25, 2018, Pages 103-110, ISSN 1544-6123, *https://doi.org/10.1016/j.frl.2017.10.012*.

Baur, Dirk G. & Hong, KiHoon & Lee, Adrian D., (2018). Bitcoin: Medium of exchange or speculative assets?, *Journal of International Financial Markets, Institutions and Money,* Volume 54, 2018, Pages 177-189, ISSN 1042-4431, *https://doi.org/10.1016/j.intfin.2017.12.004*

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, and Catherine Casamatta, 2017, The blockchain fold theorem, Preliminary Work in Progress.

Blau, Benjamin M., 2017. "Price dynamics and speculative trading in bitcoin," *Research in International Business and Finance*, Elsevier, vol. 41(C), pages 493-499.

Borri, Nicola, Conditional Tail-Risk in Cryptocurrency Markets. *Journal of Empirical Finance*, Volume 50, 2019, Pages 1-19, ISSN 0927-5398,

Bouri, Elie & Molnár, Peter & Azzi, Georges & Roubaud, David & Hagfors, Lars Ivar (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, Volume 20, 2017, Pages 192-198, ISSN 1544-6123, *https://doi.org/10.1016/j.frl.2016.09.025*

Buchholz, M., J. Delaney, J. Warren, and J. Parker. 2012. "Bits and Bets, Information, Price Volatility, and Demand for BitCoin, Economics 312." *www.bitcointrading.com/pdf/bitsandbets.pdf*

Burniske, Chris & White, Adam (2017). Bitcoin: Ringing The Bell For A New Asset Class, ARK Invest. 3/26/2019 Web. *https://research.ark-invest.com/hubfs/1_Download_ Files_ARK-Invest/White_Papers/Bitcoin-Ringing-The-Bell-For-A-New-Asset-Class.pdf*

Calum G. Turvey, Paitoon Wongsasutthikul, (2016) "An autoregressive approach to modeling commodity prices as a quasi-fractional Brownian motion", *Agricultural Finance Review*, Vol. 76 Issue: 1, pp.54-75, *https://doi.org/10.1108/AFR-01-2016-0004*

Caporale, Guglielmo Maria & Gil-Alana, Luis & Plastun, Alex (2018). Persistence in the cryptocurrency market, *Research in International Business and Finance*, Volume 46, 2018, Pages 141-148, ISSN 0275-5319, *https://doi.org/10.1016/j.ribaf.2018.01.002*

Cheah, E.T., Fry, J., (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, Volume 130, 2015, Pages 32–36. ISSN 0165-1765, *https://doi.org/10.1016/j.econlet.2015.02.029*

Chen, B. & Hong, Y. (2012). Testing for smooth structural changes in time series models via nonparametric regression. Econometrica 80, 1157–1183.

Cheung, A., Roca, E., Su, J.-J., (2015). Crypto-currency bubbles: An application of the Phillips-Shi-Yu (2013) methodology on Mt.Gox bitcoin prices. *Applied Economics, 47(23), 2348–2358. doi:10.1080/00036846.2015.1005827*

Ciaian, Pavel & Rajcaniova, Miroslava & d'Artis Kancs (2016). The economics of BitCoin price formation, *Applied Economics*, 48:19, 1799-1815, *DOI: 10.1080/00036846.2015.1109038*

Coinmarketcap (2021a). Active Trading Cryptocurrencies, Web, 06/27/2021. *https://coinmarketcap.com/*

Coinmarketcap (2021b). Cryptocurrencies Market Cap, Web, 06/27/2021. *https://coinmarketcap.com/*

Cong, Lin and He, Zhiguo and Li, Jiasun, Decentralized Mining in Centralized Pools (February 1, 2018). George Mason University School of Business Research Paper No. 18-9. Available at SSRN: *https://ssrn.com/abstract=3143724*

Digital signature (2018). https://en.wikipedia.org/wiki/Digital_signature (Accessed 1/6/2018)

Dyhrberg, A. H., (2016a). Bitcoin, gold and the dollar – A GARCH volatility analysis, *Finance Research Letters*, Volume 16, 2016, Pages 85-92, ISSN 1544-6123, *https://doi.org/ 10.1016/j.frl.2015.10.008*

Dyhrberg, A. H., (2016b). Hedging capabilities of bitcoin. Is it the virtual gold?, *Finance Research Letters*, Volume 16, 2016, Pages 139-144, ISSN 1544-6123, *https://doi.org/10.1016/j.frl.2015.10.025*

Easley, David and O'Hara, Maureen and Basu, Soumya, From Mining to Markets: The Evolution of Bitcoin Transaction Fees (May 1, 2018). *Journal of Financial Economics (JFE)*, Forthcoming. Available at SSRN: *https://ssrn.com/abstract=3055380*

Efficient-Market Hypothesis (2019). Wikipedia, Wikimedia Foundation, 24 Mar. 2019, *en.wikipedia.org/wiki/Efficient-market_hypothesis#cite_note-12*

Enders, W., (2010). Applied Econometric Time Series. Wiley.

Gkillas, Konstantinos and Katsiampa, Paraskevi (2018). An application of extreme value theory to cryptocurrencies, *Economics Letters*, Volume 164, 2018, Pages 109-111, ISSN 0165-1765, *https://doi.org/10.1016/j.econlet.2018.01.020*

Grinberg, Reuben (2011). Bitcoin: An Innovative Alternative Cryptocurrency (December 9, 2011). *Hastings Science & Technology Law Journal*, Vol. 4, p.160. Available at SSRN: *https://ssrn.com/abstract=1817857*

Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in us labor productivity. Journal of Economic Perspectives 15, 117–128.

Hayes, Adam S. (2019) Bitcoin price and its marginal cost of production: support for a fundamental value, *Applied Economics Letters*, 26:7, 554-560, *DOI: 10.1080/13504851.2018.1488040*

Hong, Y., Wang, X. & Wang, S. (2017). Testing strict stationarity with applications to macroeconomic time series. International Economic Review 58, 1227–1277.

Huberman, Gur and Leshno, Jacob and Moallemi, Ciamac C., An Economic Analysis of the Bitcoin Payment System (March 13, 2019). Columbia Business School Research Paper No. 17-92. Available at SSRN: *https://ssrn.com/abstract=3025604*

Jiang, Yonghong & Nie, He & Ruan, Weihua (2018). Time-varying long-term memory in Bitcoin market, *Finance Research Letters,* Volume 25, 2018, Pages 280-284, ISSN 1544-6123, *https://doi.org/10.1016/j.frl.2017.12.009*

Kaplanov, Nikolei, Nerdy Money: Bitcoin, the Private Digital Currency, and the Case Against Its Regulation (March 31, 2012). Temple University Legal Studies Research Paper. Available at SSRN: *https://ssrn.com/abstract=2115203*

Katsiampa, P., (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, Volume 158, 2017, Pages 3-6, ISSN 0165-1765, *https://doi.org/10.1016/j.econlet.2017.06.023*

Kirman, Alan (2009). "Economic theory and the crisis." Voxeu. 14 November 2009.

Klein, T. & Thu, H. P. & Walther, T., (2018) Bitcoin is not the New Gold – A comparison of volatility, correlation, and portfolio performance, *International Review of Financial Analysis*, Volume 59, 2018, Pages 105-116, ISSN 1057-5219, *https://doi.org/10.1016/j.irfa*. 2018.07.010.

Koutmos, Dimitrios (2018). Bitcoin returns and transaction activity. *Economics Letters*. Volume 167, June 2018, Pages 81-85

Kristoufek, Ladislav (2015). What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis. *PLoS ONE* 10(4): e0123923

Krugman, Paul (2018). "So Bitcoin Just Lost Half Its Value. Where Does It Now Stand Relative to Fundamentals? Hard to Say, Because There Aren't Any Fundamentals. More than Ever, This Looks like a Pure Bubble Https://T.co/L9LbDyuZKD." Twitter, 17 Jan. 2018, Web. *https://twitter.com/paulkrugman/status/953634522935037958?lang=en*

Lahiff, Keris. (2018) "Nobel-Winning Economist Rails against Bitcoin, Says It's a Perfect Example of 'Faddish Human Behavior'." CNBC, 13 Apr. 2018, Web. *www.cnbc.com/2018/04/13/the-bitcoin-bubble-is-an-example-of-faddish-human-behavior-shiller.html*

Lavi, R., O. Sattath and A. Zohar, 2017, Redesigning Bitcoin's Fee Market, arXiv:1709.08881.

Libra White Paper: Blockchain, Association, Reserve. (n.d.). Retrieved from *https://libra.org/en-US/white-paper/* July 2019, Web.

Lin, C.-F. J. & Terasvirta, T.¨ (1994). Testing the constancy of regression parameters against continuous structural change. Journal of Econometrics 62, 211–228.

Eyal, Ittay and Sirer, Emin Gün (2014). Majority is not Enough: Bitcoin Mining is Vulnerable. To appear in Proceedings of *Financial Cryptography*, Barbados, March 2014.

Mensi, W., Al Yahyaee, K., & Kang, S. H. (Accepted/In press). Structural breaks and double long memory of cryptocurrency prices: A comparative analysis from Bitcoin and Ethereum. *Finance Research Letters*. *https://doi.org/10.1016/j.frl.2018.07.011*

Nakamoto, Satoshi, 2008, Bitcoin: A peer-to-peer electronic cash system, Online Publication.

Paralysis Proofs: Secure Dynamic Access Structures for Cryptocurrencies and More. F Zhang, P Daian, G Kaptchuk, I Bentov, I Miers, A Juels (2018).

Robinson, P. M. (1989). Nonparametric estimation of time-varying parameters. In Statistical Analysis and Forecasting of Economic Structural Change. Springer, pp. 253– 264.

Saleh, Fahad (2018). Volatility and Welfare in a Crypto Economy (December 22, 2018). Available at SSRN: *https://ssrn.com/abstract=3235467*

Tarnopolski, Mariusz (2017). "Modeling the price of Bitcoin with geometric fractional Brownian motion: A Monte Carlo approach," Papers 1707.03746, arXiv.org, revised Aug 2017.

Thiesa, Sven and Molnár, Peter (2018). Bayesian change point analysis of Bitcoin returns. *Finance Research Letters.* Volume 27, pp. 223-227.

Tradingview (2021). Market Cap Btc Dominance. Web. June 27, 2021. *https://www.tradingview.com/symbols/CRYPTOCAP-BTC.D/*

Trimborn, Simon & Härdle, Wolfgang Karl, 2018. "CRIX an Index for cryptocurrencies," *Journal of Empirical Finance,* Elsevier, vol. 49(C), pages 107-122.

Wheatley, Spencer and Sornette, Didier and Huber, Tobias and Reppen, Max and Gantner, Robert N., (2018). Are Bitcoin Bubbles Predictable? Combining a Generalized Metcalfe's Law and the LPPLS Model (March 15, 2018). *Swiss Finance Institute Research Paper* No. 18-22. Available at SSRN: *http://dx.doi.org/10.2139/ssrn.3141050*

3. CHAPTER 3

Horses vs. Tractors? Old Order Amish Population Growth and New York Farmland Markets

Old Order Amish and Mennonite (Plain) communities are a major part of farming systems in several states and continue to grow despite limited adoption of modern technology. We study the relationship of farmland markets and Amish settlement growth, using a unique, spatially explicit dataset on growth of New York Amish communities and all arms-length New York farmland transaction records from 2007 to 2015. Amish influence is measured by the number of Amish districts (church groups) within a 10mile radius of each parcel transacted. We first develop a conceptual framework for how Amish farmers who are not operating on the production technology frontier could be profitable and compete with farmers that do. Then we estimate a standard hedonic model for farmland prices, with Amish influence as the key independent variable. To address the potential endogeneity between Amish population growth and farmland markets, we use an enclaving or shift-share like IV, based on an established approach in the labor economics literature. Our results suggest that there is not a robust, statistically significant relationship between Amish population growth and farmland price. This implies that Amish farms are relatively profitable and bid for farmland at a level similar to conventional farms. Our study contributes to the literature on (1) the inverse farm productivity–size relationship and (2) competitiveness of small farms.

3.1. Introduction

Economic development in agriculture worldwide is typically characterized by adoption of advanced production technologies and increasing farm size, as a part of the classical "structural transformation" (Johnston and Mellor, 1961). At the same time, there is a vast literature and ongoing debate on farm size and productivity in developing countries. Many studies support the inverse farm size-productivity relationship at least in some situations and typically measure productivity using gross output such as yield (Rada et al., 2019; Gautam and

Ahmed, 2018; Barrett et al., 2010), which indicates that small farms are more productive and potentially competitive on a per-acre basis, while it may vary by crops (Hoken and Qun, 2019). Recent research suggests a focus on profits and total factor productivity instead of yield may explain productivity more realistically (Michler et al., 2019; Muyanga and Jayne, 2019). Michler et al. (2019) suggests that farmers' choice of adopting certain technology may not solely based on its effect of increase in yield, but more due to reduction in costs or increase in profits. In the U.S. setting, what is seen as an inevitable structural transformation to larger, technologically advanced farms (Peter Timmer, 1988) is belied by the growth of Amish and other religious communities that limit adoption of modern agricultural technology (tractors, electronics, and cars) and continue operating their family farms in small-scale (Cross, 2014). In this study, we use the relationship between farmland prices and the population growth of small-scale, less productive but staying competitive Amish farmers, to explore the persistent coexistence of different farming systems that include ranging farm sizes and different farm size-productivity relationships.

In this study, we provide a conceptual model for the coexistence of Amish and conventional farming; and empirically test whether the primary farm asset values reflect growth of Amish populations. Farmland market data provide an opportunity to better understand the competition for resources with the growth of religious communities that limit technology adoption, especially in the absence of other data sources. We first develop a conceptual framework that formalizes how Amish farmers who are not operating on the production technology frontier could compete with conventional farmers, under the assumption that they have less efficient production technology but lower labor costs. We then hypothesize that any differences in profitability of farming systems would be reflected in farmland prices and analyze whether farmland prices are influenced by Amish population growth. We use a unique, spatially explicit dataset on the growth of Old Order Amish communities and all arms-length New York farmland transaction records from 1999 to 2015. Estimating a standard hedonic model, we find that Amish community density does not have a statistically significant relationship with farmland prices, even after implementing a shift-share instrumental variable (IV) based on historic Amish settlement patterns. Our results suggest

that Amish farmers compete for farmland in a manner comparable with conventional farmers, reflecting a similar level of profitability, despite the scale or technology adoption. Effectively, religious traditions that keep production costs low for these communities may help them survive alongside otherwise more efficient farmers that face higher labor (or lifestyle) costs.

The farm size-productivity relationship has been intensively debated, most of the studies support the inverse relationship because they find the small farms usually spend more intensive human labor efforts in addition to the standard machinery operations that are also applied in the larger farms (Sen, 1962; Collier, 1983; Carter, 1984; Rios and Shively, 2005). However, some studies have found the opposite result and claimed that the inverse relationship is spurious and not widely applicable, potentially due to market failure, measurement error, or omitted variable bias (Feder, 1985; Benjamin, 1995; Bhalla and Roy, 1988). A U-shaped relationship between farm size and productivity in some developing countries is also found possible, where very small farms have inverse, small farms have flat, and then medium to large farms have positive relationships with farm productivity (Muyanga and Jayne, 2019; Sheng et al., 2019). Yet, the majority of the literature supports an inverse relationship, while Amish farms have a direct farm size-productivity relationship. Because they are relatively small scale (Zook, 1994), and also typically less productive (defined by yield per acre or cow) due to lower technology usage. Thus the growth of these farming communities does not easily fit within classical economic paradigms. The missed productivity or competitiveness due to lower technology use could potentially be compensated by lower family living expenses, high savings rates, and community risk-sharing and collaboration (Delpierre et al., 2019; Reid, 2015). Consistent with this argument, a small sample of 27 Plain\footnote{Plain people are Christian groups characterized by separation from the world and by simple living, including plain dressing in modest clothing. Plain people refer to Old Order Amish, New Order Amish, Old Order Mennonites, Hutterites, and many other communities with slightly varying religious traditions.} dairy farms collected by Farm Credit East shows that lower labor, machinery and overall costs lead to comparable profits and return on assets as conventional farms. Efficient small dairy farm can be cost competitive with larger farms, while the less efficient small farms cannot (Tauer, 2001). We believe that many

Amish farmers are on the efficient side, as religiosity indicates greater commitments to farm practices (Decker et al., 2014).

Very little economics research considers Amish or other Plain26 community farming. Most academic research about Amish communities is descriptive and in the following areas: (1) rural sociology, occupation, and education; and (2) religious rules and lifestyle. Kreps et al. (1994) and Hill (2006) document the farming-dominated occupational structure and movement to nonfarm occupations of some communities. Amish communities have developed successful enterprises such as woodworking, tourism, construction, and small shops (Kreps et al., 1997; Hovinen, 1995). The Amish's strong religious networks may have supported the growth of successful farm and non-farm businesses (Jeong, 2013). The Amish educational system emphasizes practical skills, with formal education ending after 8th grade, while ongoing education relies on an apprentice system for practical vocational skills (Hynes and Edgington, 2018). Amish restricts secular education as they think it can boost people's wage labor productivity and thus more likely to leave the community that may harm the community welfare with fewer participation (Wang, 2020).

Old Order Amish and Mennonite (Plain) communities are a major and often-growing part of farming systems in several states that also have a large presence of conventional or commercial agriculture, including Pennsylvania, Wisconsin, Ohio, Indiana, and New York (Elizabethtown College, 2017). Farming is an integral part of their religious convictions, which emphasize humility, simplicity, altruism, family ties, and community collaboration (Kraybill et al., 2013; Choy, 2020). The Amish have become influential in New York agriculture and other concentrated local markets (Reid, 2015; Moss and Schmitz, 2003). There are some commonalities to their concentration: Amish tend to settle in rural locations with nearby small commercial centers and sparse population, that are conducive to small-scale farming (Anderson and Kenda, 2015), e.g. moderately rolling hills and lower farmland prices (Ifft and Gao, 2019). Amish prefer to settle near established Amish communities (Anderson and Kenda,

[26] Plain people are Christian groups characterized by separation from the world and by simple living, including plain dressing in modest clothing. Plain people refer to Old Order Amish, New Order Amish, Old Order Mennonites, Hutterites, and other communities with slightly varying religious traditions.

2015), the religious concentration is also correlated with indicators of investment into farmland such as tillage (Decker et al., 2014). This settlement pattern motivates our construction of an enclaving or shift-share instrumental variable, similar to the immigration and labor literature, e.g. Bartel (1989), Lewis and Peri (2015), and Jaeger et al. (2018).

Our study makes several contributions. We contribute to the still-inconclusive literature on the inverse farm size-productivity relationship and competitiveness of small farms. Our novel application of a shift-share/enclaving IV allows for more accurate measurement of the impacts of a growing farming community, and could potentially be applied to other analyses of farmland markets. Further, this is also the first study to quantify the relationship between Amish community size and farm real estate prices, which make up over 80 % of U.S. farm sector assets. This study was originally motivated by anecdotes that "Amish drive up land prices"; our empirical finding suggests that Amish are actually competing in a manner similar to their non-Amish neighbors.

The rest of the paper is structured in the following way. Section 2 details the conceptual framework to demonstrate the equilibrium land price under the coexistence of conventional and Amish farmers. In section 3, we introduce data source. Then followed by section 4 to discuss empirical model and identification strategy. Section 5 presents empirical results and discusses whether Amish population growth impacts farmland prices in New York state. In section 6, we conclude the paper, and discuss implications of our work for farmland economics research, production economics, and plain communities.

3.2. Conceptual Model

The problem of two quite different farming production styles competing with the same land market to achieve the equilibrium is of interest to us. We refer to the canonical Ricardian approach (Ricardo, 1817) for land rent and then model the production isoquants difference between intensive and extensive farming style. Ricardian rent theory's key assumptions of limited land supply and varying land quality result in heterogeneous land productivity and farmland price. Modern economists, Malthus explains the concept of residual surplus, and von

Thünen introduces the theory of rent differentials resulted by distance from a central market. These form the basis of our modern understanding of land rents and land values (Barkley, 1986). Our conceptual framework is built based on the capitalization model, that is widely applied to formalize core factors of land values in classical theoretical basis (Nickerson and Zhang, 2014). Capitalization models works well for long run though not short run, thus time series models are further developed to address the concern (Melichar, 1979; Featherstone and Baker, 1987). In addition, hedonic model is developed to evaluate the influencing factors of farmland prices, utilizing the increasing access to spatially explicit and cross-sectional data that reflects spatial and temporal differences (Irwin et al., 2010). More importantly, hedonic model recognizes that farmland returns or its valuation may also come from "nonfarm" sources, such as close proximity to urban areas (Capozza and Helsley, 1989), land characteristic and soil quality (Miranowski and Hammes, 1984; Ervin and Mill, 1985), land use restrictions (Chicoine, 1981), and nearby amenities (Johnston and Duke, 2009). The quantifiable data on the Amish religious and farming habits characteristics is rare, so we consider this group represents their characteristics as a proxy, or a bundle of characteristics (lower farming technology use, larger family and lower labor costs, high savings or less expenditure lifestyle, and community risk sharing) that are internalized in their communities. We empirically test the role of Amish population growth in land value using a hedonic approach capturing wide range of agricultural and non-agricultural characteristics, in addition to the capitalization model in below that provides a general framework of farmland revenue potential.

The primary goal of this paper's conceptual model is to help understand the coexistence of two very different farming systems. It is motivated by Farm Credit East dairy farm survey data from 305 conventional farms and 27 Plain (Old Order Amish and Mennonite) farms. Although these data are not representative, they support our underlying intuition that Plain farmers have lower input costs and their family labor is implicitly "undervalued" relative to the conventional producers, due to their religious and cultural traditions (Reid, 2015). Plain farms on average produce less milk per cow (17% lower), but other factors support their profitability. Lower input costs (including 30% lower labor cost per unit milk produced) help

them achieve a similar return on assets, 1.4% vs. 1.5% for plain vs. conventional farms, respectively, and higher net farm income per cow ($346 vs. $83). Plain communities may also benefit from higher savings due to lifestyle (which is related to labor costs) and community risk-sharing (Delpierre et al., 2019; Choy, 2020; Reid, 2015), although we focus on the labor costs in our model.

The conceptual model assumes two groups of producers who may produce under completely different combinations of inputs or various techniques, while still achieving the same level of output or profit level. We refer to the foundation of Feller (1972), where he explains the linkages and potential concerns between theory and empirical analysis, in evaluating the effect of changing factor prices on relative factor shares considering a technological change or technique choices. He argued that the use of isoquant is sometimes mixing up the installed techniques versus the alternative methods of production with different input combinations. He points out that the conventional production isoquant represents only a given underlying technique while there are other techniques by choice or by constraint. Feller (1972) develops the concept of isoquants envelope that contains input combinations to produce a certain output for all existing techniques with each has a different isoquant. All of these techniques represent the society's technology, with some of the producers' technique is more advanced than others. In the case of our model, we plot two isoquants with Amish farmers more to the right of the society's envelope as to describe they choose to use the less mechanical technique; while the conventional farmers use more advanced technique and their isoquant is more to the left, shown in the left panel of figure 3.1.

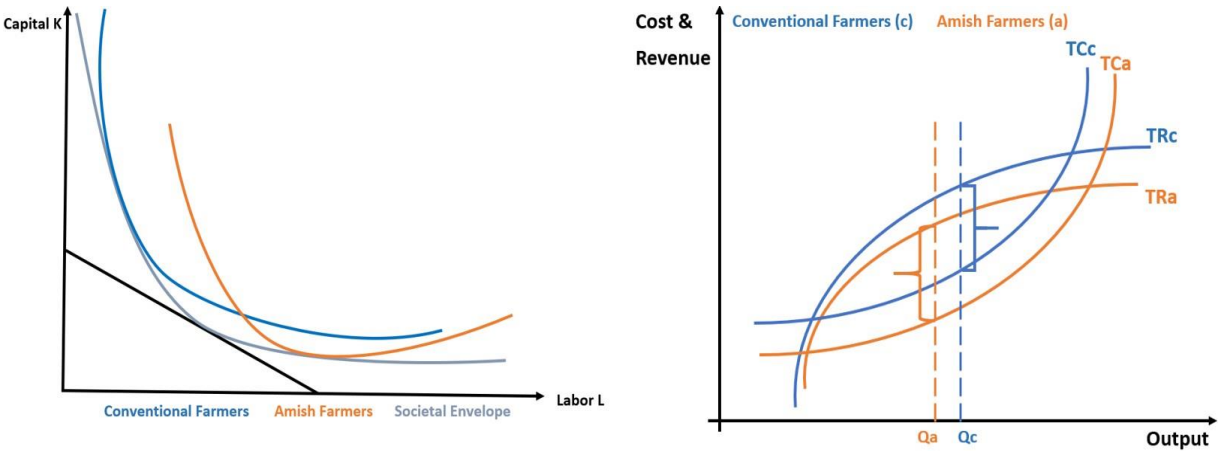


Figure 3.1: Amish Production and Equilibrium Conceptual

We model production for two types of farmers: (1) capital intensive, representing conventional farmers; and (2) labor intensive, referring to Amish farmers. In other words, type 2 farmers use more labor in production, to compensate for the limited use of capital, in order to reach the same level of production or profit that enables them to stay competitive in acquiring major inputs such as farmland. E.g. Amish farmers use little to no machinery, electricity or other capitals; but spend much more labor per unit of land, as they generally have larger family and thus more family members to participate in production (Choy, 2020). We model these two types of farmers' capital or labor insensitivity in general production framework considering input trade-offs. Then we present the production, profit maximization, market clearing conditions, equilibrium solving for this model.

3.2.1. Theoretical Background: Ricardian Theory of Rent

The land value, rent and potential farming profit are in the context of the Ricardian theory with the core decision making of marginal substitution, where the highest quality land is the first occupied, then the next coming farmer would need to choose to pay the rent which is the difference of productivity and income, or to occupy a less productive piece of land, so on and so forth until there's no more land available or any sort of boundary condition is reached, such as a country or a state. In this study, we measure New York State farmland market, thus we would consider the boundary is reached and farmers compete under limited land supplies rather than move to find and occupy new available free land. We assume there are relatively much more available labor than land, people will compete to their own subsistence level income or standard of living, so the wage can go even below the theoretical marginal level, for instance, Amish may pay themselves lower wages in order to still maintain their competitiveness in the market, as they may consume less for computers, TVs, alcohol, electricity, or any modern leisure-related goods (Kraybill, 2021). The consequence is that this group of people may drive down the wage of the entire labor market, and in addition to influence profit and produced goods allocation.

In the context of Ricardian rent theory (Ricardo, 1817), rent is produced goods' value after paying labor. In our context here, we would like to analogously model the rent as the profit, where more fertile lands yield higher profit obtained to the farmers, which allow them

to be always ahead of the competition as they have the wealth accumulation to bid out other competitors. However, the existence and expansion of the Amish population may change this market dynamic, where less fertile lands Amish owners may give themselves lower wages or save the profit more, which they can bid higher to compete with the conventional farmers who used to be owning more fertile lands.

3.2.2. Capitalization Model and Production Profit Maximization

Based on the basic capitalization model discussed by Nickerson and Zhang (2014), farmland enters as a fixed cost of production that generates a stream of economic returns, discounted to the present value. The returns are defined as the return above all variable factors of production and with consideration of land quality as well as technique use. The model is farming annual return based, with return accumulation in years of production the land may have to denote the farmland price, or think of as the farmland price equals to the potential future returns from farming on the land. Under Ricardian theory framework with expected farming returns to evaluate farmland rent and values, level of land's productivity closely link into its farmland value. Reaching comparable level of productivity or profit is the key to farmers ability bidding at same level of farmland price. The land price equilibrium reflects farmers' recognition of potential amounts of goods to be produced with their own production methods to be profitable.

Formally, the model is written as

$$P_t\ (m,p,w) = \int_{t=0}^{\infty} R_{it}(m,p,w) * e^{-rt} dt. \qquad (3.1)$$

where P_t is the farmland price of period t; R_{it} is annual net farming returns that is closely related to profits for farm i in time t, with factors of farm characteristics m, such as soil quality and level of machinery usage, output price p (produced crop, i.e. corn, soybeans, wheat and etc.) and input prices w_n that will be specified later in equation (4). And r is the discount rate or risk-free rate. We then model farm conduct as a farm household profit-maximizing problem (instead of a production problem):

$$\pi\ (\boldsymbol{p}, \boldsymbol{w}, \boldsymbol{l}) = max_x\{\boldsymbol{p} * f(\boldsymbol{x}, \boldsymbol{m}) - w_k * \ x_k\}. \quad (3.2)$$

where *f*(*x,m*) is the production function (Cobb Douglas form) we model in the following:

$$Q = f(\boldsymbol{x}, \boldsymbol{m}) = A(\boldsymbol{m}) * \boldsymbol{B}_j \prod_{n=1}^{N} \boldsymbol{x}_n{}^{\lambda_n}. \qquad (3.3)$$

Here *j* = 1,2 indicates 1 as type 1 farmer who is operating through the system primarily hired labor (relates to the prior discussed conventional farmers) and 2 as type 2 farmer who is mainly under the system of solely family labor (reflects emerging settlement of Amish farmers). *A*(*m*) is an efficiency parameter considering land characteristics which is a total factor productivity (TFP) like parameter; *B* is a behavioral indicator that illustrates the intensiveness of farming operational nature (we simplify to have only two stages: intensive and extensive, instead of continuously infinite levels). To keep the focus, the inputs we model here are farmland *K* and labor L_j that are the non-negative quantities utilized, where *j* = 1,2; and the corresponding input prices P_t (endogenous) and g_j wage. For the labor and wage, we have two types: hired labor (*j* = 1) and family labor (*j* = 2), assume hired labor is pricier and more skilled than family labor.

Convert the equations above, we can indicate the farmland price equilibrium as

$$\pi\ (p, w, m)\ =\ max_x\{p\ *\ f(x, m)\ -\ P_t\ *\ K\ -\ g_j\ *\ L_j\}. \quad (3.4)$$

Primary assumptions about output and input mobility:

i. Labor *L*: hired labors L_1 and family labors L_2 are not freely mobile, as under this paper's setting that conventional and Amish farmers are distinct communities, the Amish have strong religious beliefs and community gathering nature who don't usually go out of their communities to seek employment, nor they often hire outside employees unless the skills are not met within the community such as driving trucks (Kraybill et al., 2011).

ii. The other input (farmland) *K* and output (crops and animal products) are assumed to be traded across all communities without religious belief distinctions.

As discussed earlier there are two types of farming styles: 1. intensive farming with hired labor, and 2. extensive farming with family labor. The major differences represent in B, operational nature's intensiveness behavior $B_1 > B_2$ and their elasticities of inputs for the two types of farmers respectively. Here λ_n is an elasticity parameter for input n where $n = K$ and L. Here $\lambda_k > \lambda_{L1}$ indicates the conventional farmers are capital intensive and $\lambda_k < \lambda_{L2}$ indicates the Amish farmers are labor intensive. Formulation and intuitively discussion is as follow:

$$\pi\ (p, w, m) = max_x\{p * A(m) * B_j * K^{\lambda_k} * L_j{}^{\lambda_{Lj}} - P_t\ *\ K - g_j\ *\ L_j\}. \quad (3.5)$$

A type 1 farmer (conventional) was assumed to be more productive and thus higher farm revenue than a type 2 farmer (Amish), but this type 2 farmer has lower expenditure in labor, so the profit margin might be similar, then two farmers may bid to relatively the same level of farmland price to achieve equilibrium. In this setting, the savings of labor cost for farmer 2 catch up/ offset farmer 1's higher productivity and consequently higher revenue.

The above case may show two types of farmers compete and achieve same equilibrium of farmland prices; however, it is not what we observe in practical and some basic empirical findings where transacted farmland prices tend to be lower in Amish influenced areas than non-Amish influenced areas (shown by graphs in appendix). The two different prices may correspond to different locations, market conditions and land quality. This also aligns with the differential theory of rent of Ricardo's.

3.2.3. Market Clearing Conditions and Equilibrium

Market clearance is based on the supply and demand to be equal and thus result competitive equilibrium, the equilibrium is defined by these following conditions:

1) The marginal rate of technical substitution (MRTS) equals to price ratio.

$$\frac{K'}{L_j{}'} = \frac{P_t}{g_j} \quad (3.6)$$

Under the assumption that the Amish uses relatively larger number of labors than conventional farmers ($L_2 > L_1$) to compensate the lack of machinery use and lower productivity in order to compete. Therefore, MRTS-Conventional is greater than MRTS-Amish, or say

steeper in slope, graphed with blue and orange lines in the left panel of figure 3.1. We illustrate that Amish farmers generally produce less output per land and thus have lower iso-quants. Specifically looking at the two curves that are quite far from each other in the capital axis, conventional farmers who are capital intensive, utilizing the efficiency from machines and other capitals. To the labor axis that we conceptualize the Amish farmers use more labor than conventional farmers. With the trade-off in input, the two groups will be asymptotically converging to have same amount of output, and stay within the same isoquant envelope.

2) Zero profit condition: farmers maximize profits by choosing production quantities Q such that unit cost equal output price for all sectors:

$$p = c(P_t, g_j) = P_t * a_K + g_j * aL_j. \tag{3.7}$$

Where $a_K = \frac{K}{Q}$ and $a_{L_j} = \frac{L_j}{Q}$ are the unit productivity of corresponding input.

3) Input market clearance condition: two inputs: labor L_j is not mobile across farmer types due to religious groups, but farmland K is.

$$\sum_i L_{ij} = L_j = \sum Q * a_{L_j}, Where\ j = 1,2\ and\ L_1 + L_2 = L \tag{3.8}$$

$$\sum_i K_i = K = \sum Q * a_K \tag{3.9}$$

3.2.4. Intuitive Discussion of Equilibrium

We assume that Amish always produce less output than conventional, and we additionally find other sources for Amish be able to compete at the similar level on land markets. One main source is as described above from the lower cost of labor, another main source is that Amish spend less in daily life, i.e. less entertainments. Therefore, there's additional savings from household expenditures. We model two sources of savings into one general savings for simplicity. Equilibrium quantity produced maybe lower or higher for Amish, if profits lower, the savings can make up for the same level of competitiveness. So to conclude this thinking: even if Amish generates less revenue from lower output, they may reach to similar levels of

competitiveness, under two scenarios: (1) similar profits due to lower labor costs; (2) higher savings to compensate. Figure 1 illustrates this by displaying the profit brackets under different input and technology choices.

Table 3.1: Farm Credit East dairy farm 2018 survey data: Conventional vs. Plain Farms

FCE 2018 Dairy Farms	Conventional	Plain Farm
Production per cow lbs./year	25,264	20,974
Pounds milk per worker	1,255,687	832,830
Labor, living, inc. tax $/cwt	3.66	2.5
Breakeven milk price $/cwt	17.7	16.2
Gross farm revenue $/cow	5,104	4,003
Gross farm expenses $/cow	5,021	3,657
Net farm income $/cow	83	346
Return on assets	1.51%	1.43%
Return on equity	2.26%	2.06%

As conceptualized in the left panel of figure 3.1, with the descriptive evidence from representative portion of the Farm Credit East dairy farm survey data in table 3.1. Therefore, we assume conventional farmers generally have higher total costs, outputs and thus total revenues than Amish farmers. However, it's very possible that two groups reach same level of profits. Such phenomenon is also backed by Farm Credit Statistics data, showing that Amish group spend less in capital and have lower revenue, whereas conventional commercial farmers spend more capitals and gain higher revenue, but the two groups' profits are comparable, which provide some evidence that Amish stay competitive with their farming style. Conventional farmers are less elastic with labors and more elastic with capital than Amish farmers, as conventional farmers are less comparatively advantageous with labors due to limited family labors and growing costs with hired labors. The situation of similar level of profits between two groups are illustrated in the right panel of figure 1. The core of the farming style difference is that Amish input more labor and less capital than conventional farmers is still competitive.

3.3. Data Source and Descriptive Statistics

Our dataset includes farmland transactions from 1999 to 2015 of New York state, obtained from New York's Office of Real Property Tax Services, which gathers the data from property assessment offices in each county. We have 21,137 total raw observations in the dataset,

foreach transaction we have information such as sales price, acreage, and detailed location and at the parcel level. We then append the location matched Old Order Amish population data that is estimated based on the Amish church districts. The variable is named as Amish Trend and is constructed to indicate the number of Amish Churches within the 10 miles radius of the farmland parcel transaction observations. The raw data of Amish settlement and church groups was gathered from directory of ministers (Raber's Almanac) for 1995 to 2015.

Some observations are dropped due to: outlier prices that are higher than 50,000 (likely buildings and improvements) and lower than 100 dollars per acre; as well as missing important attribute especially detailed spatial identification (latitude and longitude). There were in total of 17,481 land transactions left after dropping. We tested various sub periods to evaluate temporal differences. We chose to conduct sub-period 2007-2015 instead of the full sample from 1999 to 2015, because of three main considerations: (1) Amish communities were still settling in and developing in the first half decade of 2000s as shown in figure 3.3; (2) agricultural commodities price especially corn price has been surging since 2007; and (3) global financial crisis started around 2007 where real estate markets became so volatile.

In below, table 3.2 displays a selected portion of the summary statistics. The entire list and statistics by year are in the appendix. Figure 3.2 displays the substantial Amish population growth and land purchases from 1999-2015. Emerging trend of land purchases in Amish influenced areas is particularly obvious in early 2000s shown in figure 3.3. Each dot in figure 4 represents a parcel of farmland sold sometime between 1999-2015, specially colored parcels are within 10 miles of an Amish district that are concentrated regions in New York State.

Table 3.2: Key variables summary statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
Deflated per acre farmland price	6806	4314.251	6430.82	103.516	63607.199
Log deflated per acre farmland price	6806	7.834	.969	4.64	11.06
Amish trend (church districts)	6806	1.078	2.158	0	16.824
Number of parcels in sale	6806	1.398	3.31	0	266
Distance to large farms	6806	15380.294	18065.697	0	139452.17
Distance to New York City(ft)	6806	337480.34	94466.349	61957.836	510181.34

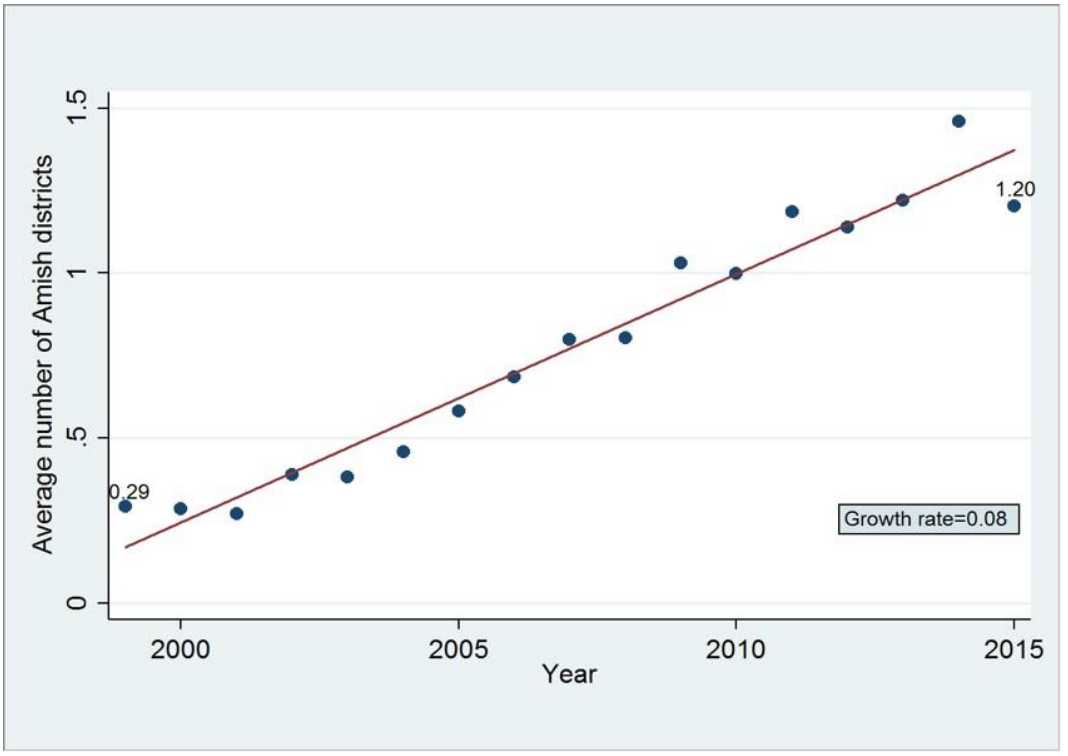


Figure 3.2: Average number of Amish districts located with 10 miles of an arms-length NY farmland transaction (Ifft and Gao, 2019)

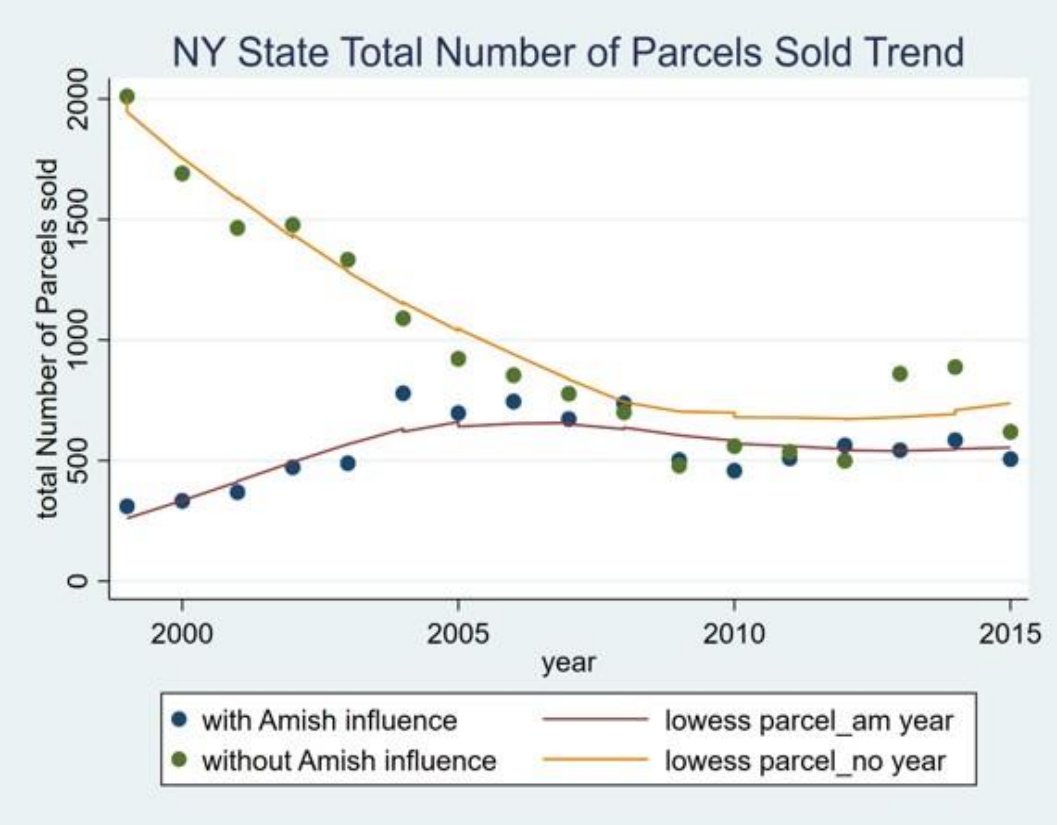


Figure 3.3: New York state total number of parcels sold trend

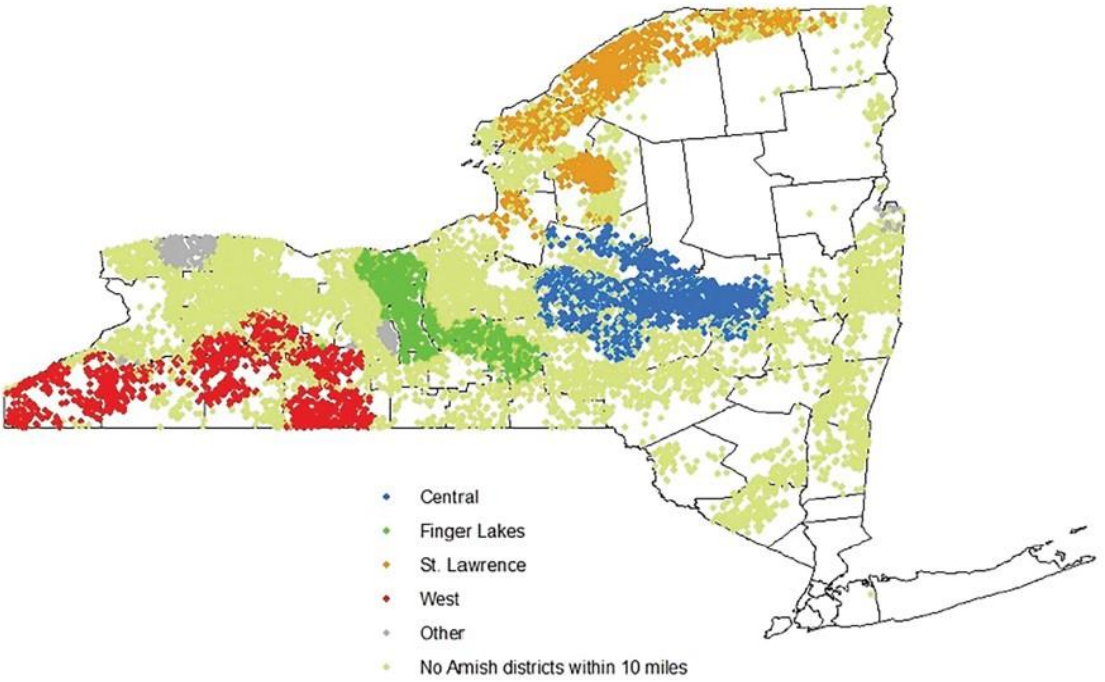


Figure 3.4: New York farmland sales with nearby Amish districts, 1999-2015 (Ifft and Gao, 2019)

3.4. Empirical Model and Identification Strategy

This section discusses our empirical methods and the identification strategy. We use a standard hedonic approach to model farmland prices with influencing factors including land characteristics and non-agricultural characteristics, that reflects the potential income be generated on this land based on the capitalization model conceptual framework earlier. Among these factors, our key explanatory variable (denoted as Amish Trend) is the number of Amish church districts within the 10-miles radius of farmland transactions. The number of church districts could represent Amish population growth as each Amish church has a cap for about 30-40 families or it will divide to form a new church district (Ifft and Gao, 2019). This religious group of farmers have certain farming practices that differ from conventional farmers, for which we cannot quantify and thus we consider these farming practices and related characteristics (lower farming technology use, larger family and lower labor costs, high savings or less expenditure lifestyle, and community risk sharing) to be composed and reflected in their community and population growth. We construct a shift share like enclaving instrumental variable (IV) approach, to address endogenous choices of settlement and farmland purchases in analyzing the impact of Amish population growth on farmland prices.

3.4.1. Standard OLS Model: Farmland Price and Amish Population

We use farmland price as the key dependent variable, and Amish population influence as the main explanatory variable. The model includes 31 control variables with agricultural productivity as well as nonagricultural influences, such as distance to urban areas, small towns, and recreation water. We select control variables based on the LASSO estimator developed by Ifft and Yu (2019), who employed least absolute shrinkage and selection operator (LASSO) method to select among a total of 192 variables representing parcel characteristics. These abundant variables include a large portion that indicate soil quality and infer productivity, that represent the farmland value. Many of these factors appear to be highly co-linear, thus the LASSO selection they've conducted deliver a reasonable solution to reduce issue of overfitting and thus keep the most necessary control variables. The model also uses region[27] and year effects to control for unobserved spatial and temporal factors that may influence farmland markets. We start with estimating the standard OLS regression model, the baseline regression equation is expressed in below:

$$P_{it} = \beta_0 + \beta_1 AT_{it} + \beta \boldsymbol{X_i} + \tau_t + \gamma_s + \varepsilon_{it} \tag{3.10}$$

Where P_{it} is the deflated and natural logarithm-ed price of farmland per acre of farmland i in year t. AT_{it} is the constructed Amish trend variable that indicates number of Amish Churches within the 10-mile radius of the transacted farmland i in time t. X_i is a vector of farm and operator characteristics that contain our control variables, including geographic characteristics, soil characteristics, and nonagricultural characteristics. τ_t and γ_s are year and

[27] https://esd.ny.gov/regions specifies 9 regions in New York state. Region 1. Western New York: Allegany, Cattaraugus, Chautauqua, Erie, Niagara. 2. Finger Lakes: Genesee, Livingston, Monroe, Ontario, Orleans, Seneca, Wayne, Wyoming, Yates. 3. Southern Tier: Broome, Chemung, Chenango, Delaware, Schuyler, Steuben, Tioga, Tompkins. 4. Central New York: Cayuga, Cortland, Madison, Onondaga, Oswego. 5. Mohawk Valley: Fulton, Herkimer, Montgomery, Oneida, Otsego, Schoharie. 6. North Country: Clinton, Essex, Franklin, Hamilton, Jefferson, Lewis, St. Lawrence. 7. Capital Region: Albany, Columbia, Greene, Saratoga, Schenectady, Rensselaer, Warren, Washington. 8. Mid-Hudson: Dutchess, Orange, Putnam, Rockland, Sullivan, Ulster, Westchester. 9. New York City: Bronx, Kings, New York, Richmond, Queens.

region fixed effects. ε_{it} is a white noise error term. Year effects allow us to capture farmland and crop price movements in each year and crop season differences due to weather and macroeconomic situations. Given the general increase in commodity prices across our study period, the year effect is an important control for these trends. For this and all other models we estimate standard errors to be robust to correlation at the county level. Acre weights are used following Bigelow et al. (2020) that identifies different share of large and small farms and their heterogeneous participation's impact on land markets.

3.4.2. Potential Endogeneity

The key identification challenge is the potential endogeneity of the Amish population growth resulting from non-independent change in farmland prices. There are two main possible sources of endogeneity problem that could bias the estimate in different directions. (1) Reverse causality: the Amish farmers make decisions related to farmland purchases and settlement locations simultaneously. Amish tends to settle in areas with lower prices land (Ifft and Gao, 2019), and we find basic evidence as shown in figure 5. In other words, the low-price land area attracts Amish and reversely causing the growth of Amish population. Therefore, the farmland price and Amish population growth could be negatively correlated, and we may underestimate the impact. (2) Correlated unobservable: the number of Amish districts may be correlated with unobservable factors that are related to farmland sales; as Amish are known to carefully consider farmland market characteristics when they buy land (Johnson-Weiner, 2017). It is possible that land purchase decisions and prices are influenced by unobserved, time varying farm- or market-level factors. Some factors may be among the commonalities of their concentration that we discussed earlier in section 1, i.e. nearby small commercial centers and sparse population, conducive to small-scale farming (Anderson and Kenda, 2015). Along with other unobserved factors, some maybe positively and others negatively correlated with farmland prices. Therefore, the estimator could either be underestimated or overestimated. Considering these possible sources of endogeneity, we implement enclaving IV to address the endogeneity and evaluate the impact.

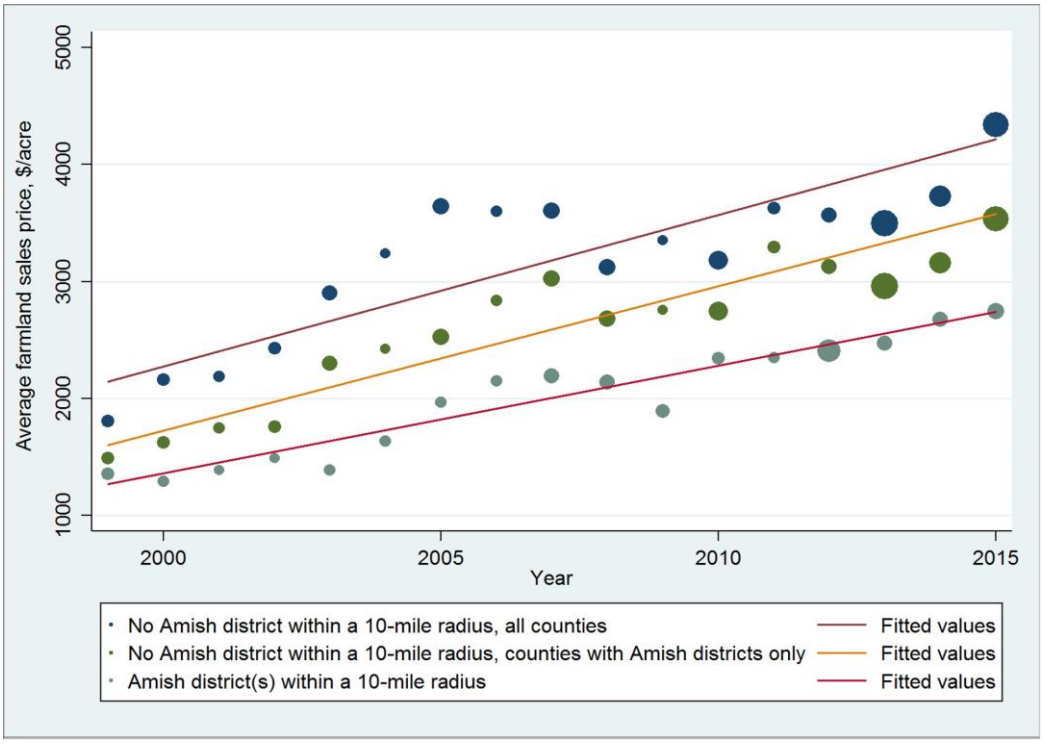


Figure 3.5: Amish Settle in Lower Farmland Price Areas in NY

3.4.3. Instrumental Variable Identification Strategy

To address the concern of non-random choices of Amish settlement area and the potential endogeneity issue, we conduct a two-stage least squares (2SLS) regression, with considering Amish enclaving as an instrumental variable. We constructed this enclaving IV (a shift-share like IV) based on a well-established approach in the labor economics literature, e.g. Lewis and Peri (2015) and Jaeger et al. (2018). The identification basis for these shift-share IVs is that immigrants tend to settle in areas near people from their home country (Bartel, 1989). We refer to this concept and assume there is a similar pattern for the Amish population, that we also find support from the anecdotal evidence. Amish and other religious communities tend to move into places where the territory already have their communities live in. This religious tie and closer settlement choice pattern of Amish has eventually established their enclaves, which validates the inclusion restriction assumption of the IV.

This instrument derives from the settlement patterns of Amish population across New York State 10 years prior to our evaluation of land markets, that has evolved overtime as Amish population to grow. The enclaving instrument's exogeneity assumption hinges on the fact that

the earlier location choice of migrated Amish population is exogenous to the current land competition conditions. This reasonably satisfies the exclusion restriction in the underlying assumptions of the IV. In the exact construction, we use the initial share of Amish church districts in a county to the state in 1995 multiplied by change in the number of church districts from 1995 to 2005. C_j^{95} is the number of church districts of county j in 1995. The IV aims to identify the variation at the county level where the Amish settlement differs from county to county. Because the IV contains the ”shift” which is the population change from year to year, thus the temporal variation of Amish settlement is also captured.

$$Z_{ij} = \frac{C_j^{95}}{\sum C_j^{95}}(C_j^{05} - C_j^{95}). \qquad (3.11)$$

Another instrumental variable was considered, the ratio of county farm real estate value per acre to corn yield in 1997 (the earliest year available), to proxy for farmland affordability. The data is obtained from USDA Census and the construction is at the county level. The real estate value to yield ratio is becoming more common to instrument for farmland affordability with the available data especially after the 2000s. The land affordability is assumed to be carefully considered by Amish when they migrate in and grow their communities, while the historical affordability may be relatively independent from recent farmland conditions.

3.5. Empirical Results

In this section, we first present the estimated impact of Amish population trend on farmland prices using the standard OLS model. We then present the result of Two-stage Least Squares (2SLS) method, with using the constructed enclaving instrumental variable, that addresses the endogenous issues in non-independent area choices between Amish settlement and farmland purchase. All model specifications presented in section 3.5 include all 31 control variables, while we omit presenting them in the tables here, the entire list of results is provided in the appendix. The control variables are determined according to the LASSO estimator developed by Ifft and Yu (2019), who select among a total of 192 variables representing farm parcel

characteristics. All these model specifications also include year and region fixed effects, to capture the within year and region correlation among observations.

3.5.1. Standard OLS Model Results

Table 3.3 shows the regression results of baseline standard OLS model specifications in Equation 10. The model reveals that Amish population trend does not have a statistically significant relationship with farmland prices. The key variable is the constructed population indicator denoted as Amish trend. We sequentially add a vector of farm and operator characteristics that contain our control variables, including geographic characteristics, soil characteristics, and non-agricultural characteristics. Considering endogeneity of large farms, we add CAFO variables (distance to large farms and number of large farms nearby) in column (2), which also appears to be not significant. We suspect small acre sales have more than necessary weights that may swing the coefficients even with using acre weights specified in column (1) and (2). Thus, we consider excluding transactions with acres smaller than 5 and 10 specified in column (3) and (4). While dropping small acre sales slightly increases statistical significance, we should not take the impact to be verified yet. Because of the remaining endogeneity issue, that area choices of Amish settlement and farmland purchases are endogenous, this concern will be addressed with an enclaving IV in the next section.

Table 3.3: Regular OLS regression results

2007-2015	Standard OLS Regresion			
Variable	Standard	+CAFOs	Drop <5 Acres	Drop <10 Acres
Amish trend	-0.0129	-0.0123	-0.0192**	-0.0174*
	(0.0126)	(0.0126)	(0.0088)	(0.00931)
Constant	9.143***	9.061***	9.510***	9.202***
	(0.472)	(0.479)	(0.46)	(0.435)
Observations	6,806	6,806	6,417	5,950
R-squared	0.309	0.309	0.282	0.311
Acre Weighted	YES	YES		
County Cluster	YES	YES	YES	YES

Robust standard errors are shown in parentheses; asterisks denote ***α≤0.01, **α≤0.05, *α≤0.10

3.5.2. Enclave IV Two-stage Least Squares (2SLS) Model Results

With including the enclave IV (shift share like IV, constructed in Equation 11) and running two-stage least squares (2SLS) to address the possible endogeneity issues in Amish settlement and

farmland purchase area choices, we see Amish trend still does not have a statistically significant relationship with farmland prices, shown in Table 3.4. In column (2) we can see the first stage is statistically significant with F stat above 10, which indicates it's not a weak IV and satisfies the inclusion restriction that Amish enclaves in earlier years is closely related to the later Amish population growth trend. The IV model results have slightly lower (more negative) coefficients than the OLS model results in Table 3.3. This may align with our one of earlier conjectures that without addressing the endogeneity, our estimator in the OLS model could be overestimated or say biased upward, which may mistakenly conclude that Amish population growth drives up farmland prices. However, the estimator bias direction could be ambiguous due to other unobserved endogeneity issues illustrated in section 4.2. The standard error for both sets of models in Table 3.3 and 3.4 are relatively large and the varying range will cover the 0 point around the point estimator, which further enhancing the indication of the estimator being not statistically significantly different from zero.

Table 3.4: Enclave IV two-stage least squares (2SLS) regression results

2007-2015	IV Enclave Shift Share		
Variable	IV	1st Stage	+CAFOs
Amish trend	-0.0418		-0.0409
	(0.0431)		(0.0437)
IVZ_Enclave		0.962***	
		(0.102)	
Constant	9.124***	2.756*	9.051***
	(0.48)	(1.425)	(0.486)
Observations	6,806	6,806	6,806
R-squared	0.305	0.431	0.306
Acre Weighted	YES	YES	YES
County Cluster	YES	YES	YES

Robust standard errors are shown in parentheses; asterisks denote ***α≤0.01, **α≤0.05, *α≤0.10

Table 3.5 is the 2SLS model results with including the enclave IV that drops the small acre observations, which should be viewed in comparison with column (3) and (4) of Table 3.3. In Table 3.3 we see slight statistical significance of the key variable, here including enclave IV the key variable Amish Trend is statistically insignificant, consistent with other specifications in Table 3.4. The coefficients with dropping small acre observations shown in table 3.5 is slightly lower than those in table 3.4, inferring the potential upward bias those small acre

observations may have caused primarily due to non-agricultural uses common to small parcels. However, as Amish farmers tend to be small-scale and thus purchase small parcels, we cannot be certain whether these small parcels bias directions are resulted from commercial uses or Amish farmers. Checking all these alternative specifications solidify our finding that Amish population does not have a statistically significant relationship with the farmland prices.

Table 3.5: Enclave IV (dropping small acres)

2007-2015	IV1 No Acre Wgts but Excluding Small Acres (<5 and <10)					
	Excluding Acres <5			Excluding Acres <10		
VARIABLES	IV1	1st Stage	+CAFOs	IV1	1st Stage	+CAFOs
Amish trend	-0.0599		-0.0633	-0.0545		-0.0564
	(0.0387)		(0.0393)	(0.0375)		(0.0382)
IV1_Enclave		0.914***			0.922***	
		(0.125)			(0.123)	
Constant	9.421***	1.305	9.401***	9.124***	1.439	9.125***
	(0.424)	(1.12)	(0.428)	(0.41)	(1.118)	(0.414)
Observations	6,417	6,417	6,417	5,950	5,950	5,950
R-squared	0.276	0.432	0.276	0.305	0.431	0.305
County Cluster	YES	YES	YES	YES	YES	YES

Robust standard errors are shown in parentheses; asterisks denote ***α≤0.01, **α≤0.05, *α≤0.10

What we estimated is the Local Average Treatment Effect (LATE) of Amish population growth on farmland prices, or say the Average Treatment Effect (ATE) for the subgroup of Amish population that are enclaving (compliers who align with our identification and settle close to their existing religious community). Because different instrumental variables may identify particular population subgroup's ATE, and thus different LATE are estimated by various IVs (Imbens and Wooldridge, 2009). The non-compliers (Amish who don't settle into enclaves) could be those who are restricted to settle into the enclaves, for various reasons. For example, it could be due to some local governments implemented strong regulation that is not friendly for later Amish to move in or even for existing residents [28]. This could be a cause for the Amish who want to settle into their related communities but could not. We do not observe these types of restrictions in New York state. Therefore, if all New York State Amish population settled because of enclaves (the compliers are the whole sample), then our LATE

[28] https://www.bridgemi.com/michigan-environment-watch/michigan-county-threatens-bulldoze-amishhomes-poop-dispute

could be equivalent to ATE. Though because we cannot assure the portion of compliers and non-compliers, we cautiously conclude our estimator is the LATE.

3.5.3. Robustness Check: Alternative IV

Table 3.6 displays the results of 2SLS model with using an alternative instrumental variable as a robustness check. This 2nd IV is the ratio of county farm real estate price per acre to corn yield in 1997, a proxy for farmland affordability. We can see the first stage result is not statistically significant and indicates a weak IV problem, hence the slight significance in Amish Trend coefficients are not meaningful. We also tested employing both enclave IV and this 2nd IV but IV2 still remains statistically insignificant, and so do the Amish Trend coefficients. Therefore, we keep the focus on the enclave IV presented in earlier sections as the main models.

Table 3.6: Robustness check: Alternative IV value yield

2007-2015	IV2 Value Yield		
Variable	IV2	1st Stage	+CAFOs
Amish trend	-0.567*		-0.538*
	(0.334)		(0.283)
IV2 value yield		-0.0779	
		(0.0495)	
Constant	8.763***	0.948	8.880***
	(1.418)	(2.404)	(1.283)
Observations	6,806	6,806	6,806
R-squared		0.316	
Acre Weighted	YES	YES	YES
County Cluster	YES	YES	YES

Robust standard errors are shown in parentheses; asterisks denote ***$\alpha \leqslant 0.01$, **$\alpha \leqslant 0.05$, *$\alpha \leqslant 0.10$

3.6. Conclusion

Plain sects, including the Amish, are growing and prospering in several states that have strong commercial/conventional agriculture. Land market activity allows us to better understand this growth and coexistence of different farming systems. This paper analyzes farmland markets and the emerging Amish communities who rely heavily on farming while use limited modern technology.We conceptualize farmland market competition between conventional farmers

who are more capital intensive and the emerging Amish farmers who are more labor intensive. We assume the two groups have similar level of profitability, even if Amish are less productive due to limited technology use. Because Amish typically has larger families thus lower labor costs, and lower capital expenses with using older technology. Then we estimated a series of standard hedonic models for farmland prices, to study the impact of the Amish population growth on farmland market. We find that Amish community density does not have a statistically significant relationship with farmland prices. Even though Amish limit their adoption of modern production technologies, they successfully compete for farmland with conventional farmers. Amish farmers' competitiveness may stem from their religious traditions with larger families and intensive human-nature interactions.

It may improve the model accuracy to include the sociodemographic information for identifying Amish and conventional farmers' backgrounds to further detail their choice mechanism using revealed preferences, though we are lacking access to this type of data. Our study focuses on Amish group who make up a large share of plain communities, though it would be nice to test with other groups as well. While we observe farmland transactions and utilize farmland markets dynamic as the channel to pass through farmer characteristics, we do not have the production input and output information of farmers that may more accurately capture their costs, labor intensiveness, and capital purchases.

Our study contributes to the literature on (1) the inverse farm productivity–size relationship, and (2) competitiveness of small farms. We showcase a way of measuring population and its impact by constructing Amish Trend based on Church groups that are generally capped in number of participants, that can be applied in further research. Our identification strategy mitigates bias from endogenous growth of Amish settlements and may be replicable in other settings.

3.7. References

Anderson, C. and L. Kenda (2015). What kinds of places attract and sustain Amish populations? *Rural Sociology 80*(4), 483–511.

Barkley, P. W. (1986). Barlowe, Raleigh. Land Resource Economics: The Economics of Real Estate, 4th ed. Englewood Cliffs NJ: Prentice-Hall, 1986, x + 559 pp. *American Journal of Agricultural Economics 68*(4), 1030–1030.

Barrett, C. B., M. F. Bellemare, and J. Y. Hou (2010). Reconsidering conventional explanations of the inverse productivity–size relationship. *World Development 38*(1), 88 – 97.

Bartel, A. P. (1989). Where do the new u.s. immigrants live? *Journal of Labor Economics 7*(4), 371–91.

Benjamin, D. (1995). Can unobserved land quality explain the inverse productivity relationship? *Journal of Development Economics 46*(1), 51 – 84.

Bhalla, S. and P. Roy (1988). Mis-specification in farm productivity analysis: The role of land quality. *Oxford Economic Papers 40*(1), 55–73.

Bigelow, D., J. Ifft, and T. Kuethe (2020). Following the market? hedonic valuation using Sales prices versus self-reported values. *Land Economics (forthcoming)*.

Capozza, D. and R. Helsley (1989). The fundamentals of land prices and urban growth. *Journal of Urban Economics 26*(3), 295–306.

Carter, M. R. (1984). Identification of the inverse relationship between farm size and productivity: An empirical analysis of peasant agricultural production. *Oxford Economic Papers 36*(1), 131–145.

Chicoine, D. L. (1981). Farmland Values at the Urban Fringe: An Analysis of Sale Prices.

Land Economics 57(3), 353–362.

Choy, J. P. (2020). Religious rules as a means of strengthening family ties: Theory and evidence from the Amish. *Journal of Comparative Economics 48*(3), 729–748.

Collier, P. (1983). Malfunctioning of African rural factor markets: Theory and a Kenyan example. *Oxford Bulletin of Economics and Statistics 45*(2), 141–172.

Cross, J. A. (2014). Continuity and change: Amish dairy farming in Wisconsin over the past decade. *Geographical Review 104*(1), 52–70.

Decker, C. S., C. S. M. E. J. Wohar, and C. Denny (2014). Religiosity's impact on agricultural investment on the american frontier. *Great Plains Research 24*, 37–46.

Delpierre, M., C. Guirkinger, and J.-P. Platteau (2019). Risk as impediment to privatization? the role of collective fields in extended agricultural households. *Economic Development and Cultural Change 67*(4), 863–905.

Elizabethtown College (2017). Elizabethtown, PA. Amish Population Change 2012-2017. https://groups.etown.edu/amishstudies/files/2017/08/Population Change 20122017.pdf.

Ervin, D. E. and J. W. Mill (1985). Agricultural land markets and soil erosion: Policy relevance and conceptual issues. *American Journal of Agricultural Economics 67*(5), 938– 942.

Featherstone, A. M. and T. G. Baker (1987). An examination of farm sector real asset dynamics: 1910–85. *American Journal of Agricultural Economics 69*(3), 532–546.

Feder, G. (1985). The relation between farm size and farm productivity: The role of family labor, supervision and credit constraints. *Journal of Development Economics 18*(2), 297 – 313.

Feller, I. (1972). Production isoquants and the analysis of technological and technical change.

The Quarterly Journal of Economics 86(1), 154–161.

Gautam, M. and M. Ahmed (2018). Too small to be beautiful?: The farm size and productivity relationship in Bangladesh. *Development Economics: Regional Country Studies eJournal*.

Hill, A. (2006). Fields with dreams: The distribution of farmland with publicly-funded conservation easements in Pennsylvania.

Hoken, H. and S. Qun (2019). An analysis on the inverse relationship between yield and farm size in rural China in the 1930s, edited by Hao Hu, Funing Zhong, and Calum G. Turvey. In *Chinese Agriculture in the 1930s*, pp. 185–188. Palgrave Macmillan, Cham.

Hovinen, G. R. (1995). Heritage issues in urban tourism: An assessment of new trends in Lancaster county. *Tourism Management 16*(5), 381 – 388.

Hynes, J. W. and W. D. Edgington (2018). When everything old is new again: Amish career and technical education. *Bia lostockie Studia Prawnicze 23*(4), 125–133.

Ifft, J. E. and Y. Gao (2019). Old order Amish settlements and New York farmland markets. *The Journal of the ASFMRA*, 100–107. Copyright - Copyright American Society of Farm Managers and Rural Appraisers 2019; Last updated - 2019-12-17.

Ifft, J. E. and A. Yu (2019). The impact of the 2008 shale gas moratorium on New York farmland values. 2019 Annual Meeting, July 21-23, Atlanta, Georgia 291232, Agricultural and Applied Economics Association.

Imbens, G. W. and J. M. Wooldridge (2009, March). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Irwin, E. G., A. M. Isserman, M. Kilkenny, and M. D. Partridge (2010). A century of research

on rural development and regional issues. *American Journal of Agricultural Economics 92*(2), 522–553.

Jaeger, D., J. Ruist, and J. Stuhler (2018). Shift-Share Instruments and the Impact of Immigration. NBER Working Papers 24285, National Bureau of Economic Research.

Jeong, S. (2013). The role of social capital for Amish entrepreneurs in pursuing informal economic opportunities. *Journal of Amish and Plain Anabaptist Studies 1*(1), 127–166.

Johnson-Weiner, K. M. (2017). In *New York Amish Life in the Plain Communities of the Empire State*. Cornell University Press, Ithaca, NY.

Johnston, B. F. and J. W. Mellor (1961). The role of agriculture in economic development. *The American Economic Review 51*(4), 566–593.

Johnston, R. J. and J. M. Duke (2009). Willingness to pay for land preservation across states and jurisdictional scale: Implications for benefit transfer. *Land Economics 85*(2), 217–237.

Kraybill, D. B. (2021). Amish, https://www.encyclopedia.com. Web. 07.04.2021.

Kraybill, D. B., K. M. Johnson-Weiner, and S. M. Nolt (2013). *The Amish*. Baltimore: Johns Hopkins University Press.

Kraybill, D. B., S. Nolt, and E. J. Wesner (2011). Sources of enterprise success in Amish communities. *Journal of Enterprising Communities: People and Places in The Global Economy 5*, 112–130.

Kreps, G. M., J. F. Donnermeyer, C. Hurst, R. Blair, and M. Kreps (1997). The impact of tourism on the Amish subculture: A case study. *Community Development Journal 32*(4), 354–367.

Kreps, G. M., J. F. Donnermeyer, and M. W. Kreps (1994). The changing occupational structure of Amish males1. *Rural Sociology 59*(4), 708–719.

Lewis, E. and G. Peri (2015). Immigration and the economy of cities and regions. In *Handbook of Regional and Urban Economics*, Volume 5, Chapter 10, pp. 625–685. Elsevier.

Melichar, E. (1979). Capital gains versus current income in the farming sector. *American Journal of Agricultural Economics 61*(5), 1085–1092.

Michler, J. D., E. Tjernstr¨om, S. Verkaart, and K. Mausch (2019). Money matters: The role of yields and profits in agricultural technology adoption. *American Journal of Agricultural Economics 101*(3), 710–731.

Miranowski, J. A. and B. D. Hammes (1984). Implicit prices of soil characteristics for farmland in Iowa. *American Journal of Agricultural Economics 66*(5), 745–749.

Moss, C. and A. Schmitz (2003). *Government Policy and Farmland Markets: The Maintenance of Farmer Wealth*. Wiley-Blackwell; 1 edition (October 6, 2003).

Muyanga, M. and T. S. Jayne (2019). Revisiting the farm size-productivity relationship based on a relatively wide range of farm sizes: Evidence from Kenya. *American Journal of Agricultural Economics 101*(4), 1140–1163.

Nickerson, C. and W. Zhang (2014). Modeling the determinants of farmland values in the United States. In *The Oxford Handbook of Land Economics*, Chapter 5, pp. 111–138.

Peter Timmer, C. (1988). Chapter 8 the agricultural transformation. *1*, 275–331.

Rada, N., W. Liefert, and O. Liefert (2019). Evaluating agricultural productivity and policy in Russia. *Journal of Agricultural Economics 71*(1), 96–117.

Reid, J. (2015). Old order Mennonites in New York: Cultural and agricultural growth. *Journal of Amish and Plain Anabaptist Studies 3*(2), 212–221.

Ricardo, D. (1817). *On the principles of political economy, and taxation*. J. Murray.

Rios, A. R. and G. E. Shively (2005). Farm size and nonparametric efficiency measurements for coffee farms in Vietnam. (19159).

Sen, A. (1962, 27). An aspect of Indian agriculture. *Economic and Political Weekly Vol. 14*(Issue No. 4-5-6).

Sheng, Y., J. Ding, and J. Huang (2019). The relationship between farm size and productivity in agriculture: Evidence from maize production in northern China.

Tauer, L. (2001). Efficiency and competitiveness of the small New York dairy farm. *Journal of Dairy Science* 84(11), 2573 – 2576.

Wang, L. C. (2020). Religious prohibition and sacrifice: evidence from the Amish restriction on high school education. *Journal of Demographic Economics 86*(3), 403–434.

Zook, L. (1994). The Amish farm and alternative agriculture. *Journal of Sustainable Agriculture 4*(4), 21–30.

4. CHAPTER 4

The Willingness to Offer Livestock Insurance in Rural China: A Discrete Choice Experiment Among Chinese Insurance Agents

This study implements discrete choice experiments (DCE) and survey among Chinese agricultural insurance agents, to investigate their Willingness to Offer (WTO) livestock insurance based on the variations of eight main attributes of livestock insurance. So far as we are aware this is the first DCE designed around the supply of insurance products with the subjects being insurance agents, marketers, and executives. The experiment primarily evaluates how government subsidies affect agents' WTO. We find that the effect is enormous; a one level increase of subsidy leads to 3.166 times higher probability to offer. This subsidy effect is important as it confirms the endogenous structure between price and quantity in insurance offering, where subsidy does not only incentivize demand, but also the supply. Another main factor of insurance investigated, is the impact of different coverage types on agents' WTO. We find that agents prefer mortality insurance the most, then followed by revenue insurance and profit insurance, while Index-Based Livestock Insurance (IBLI) is the least preferred to offer. Agents' knowledge about these newer types of insurance supports their WTO as well, thus proper education is necessary to promote the more advanced types of livestock insurance. Our findings and contribution are critical to the reform and improvement of livestock insurance.

4.1. Introduction

Agricultural insurance is an important tool to support agricultural development and poverty reduction. Nevertheless, countries face various challenges in providing and operating agricultural insurance products that farmers really want and benefit from. The demand for agricultural insurance has been extensively studied (Kong et al, 2011; Chantarat et al, 2007; Belasco et al, 2019), but research on the supply side of agricultural insurance is rare and understudied. Yet to fully understand the issues of delivering agricultural insurance in limited

resource or developing agricultural economies understanding both supply and demand is critically important. In this paper, we implement a novel choice experiment to explore the attributes of supplying or selling agricultural insurance, and in particular livestock insurance, by insurance industry brokers and managers in China. China, the world's second-largest agricultural insurance market in terms of size, is still far from mature. China has developed agriculture insurance for almost four decades and experienced two main phases (1982-2002, and 2003-present). The take-up during the first phase was quite low and actually started decreasing after 1992, mainly because insurers were not interested in offering agricultural insurance products as they are unprofitable (www.agroinsurance.com). The second phase initiated multiple pilot subsidy programs in 2003. Particularly, the state council's declarations to emphasize the importance of agricultural insurance in 2006 and 2007 boosted product offerings and farmers' participation. Similar to the U.S., agriculture insurance provided in China is also primarily led by government agencies and state-owned companies. Therefore, governmental policy and structural guides in the design and operation of insurance products play a critical role in insurers' willingness to offer. Current insurance premium subsidies seem to be mostly assisting farmers rather than insurers (Global Facility for Disaster Reduction and Recovery, 2020), while the endogenous structure of supply and demand will show us how premium subsidies will increase supply incentives.

A recent state council document of China aims to provide guidance to expedite the development of high-quality agricultural insurance (China Government, 2019-#102). The guide has targeted specific levels of major insurance attributes as a development goal, with a special focus on developing revenue insurance and insurance-futures synthetic products (a form of price index insurance). This focus was based on the pilot experience that found some particular insurance types (weather index insurance, input cost protection, or output protection) were inefficient in operation or insufficient to support farmers. Specifically, some of this inefficiency comes from supplier scalability considerations. To understand the supply of agricultural insurance and contribute to policy designs of being more optimally providing

insurance, we investigate insurers' willingness to offer (WTO) livestock insurance29 in rural China. Precisely, we determine how core attributes of mainstream existing and piloting insurance products (through discrete choice experiment) and provider characteristics (via a survey) affect their WTO. Discrete choice experiment (also referred to as conjoint analysis) is widely applied in consumer choice research, and has proven to be a valid method to reveal the participants' preferences through combination and rank of attributes and levels. To a much lesser extent choice experiments have been used to investigate agricultural insurance (Wang et al, 2020, and Shee et al, 2021). With the exception of Sheet et al. (2021), we believe this is the first paper to explore offerings of agricultural insurance from the supply side. Within the experimented attributes, we extensively evaluate how government subsidies on the insurance premium costs affect insurers' WTO. We also develop a conceptual framework to demonstrate this effect.

There's abundant research on agricultural insurance, covering a wide range of insurance product types that study their production impact, welfare improvement, cost-effectiveness, willingness to pay (WTP), actuarial design, and operational difficulties. Insurance provision is generally considered to allow farmers to invest in higher-return production activities that are usually riskier (Cole, Giné, and Vickery, 2017). With increased insurance coverage, farmers could increase their borrowing and invest in expanding production of insured crops, which support longer-term development and improve farmers' welfare (Cai, 2016). Financial innovations in insurance products that help pool and diversify risk serve an important role, that stabilizes farmers' income, reducing poverty, while boosting agricultural economic development. However, the success of this boosting potential depends on further improvements to product design and delivery that can alleviate "barriers" to insurance uptake, including high premium price, low coverage level, administrative costs, limited understanding, basis risk, and distrust of insurers (Cole et al. 2013; Clarke 2016; Rampini and Viswanathan 2015).

[29] We emphasize the research on livestock insurance because it has less mature development than crop insurance. Also because of the timing, both farmers and insurers as well as the government realized the importance more after the African Swine Fever outbreak during 2018-2019 in China.

The evolution of agricultural insurance includes the invention, design, and implementation of various types of insurance products that address different perils: from variable input cost, output quantity, and output price; to weather index, revenue, and margin. Hypothetically revealing in the product design, weather index insurance displays advantages over other insurance products in addressing moral hazard and adverse selection concerns; it also costs less administrative fees in loss assessment but requires precise data and analysis (Mahul and Skees, 2007; Barrett, Mude, and Turvey 2007; Belasco, Cooper, and Smith, 2019). Chinese farmers expressed interest in precipitation insurance (Turvey, Kong, and Belltawn, 2009), though the insurers found weather index insurance challenging to operate because of the basis risk and unavailability of precise weather data (Jensen et. al, 2016; Long, 2021; Woodard and Garcia, 2008). Revenue and margin insurance gained more attention among Chinese policymakers, as they observe the market risk is also hurting farmers significantly in addition to natural damages. Revenue insurance also outplays other products as the premium is generally lower because output quantity and price are inversely correlated, creating what is referred to as a natural hedge (Skelton and Turvey, 1994; and Turvey, 1992). Livestock insurance shares a similar development trend and research findings that Index - Based Livestock Insurance (IBLI) and revenue insurance are the new pursuits in several developing countries (Chantarat et al, 2013; Skees and Barnett, 2006; and Belasco et al, 2015). This is especially the focus in China because the African Swine Fever (ASF) epidemic in 2018 – 2019 has killed many pigs and made the hog market volatile. A large number of farmers suffered significant losses to reach bankruptcy, requiring the Chinese government to intervene and provide assistance compensation to them.

Most studies about insurance participation focus on the demand side, that is to understand farmers' adoptions and reactions to agricultural insurance (Wang et al., 2020). Farmers are generally not willing to pay for insurance that above 30% of its actuarial premium, the rest is generally subsidized by several levels of government (Long, 2021). The demand for drought insurance is empirically shown to be downward sloping and elastic, substantial premium subsidies are needed for wide adoption (Coble and Barnett, 2013; Yu and Sumner, 2018; Miranda and Farrin, 2012; Kong et. al, 2011). Kong et. al (2011) also show that farmers'

risk averseness and WTP are higher when variations in revenues and yields increase. Insurance take-up may be influenced by recently experienced events, as farmers tend to be more attentive to the benefit of insurance after losses and payments have occurred and financially educated (Cai, 2020; Cai et al., 2021).

Our study makes several contributions. We contribute to the agriculture insurance participation literature, especially in the rarely studied supply side. In a closely related field of insurance-linked credit products, there are conflicting demand and supply-side preferences (Shee et al, 2021). Our work hopes to investigate the preference differences between the two sides which can lead to better agricultural insurance product design. This study provides further validation in the DCE research method that could potentially be applied to different analyses: using choice experiments to study insurers and reveal their preferences, through combinations of various levels of core attributes for insurance products. Identifying the desired attributes and levels for insurance companies can provide optimal policy guidance in insurance packaging design that motivates market participation and benefits agricultural development. Among these attributes, we intensely analyze the impact of subsidies on insurers' willingness to offer, both conceptually and empirically. This would support the investigation of optimal subsidy levels, which is among the top interests of researchers and policymakers in regard to improving the cost-effectiveness of government funds.

It still remains questionable whether newer types of insurance products (revenue and index type) are going to be beneficial and widely adopted by farmers in China and other places. Because of challenges faced by farmers and insurers with respect to implementing weather index insurance, it is important to consider and evaluate alternative types of insurance including revenue insurance, livestock insurance, and margin insurance. Indeed, insurance agents would most likely be in a position to expand their book if they had greater flexibility in their offerings of a suite of insurance products to farmers depending upon the specific perils they faced. In addition, our study can lead to improvements in the general governance of insurance markets, particularly with respect to concerns by insurers in regards to agriculture insurance being unprofitable, county financial departments don't have enough funds to meet

the required portions of the premium subsidies, and the central government now focuses on promoting the development of revenue type insurance.

The rest of the paper is structured in the following way. In section 2, we introduce the choice experiment design and empirical model for estimating supply preferences. Section 3 lays out data that we collected from the field and presents empirical analysis including interpretations. In section 4, we conclude the paper and discuss the economic implications of our work for product design and subsidy policy to improve participation, overall welfare, and cost-effectiveness.

4.2. Empirical Methodology

4.2.1. Discrete Choice Experiment Method

In this study, we use stated preference discrete choice experiments (DCE) to derive willingness to offer (supply side) for various attributes on livestock insurance. Respondents' preferences are uncovered from decisions to hypothetical choices rather than actual market decisions. In a DCE, respondents are given pre-designed choice cards and asked to choose one of several alternatives (choices) containing combinations of different levels of attributes. Exogenous variation of certain attributes permits measurement of the marginal values for various attributes that affect non-monetary choices of insurance supply and can be used as an end in itself, or as the first step in the development of more costly randomized control trials (RCT) or small-scale pilot programs. This hypothetical design includes product design choices with a wide range of variations. Thus, for these designed insurance products, their adoption decisions are not directly observable without the hypothetical approach. This feature of DCE allowing us to evaluate supply preferences for attributes and is particularly relevant when revealed preference approaches for which preference estimations are impossible to measure. In fact, this advantage of DCE over revealed preference methods has been examined by Adamowicz et al. (1998), Carlsson and Martinsson (2001), and Lusk and Schroeder (2004), who found no statistically significant difference between the two approaches. DCE provides exogenous variations that can address endogeneity concerns, such as endogenous structure

between premium and coverage level that may lead to biased elasticity estimates in empirical analysis (Woodard and Yi, 2020). A choice experiment, for example, can exogenously alter different levels of premiums and coverage to test elasticity.

The theoretical foundation of DCE is based on the attribute approach of the Lancastrian theory of demand to derive individuals' utilities and preference for a good via its attributes that can be decomposed from a product (Lancaster, 1966). In our setting, the good of interest here is the livestock insurance, which can be considered as a gathering of its attributes including coverage level, risk environment, actuarial premium rate (with loading), premium subsidy, coverage types (index, mortality, revenue, margin), catastrophic and property-casualty inclusion. The tradeoff between various insurance attributes can thus be examined through the choices that insurance agents have available. In addition, as we include premium price attributes in the experimental design, the willingness to offer (WTO) for other non-price-related attributes can then be estimated through the variations of different levels of these attributes (McFadden, 1974, 2001).

4.2.2. Choice Cards Design and Survey Questions

The choice experiment of this study is based on the D-optimal approach, a 6-block design, with 15 cards per block and 2 choices per card30. We do not provide an opt-out option and ask the respondents to choose the less disliked choice if neither are of interest. Each card is comprised of eight attributes as listed in Table 4.1. We determined the attributes and levels based on the existing and piloting insurance products, and provide ranges above and below current levels. We also consulted with 15 insurance experts and policymakers during the testing round of our study, which enhanced the practicality of our design to support future policy evaluations. For example, we cannot have too high of a claim starting point or too low of government premium

[30] We designed the choice cards using JMP software, consumer choice experiment package. The D-Optimal design is an algorithmic approach to maximize the determinants of the information set used in the design of experiments with multiple treatments. In this study, a fully orthogonal design across all attributes and levels as described in Table 1 would require all insurance agents to complete all possible combinations which would be 240,000 sets (8x5x3x2x5x5x4x2x5). A D-optimal design reduces the number of possibilities to 6 blocks of 15 cards with 2 choices per card (in our study) to 180 possible combinations. It is non-orthogonal in that each agent selects from only 30 out of 180 choices rather than 240,000.

subsidy. Because most agricultural insurance is generally classified as government-supporting insurance, that is meant to benefit the poor and empower the strength and resilience of the agricultural economy; but not commercial products that are primarily intended to seek profits.

Unlike many experimental studies that are conducted in controlled laboratories with university students and staff, our study was performed with actual insurance agents who design, sell, and operate livestock insurance. The original plan was to be in China and conduct in-the-field interviews with insurance agents onsite in their company offices or farms. However, due to Covid-19 travel restrictions, we were not able to travel overseas, and thus we have conducted the interviews mostly virtually (either video or audio conference calls) 31. During the search and contact of potential livestock insurance agents, we were aided by researchers and officers at the China Ministry of Agriculture32 and graduate students at China Agriculture University. Due to the time difference between the U.S. and China, sometimes we face the difficulty of scheduling. Therefore, a portion of the experiments was conducted by these graduate students, who received extensive training from us and shadowed several of our interviews to make sure the experiments they did are consistent.

Table 4.1: Attributes and Levels used in Discrete Choice Experiment.

Attribute	Description	Unit	Level
1. Claim starting point (Deductible, or coverage level = 1 – deductible)	Minimum damage percentage to get insurance compensation. E.g. if the damaged coverage is set at 20%, when actual damage is greater than 20%, then the portion that exceeds 20% will be compensated by the insurance company. E.g. asset value is ￥1000.	% of asset value	10% (protect ￥900); 15% (protect ￥850); 20% (protect ￥800); 25% (protect ￥750); 30% (protect ￥700).
2. Risk environment	Projected risk environment (both output and price)	/	Low; Moderate; High.

[31] Through various channels of contact, we reached out to 20+ companies and over 320 agents (primarily through social network and some via conferences). For which 211 of them agreed to participate in our study. Our local assistant in China was able to go visit some agents onsite and conducted 6 experiments of their reach capable. Even though China Covid cases were low, the travel was still very restrictive domestically during the time of study.

[32] Special and sincere gratitude to Dr. Wenjun Long, researcher in agricultural insurance. Director of Agricultural and Rural Affairs, China Ministry of Agriculture.

3. Market Disaster Pay	Get paid in extreme cases of market conditions. E.g., Lump Sum payout if hog/corn price ratio below 2:1.	/	Included, Not included
4. Premium actuarial (before subsidy)	Price of farmers need to pay in advance to obtain certain insurance products. (E.g. 1% is ￥9 for protecting 900, ￥8.5 for protecting 850; 15% is ￥135 for protecting 900)	% of protected value	1%, 3%, 6%,12%; 25% of the protected value.
5. Premium subsidy by gov and org	Premium subsidy by gov (including central, provincial, county government and all sources)	% of premium actuarial	50% (farmer pays 50%); 60% (farmer pays 40%); 70% (farmer pays 30%); 80% (farmer pays 20%); 90% (farmer pays 10%).
6. Coverage types	1. IBLI (rainfall, temperature index, vegetation); 2. Mortality, # of animals (live or death); 3. Gross revenue (Price*Live Weight) or price/pig; 4. Livestock margin insurance (Hog price*Live Wgt – quantity fed*price of feed).	/	IBLI (Weather-based); Mortality (death only); Revenue or Price; Net Margin (or ratio)
7. Catastrophic & P&C coverage	Uninsured catastrophic events (war, terrorism, earthquake, mudslide, typhoons). Property Casualty (P&C): Theft, Fire.	/	Included, Not Included
8. Loan Allowance	Can be used as loan collateral up to a certain level of insured or protected value.	% of protected value	0%, 30%, 50%, 70%, 90% of the protected value

In order to reach the statistical significance and rank condition of DCE, we adopted a protocol generally used in choice experiments design (Orme 1998; Johnson and Orme 2003; and Rose and Bliemer 2013) to determine the minimum sample size:

$$N \geq 500\left(\frac{l^*}{T \times A}\right) = 500\left(\frac{5}{15 \times 2}\right) = 83.333$$

Where *T* is the number of choice cards tasked to each respondent (15, in our case), *A* is the number of alternatives per choice card (2 in our case), and l^* is the largest number of levels of any of the attributes (5 for each choice). The rank condition is satisfied for N=83.333 rounded up to integer as 84, with 14 agents per block for 6 blocks. Our sample included 211 insurance agents from 13 different provinces, which well exceeds the minimum sample size

requirement. We have 35 agents per block evenly distributed, except block 3 which had 36 due to random allocation. The timing of our experiment was December 2020 to March 2021. Most agents insured pigs for farmers and facilities; some of them insure cattle, sheep, poultry, and fish; as well as some minor specialties such as horse and camel.

Our terminology and attributes are based on actual terms used in Chinese livestock insurance contracts and would be familiar to most agents. As previously mentioned, we consulted with industry experts for fine-tuning the design and wording. The claim starting point is the reverse to the coverage ratio used in the United States agricultural insurance products. It represents the minimum damage ratio of the protected value to trigger insurance payment. A claim starting point of 10% is equivalent to 90% coverage in the U.S. context. As previously mentioned, China's livestock insurance is heavily subsidized in premiums from central (40%), provincial (25%), and local (15%) governments, covering up to 80% for the main species animals. Some local governments' budgets are limited and sometimes cannot afford this portion of their subsidy responsibility; therefore, they might put a quota to the insurance companies on how many insurance policies they can sell to farmers and feeding operations. This situation is recognized by the central government, and they have been in discussions about re-allocating the subsidy shares among different levels or improve subsidy structure to become more cost-efficient and affordable to farmers as well as the local government budgets. The uptake rate of livestock insurance is still quite low at around 40%, compared to crop insurance at around 70%. The government has been striving to increase the scale of the livestock insurance uptake to benefit more livestock farmers. Thus, subsidy optimization is quite an important factor that we hope to provide implications through our study. We used 5 levels for crop insurance premium subsidies ranging from 50% to 90% incrementally.

Our design also contains an attribute that includes various coverage types, varying from a more traditional coverage type like mortality (Zhang et al., 2016) and to more modern coverage types including weather index, revenue, and feeding profit. We see some recent piloting projects that tried to popularize more innovatively designed insurance products, "imported ideas" from the western world to China. These newer types of insurance products cover different subjects than the traditional and common output quantity coverage. For

instance, China piloted weather index insurance and revenue insurance in recent years (Long, 2021). Weather index insurance, which triggers payment when pre-agreed upon weather indicators hit a certain threshold, is theoretically designed to be more objective and cost-efficient. Because it reduces the need and potential conflicts in assessing the damage (Chantarat et al, 2007). However, it is discovered as inefficient in practical operation because of basis risk and lack of accessible weather and production data, at least in China[33]. The coordination of gathering weather data, running models, educating farmers, and tackling conflicts when there's basis risk is actually higher than assessing damage in the traditional way[34]. The inefficiency may be also from supplier scalability difficulties with dealing with small-size farms. Revenue insurance protects a certain level of income for farmers (Hennessy et al, 1997; Hart et al, 2001), also one newer type of insurance recently piloted for China[35], though faces a more welcoming view by both agents and farmers. Also adding profit or output/input value ratio insurance, we consider these newer and more traditional types of insurance as different levels of an attribute, to understand the suppliers' tradeoff evaluation.

The last attribute is the loan amount that can be granted by banks to insured farmers based on the protected asset value. This insurance-loan-linked service has been on the edge of research and the concept has been emerging in China and some developing countries (Shee et al., 2021; Carter et al., 2011; Giné and Yang, 2009). The China Banking and Insurance Regulatory Commission issued "*Opinions on Strengthening the Cooperation between Agriculture-related Credit and Agriculture-related Insurance"* in 2010, to nurture such collaboration and product development. More specifically, promoting banks to offer favorable loan rates and faster loan applications to farmers with agricultural insurance, while at the

[33] Majority of the weather data are confidential or proprietary to China Meteorological Administration and its local stations. China Ministry of Agriculture and the Agricultural Insurance companies are accessing only part of the data, which cannot make the model precise, and the pricing be accurate. Besides the limited data, quantifying the relationship between production and weather condition requires an extra layer of professional modeling, that adds administrative personnel costs and also in-the-field education for farmers.

[34] We had no previous official data nor studies supporting this argument, though we hear from a lot of agents sharing their thoughts around this angle, and our analysis did reveal that insurance agents have much less willingness to offer weather index insurance than any other type of insurance products.

[35] Traditionally, livestock insurance in the United States is revenue based, while in China it is mortality based.

same time asking the insurance companies to support the collateralization of its contracts for loans. We are aware the concept is still to be publicized, and so it is important to obtain a greater understanding of how insurance providers value the addition of linking credit to their insurance products. Other attributes that are not extensively discussed 36 are briefly described in Table 4.1.

We organized the choices with different levels of attributes under the block design and presented them using cards with images below notations to make the attributes easier to understand. The two images in Figure 4.1 represent the first 2 cards of the 15-card sequence in block #1 shown as an illustration. Each agent was randomly assigned to 1 of the 6 blocks. Then the cards were always presented in order from card 1 to card 15. The detailed description of each attribute is provided in the preamble, written, and read to the agents at the very beginning of the experiment. We also explain to them how the card works: at most 4 attributes are different in one card, each row is an attribute, the levels of the attributes are changing from one another, one column is considered as one full product and there are two products per card to choose from. We ask them to consider comprehensively one complete product with all attributes rather than focus on one particular attribute. We also suggest they think carefully with the tradeoffs as if they are designing or selling the products from the standpoint of a supplier rather than a farmer.

Following the DCE, we ask the agents to complete a survey that captures demographic information, as well as knowledge and perspectives about the livestock insurance for further validations. The survey consists of four parts, including basic demographic information (part A), risk perceptions (part B), agricultural insurance offering experience (part C), as well as knowledge and hope of ideal attribute levels of livestock insurance products (part D). We use this survey information to understand the background, generate interaction terms in evaluating the nonlinear relationships, and back up the DCE data with simply asked questions as a robustness check. From our experience, the experiment took around 35-40 minutes for

[36] The preamble with detailed attributes explanations and term definitions, as well as complete choice experiment sheets is provided in the Appendix.

each agent: 5 minutes explaining the preamble, background, and experiment setup, 15 minutes on choice cards (section 1), and 15 minutes on survey questions. We also have two quick pre-experiment questions, in the beginning, to see what animals (large, medium, small) they primarily insure, and what types of clients (small farms, large farms, and corporates) they mostly work with. We asked the agents to consider the types of animals raised by their clients throughout the whole experiment, as the attribute levels may vary for different animal or client types.

Block 1 Card 1	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	20%	
Risk Environment	Low Risk	
Market Disaster Payment	Included	Not included
Premium Actuarial + Loading (before subsidy)	1% protected value	12% protected value
Premium Subsidy (by gov and org)	80% (farmers pay 20%)	
Coverage Types	Index IBLI	Profit or Ratio
Uninsured Catastrophic and (e.g., war, terrorism, etc.) **Property Casualty Coverage** (e.g., theft, fire, etc.)	Included	
Loan Allowance	Can be allowed loan for **50%** of protected value.	

Block 1 Card 2	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	30%	15%
Risk Environment	Moderate Risk	Low Risk
Market Disaster Payment	Included	
Premium Actuarial + Loading (before subsidy)	25% of protected value	
Premium Subsidy (by gov and org)	80% (farmers pay 20%)	50% (farmers pay 50%)
Coverage Types	Mortality	Index IBLI
Uninsured Catastrophic and (e.g., war, terrorism, etc.) **Property Casualty Coverage** (e.g., theft, fire, etc.)	Not included	
Loan Allowance	Can **Not** be allowed for a loan (0%)	

Figure 4.1: Choice cards in block design

4.2.3. Conditional Logit and Mixed Logit Regression Models

Both Conditional and Mixed logit regression models are applied to analyze the experimental data in our study, following the literature of DCE. A random utility framework and rational choices assumption is generally applied in the model. Respondents make choices to maximize benefit or utility when faced with a discrete set of choices. We use the selected choices by the

agent's choice on each of the 15 cards (=1 means selected, =0 means not selected). For example, if they choose A between A and B, then row A in the database will be coded as 1, and row B will be coded as 0. Homogeneous preferences among participants are the underlying assumption in the conditional Logit model, indicates that all agents have the same WTO measures. Alternatively, heterogeneous preferences are assumed for agents in the mixed Logit model, which means agents have mixed WTO measures. As we interviewed 211 agents and each agent answering to 15 cards, we have 3,165 choice sets. Each choice set includes two choices; thus, we have N=6,330 observations in our sample. We follow the conventional conduct of the logit models explained briefly below, and we refer the greater econometric details to McFadden (1974), Train (2009), McFadden et al (1977) for interested readers. We can also use the point estimates in these logit models to derive willingness to offer (Train 2009), which picks one price factor and illustrates preferences on other attributes more directly.

4.2.3.1. Conditional Logit Model

The DCE is analyzed under the framework of random utility theory (McFadden 1974). We denote that when each insurance agent respondent i faces J alternative insurance products they obtain utility U_{ij}. Random utility maximization suggests that respondent i will choose choice j if and only if

$$U_{ij} > U_{ik}, \forall\, j \neq k \tag{4.1}$$

The utility of an individual can be decomposed into 2 parts with the underlying utility function stated in equation (4.2). The first part is the supplier's indirect utility V_{ij} and is observed by the researcher through predefined specific attributes X'_{ij} in the insurance products and estimated preference estimators β. With the conditional Logit model, it is assumed that all respondents have similar preferences; thus the β is the same across all agents and is not assigned agent sub-notation. The second part of the utility function denotes the unobserved utility from other sources that our data do not capture. These unobserved characteristics are denoted by ε_{ij} and are treated as a random error in the logit models.

$$U_{ij} = V_{ij} + \varepsilon_{ij} = X'_{ij}\beta + \varepsilon_{ij} \tag{4.2}$$

In Eq 2 X'_{ij} is a vector of attributes for the j th alternative, and β is a vector of preference parameters. It is assumed that the random error ε_{ij} follows independently, identically distribution, and a Gumbel (extreme value type I) distribution. The probability that the individual i will choose alternative j among J choices can be derived as a closed-form expression (McFadden, 1974), stated below:

$$P_{ij} = \frac{exp(X'_{ij}\beta)}{\sum_{k=1}^{k=J} exp(X'_{ik}\beta)} \tag{4.3}$$

As the conditional logit model is estimated using maximum likelihood, we can further derive the estimator:

$$\hat{\beta}_{MLE} = \underset{\beta}{\operatorname{argmax}} \sum_{i=1}^{i=N} \sum_{j=1}^{j=2} y_{ij}\, ln \frac{exp(X'_{ij}\beta)}{\sum_{k=1}^{k=2} exp(X'_{ik}\beta)} \tag{4.4}$$

In our choice experiment, we have two choices (alternatives) per card, thus the latent utility for individual i to choose insurance bundle j ($j = 1,2$) is:

$$U_{ij} = \beta_1 Claim_{ij} + \beta_2 Risk_{ij} + \beta_3 Disaster_{ij} + \beta_4 Prem_{ij} + \beta_5 Subsidy_{ij} + \beta_6 Type_{ij} + \beta_7 PC_{ij} + \beta_8 Loan_{ij} + \varepsilon_{ij} \ (1.)$$

Thus, the Logit probability for individual i to choose insurance product j ($j = 1,2$) is:

$$Prob_{ij} = \frac{exp(\beta_1 Claim_{ij} + \ldots + \beta_4 Prem_{ij} + \beta_5 Subsidy_{ij} + \beta_6 Type_{ij} + \beta_7 PC_{ij} + \beta_8 Loan_{ij})}{\sum_{k=1}^{k=J} exp(\beta_1 Claim_{ij} + \ldots + \beta_4 Prem_{ij} + \beta_5 Subsidy_{ij} + \beta_6 Type_{ij} + \beta_7 PC_{ij} + \beta_8 Loan_{ij})} \tag{4.5}$$

What we observe are the choice indicators $C_i = j$ ($j = 1,2$) , not the utility value and Logit probabilities that are not directly observable. So we generate an indicator as aforementioned illustratively: if agents choose A between choices A and B, then row A in the database will be coded as 1, and row B will be coded as 0. Therefore, in utility terms we generate 2 binary choice indicators to state the preference:

$$C_{ij} = \begin{cases} 1, & U_{ij} > U_{ik}\ , \forall\, j \neq k \\ 0, & if\ not \end{cases} \tag{4.6}$$

The Conditional Logit model is constructed to evaluate alternative-specific attributes variations where the estimate varies only by attribute, not evaluate individual-specific attributes variations as the Mixed Logit model to be introduced next does.

4.2.3.2. Mixed Logit Model

The mixed Logit model is more flexible and generalized. It can tackle the limitations faced by the conditional Logit model, particularly, in its imposition of proportionate substitution pattern, inability to handle random test variations, and the correction in unobserved factors over time (McFadden and Train 2000). Consequently, the mixed Logit model does not impose the independence of irrelevant alternatives axiom as the conditional Logit does. In addition, the mixed Logit model assumes heterogeneous preference among agents rather than the similar preference among respondents conjectured in the conditional Logit model. In our context, since insurance agents come from different backgrounds and serve various types of clients, their preferences for insurance attributes may also be heterogeneous.

The latent utility function in the Mixed Logit model is similar to the conditional Logit model. One key difference is the varying preference estimator from one agent to another. Changing β in equation (2) to β_i in equation (8), implies a vector of individual-specific coefficients. The true econometric difference actually comes from the relaxation of the IIA assumption. $\bar{\beta}$ is the mean preference of all agents which is a non-individual-specific vector in conditional logit model, where $E\left[X'_{ij}\xi_i\right] = 0$. Then in the mixed logit model ξ_i can be correlated across individuals so that $E\left[X'_{ij}\xi_i\right] \neq 0$. Here, ε_{ij} is still assumed to be distributed IID.

$$\begin{aligned} U_{ij} = V_{ij} + \varepsilon_{ij} &= X'_{ij}\bar{\beta} + X'_{ij}\xi_i + \varepsilon_{ij} \\ &= X'_{ij}\left(\bar{\beta} + \xi_i\right) + \varepsilon_{ij} \\ &= X'_{ij}\beta_i + \varepsilon_{ij} \end{aligned} \tag{4.7}$$

Now with the individual specific preference, we will have the probability of respondent i choosing alternative j be conditioned on realizing β_i. The conditional probability is represented as:

$$Prob_{ij|\beta_i} = \frac{exp\left(X'_{ij}\left(\bar{\beta} + \xi_i\right)\right)}{\sum_{k=1}^{k=J} exp\left(X'_{ij}\left(\bar{\beta} + \xi_i\right)\right)} \tag{4.8}$$

In contrast to Conditional Logit estimation, here we cannot analytically solve the probability due to randomness, and thus we use statistical software to approximate it using simulation methods (Train, 2009). Particularly, STATA is used to obtain the maximum simulated likelihood estimator stated below. Standard deviations around the estimator mean are usually reported, to show the significance of heterogeneity in attribute preferences across respondent agents.

$$\hat{\theta}_{SMLE} = \underset{\theta}{argmax} \sum_{i=1}^{i=N} \sum_{j=1}^{j=2} y_{ij}\, ln \left[\frac{1}{R} \sum_{r=1}^{r=R} \frac{exp\left(X'_{ij}\beta_i{}^r\right)}{\sum_{l=1}^{l=J} exp\left(X'_{ik}\beta_i{}^r\right)} \right] \tag{4.9}$$

4.2.3.3. Willingness to Offer (WTO)

Similar to consumer demand studies that present the Willingness to Pay (WTP) measures, the insurers in this study are asked to provide a Willingness to Offer (WTO). The attribute estimators in the conditional and mixed Logit model can be used to derive the respondents' willingness to offer (WTO). WTO reveals the insurance premium charge for an agent to be willing to offer a product, given a marginal increase or decrease in a given attribute. It is equal to the negative marginal substitution rate between the given attribute and the price attribute. In our model, we treat the premium rate as the price attribute (the insurance premium price, the 4th attribute in Table 4.1). The WTO for the kth attribute as follows:

$$WTO_k = \frac{\partial U_{ij}/\partial X_{ijk}}{\partial U_{ij}/\partial Prem_{ij}} = \frac{\widehat{\beta_k}}{\widehat{\beta_p}},\ k = 1,\ \ldots,\ 8\ for\ k \neq p \tag{4.10}$$

where $\widehat{\beta_k}$ is the kth attribute estimator, and $\widehat{\beta_p}$ is the coefficient for price-related attribute: actuarial premium price with loading fee. These $\widehat{\beta_s}$ are estimated parameters from logit models.

4.3. Data and Results

The data for this study was collected through DCE and a survey from 211 agricultural insurance agents from December 2020 to March 2021. They are from 15 insurance companies and branches at 13 provinces in China where livestock is largely raised, and insurance products are broadly operated. During the search and contact stage, we reached out to over 20 insurance

firms and 320 agents who at least have some experience in livestock insurance. Those who declined to participate in our study either had no time or interest, consider themselves as not having much experience in livestock insurance, or were concerned about malicious market competition. Because insurance agents are earning a good and stable corporate income, they will not be quite incentivized by a small reward (monetary or non-monetary) for participating in our study. And some of the insurance companies actually discourage their employees to participate in external studies without their highest corporate level consent to prevent malicious competition of any sort37. Therefore, broadly inviting agents to participate in our study was not going to obtain any large-scale participation. Thus, our searching mechanism is truly based on the social network, working with China Agricultural University, China Ministry of Agriculture and Rural Affairs, and Xiamen University to reach their network in agricultural insurance companies. We asked these contacts of ours in China to ask for their friends' help and introduce us to more and more of their colleagues and friends who work in the livestock insurance industry. By following the various paths of social networks, we were able to capture a significant cross-section of the livestock insurance industry in China[38].

4.3.1. Demographic Characteristics and Selected Survey Questions Summary

Most insurance agents interviewed were aged in the range of 25-45, in their active career stage with the main contribution to the company's business activities. Approximately 75.8% of the participants were male and 24.2% were female39. Within our sample, about 79.1% primarily work with mid-size animals including pigs (the majority) and sheep; 17.5% insure large animals including cattle and dairy; 2.4% insure small animals like poultry. This pig dominance of our sample was not by design, but was a desirable outcome. Because China is

[37] Even though we specified that we are university research projects that don't use the information for any commercial purposes, some individuals are still very protective of their perspectives. Some even feel more unwilling to participate due to China-U.S. relationship tensions lately, because we are a U.S. academic institution.

[38] Even though there is no official statistics of how many people are working in the livestock insurance area in China, we heard from several agents that we covered almost all the industry, as they noted many frontier conferences in the area of livestock insurance have at most about 300-400 participants.

[39] Even though gender equality has improved a lot in China over the last several decades, the culture is still that women tend to not take jobs that require a lot of travel for business, as they work rather locally for children.

the world's largest pig producer, pig insurance is the most common, mature, and early adapted insurance among all types of animals. The client base of our respondents is quite balanced among different farm sizes, with about 38.4% of farmers primarily working with small feeding farms, 27.5% with large farms, and 29.4% with feeding companies. Educational levels were relatively high and align with the general corporate standard; about 96.2% of our participants have a college degree or higher. The mean responses on "ideal but feasible" levels of primary attributes of livestock insurance products are close to the existing insurance contracts, and hoping for slight changes that benefit the suppliers40. To be specific, the "ideal but feasible" level on claim starting rate is 14.6%, compared to the operating level of 10%. The hoped level is 6.17% for actuarial premium charges, higher than 6% is presently implemented. The responded "ideal but feasible" level for government subsidy towards premium is 76.43%, higher than the current 70%. We did interview a few executives of some insurance companies; however, our focus was on frontline agents who dealt with livestock insurance on a day-to-day basis as shown by the company and position levels in Table 4.2. Given a very young agricultural insurance industry in China, our sample of insurers had an average of 4.865 years of experience working with agricultural insurance and can be considered competent in knowing insurance products and in turn, supports the validity of our results. Respondents' knowledge about Mortality insurance is the greatest, followed by revenue insurance, profit insurance, and Index-Based Livestock Insurance that are rather more recently piloted.

Table 4.2: Some Demographic Characteristics of Insurance Agents Respondents (N = 211)

Survey Question Variables41	Mean	S.D.
Age	35.289	7.172
Education Level 1=High school, 2=College, 3=MS, MBA, 4=PhD	2.090	0.398
Affiliated company level 1=City / county subdivision; 2=Provincial branch; 3=Central national headquarter	1.360	0.671
Position level within the affiliated company, branch, subdivision 1=Entry level, 2=Middle level (Director, Manager), 3=Senior Level (President, VP, CXO)	1.872	0.893

[40] We may call it an anchoring effect, though it's not the focus of this study, maybe for future research.
[41] The detailed and actual survey questions and explanations will be provided upon request.

Agriculture / Farm / Ranch background and systematic training 1=None, 2=Some, 3=Extensive	1.910	0.566
Insurance Working Years (all types: health, home, auto, Agri, etc.)	8.531	5.610
Agricultural Insurance Working Years (Crop, Livestock, Forest, etc.)	4.865	2.876
Freedom or Input to Insurance Product Design 1=No Input (purely listen to gov's design in all); 2=Some Input 3=Extensive (only listen to the main structure of gov advice); 4=Total Freedom	2.877	0.902
Knowledge* about Index-Based Livestock Insurance	1.649	0.633
Knowledge* about Livestock Mortality Insurance	2.867	0.367
Knowledge* about Livestock Gross Revenue Insurance	2.213	0.688
Knowledge* about Feeding Profit, or Animal to Feed Ratio Insurance	1.877	0.752

*1=A little (aware of), 2=Some (trained but not yet offered), 3=Extensive (has offered for some years or researched deeply)

4.3.2. Conditional and Mixed Logit Model Results

The summary of the main specifications for both conditional and mixed logit model results is presented in Table 4.3. In addition to the basic main effects model, we have tested with adding interaction terms on knowledge about different types of insurance. The signs of the main effect model are as expected referring to the attributes in Table 4.1, estimators have positive signs for attributes that are positively related to supplier's utility and vice versa. Estimator for premium rate is significantly positive, as insurance agents want to collect higher prices of their products; more specifically, one unit increase in premium would add 1.123 to agents' utility, making them exp(1.123)=3.074 times more likely to offer the insurance product. The estimator for the risk environment is negative as providers favor lower-risk clients. Claim starting point parameter is not statistically significant, align with our observations during interviews that the regulatory body does not allow a change in claim starting point according to policy. Market disaster payment is similar where it's more ruled by the government as poverty assistance rather than the terms the insurance companies can decide. This can also be supported by our survey question on the freedom of the insurance product design, where insurance companies do not have the full flexibility (flexibility indicator is 2.877 out of 4). The premium subsidy has the most statistically significant and largest magnitude attribute estimator for agents' utility of offering insurance products. Particularly, one unit increase in

subsidy for premiums would add 3.166 to agents' utility, making them exp(3.166)= 23.712 times more likely to offer the insurance product. Looking at the surface, premium subsidy is granted to insurance buyers and is not directly go to sellers. So some people may wonder why do suppliers care about subsidies so much. This comes down to the endogenous structure of price and quantity, where suppliers understand that if the subsidy is low then farmers are not buying thus their revenue will be low. The endogenous structure with premium subsidy helps to boost the publicization of insurance and its uptake. The estimator for coverage of catastrophic events (war, terrorism, earthquake, mudslide, typhoons) and property-casualty (theft and fire) is negative which shows agents are unwilling to include these rare but high-cost events. Some respondents noted it could be an extra side contract they can add as a commercial insurance product, not for the policy-oriented type of insurance product with much broader appeal. The insurance collateralized loan allowance is not statistically significant in our model, it might be due to the inactive implementation of this concept. Anecdotally, we are aware that some insurance agents are aware of this concept from reading government documents, but have not seen it practically implemented broadly. Some agents claimed they have never heard of insurance-loan cooperation, and thus this particular attribute is not widely considered as an important attribute in their overall choice selection considerations.

Table 4.3: Conditional and Mixed Logit Results With/Without Type Knowledge Interaction Terms

	Conditional Logit Models		Mixed Logit Models			
	(1)	(2)	(3)		(4)	
VARIABLES	Base Model	Add Interaction	Base Model		Add Interaction Terms	
			Mean	SD	Mean	SD
Claim Starting Point	-0.0875	-0.0746	0.0519	5.665***	-0.0634	-5.760***
	(0.476)	(0.476)	(0.693)	(1.060)	(0.702)	(1.098)
Risk Environment	-0.348***	-0.347***	-0.451***	0.390***	-0.434***	0.335**
Linear Risk Order: Low Mid High = 1 2 3	(0.0466)	(0.0466)	(0.0654)	(0.125)	(0.0643)	(0.151)
Market Disaster Pay	0.00312	0.00485	0.0112	0.547***	-0.0208	0.519***
	(0.0621)	(0.0622)	(0.0834)	(0.124)	(0.0835)	(0.132)
Premium actuarial no subsidy	1.123**	1.129**	1.637***	-2.666**	1.646***	-2.848***
	(0.446)	(0.447)	(0.568)	(1.243)	(0.583)	(1.029)
Premium subsidy	3.166***	3.154***	4.626***	4.658***	4.738***	5.076***
	(0.254)	(0.255)	(0.471)	(0.542)	(0.495)	(0.589)
Type Index IBLI	-1.027***	-0.526	-1.396***	0.749***	-0.0192	-0.538**

	(0.0910)	(0.628)	(0.132)	(0.168)	(0.827)	(0.210)
Type Revenue or Price	-0.202*	0.539	-0.235	0.692***	1.209	0.315
	(0.106)	(0.661)	(0.145)	(0.192)	(0.876)	(0.268)
Type Margin or Ratio	-0.454***	-0.0927	-0.685***	-0.690***	0.381	-0.572**
	(0.110)	(0.640)	(0.142)	(0.242)	(0.843)	(0.226)
o. Type Mortality (Base Type)	-	-	-	-	-	-
Cata and PC Coverage	-0.208***	-0.211***	-0.263**	0.966***	-0.279***	0.891***
	(0.0679)	(0.0682)	(0.104)	(0.137)	(0.107)	(0.142)
Loan Allowance	0.107	0.115	0.244	-1.017***	0.204	-1.083***
	(0.112)	(0.113)	(0.158)	(0.242)	(0.158)	(0.222)
Type Mortality*Know		0.302			0.599**	-0.290***
Interaction Terms		(0.214)			(0.287)	(0.0694)
Type Index IBLI*Know		0.220*			0.192	-0.116
Interaction Terms		(0.122)			(0.174)	(0.106)
Type Revenue or Price*Know		0.0550			0.0794	0.240**
Interaction Terms		(0.126)			(0.179)	(0.0937)
Type Margin or Ratio*Know		0.265**			0.317**	0.0695
Interaction Terms		(0.117)			(0.160)	(0.103)
Observations	6,328	6,328	6,328	6,328	6,328	6,328

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

The type of insurance coverage attributes is also statistically significant. Setting mortality insurance as the base, we conclude that agents are less willing to offer more advanced (newly introduced) types of livestock insurance. Offering Index-Based Livestock Insurance (IBLI), Revenue Insurance, or Profit Insurance instead of Mortality insurance will lower the agent utility by 1.027, 0.202, and 0.454 respectively. In other words, mortality insurance is the most preferred type by suppliers, followed by revenue insurance, profit insurance, and then IBLI the least. This is not surprising since mortality insurance is very popular in China. In our survey, we asked agents about their knowledge and experience of these different types of insurance. As discussed briefly in Table 4.2 summary statistics, we observe that greater knowledge of a specific insurance type suggests a greater willingness to offer it. To test this formally we ran an interaction model, with interactions between insurance type and knowledge about the insurance type, (e.g., the insurance coverage subject type = Mortality interacts with Knowledge about Mortality). Results show that the estimators in the interaction model are still statistically significant with expected signs, similar to the significance and signs of the main effect model except for the insurance type attribute. The

interaction term between type and knowledge is positively significant, meaning that higher knowledge about the insurance type would increase the agent's willingness to offer such type.

Overall, we find general consistency between the patterns of the estimated results in conditional and mixed logit models, though we find the estimators are generally more extreme in mixed logit than in conditional logit. This indicates preference heterogeneity among agents in the supply for livestock insurance in China. To further verify the consistency of our model and show robustness, we split the sample by different regions and re-ran both conditional and mixed logit models. The results are presented in Tables 4.4 and 4.5. We placed the 13 provinces into 4 geographic regions according to geographic closeness, administrative district, primary feeding animal, and sample balance. The 4 regions are (1) Northern: Beijing, Hebei, Heilongjiang, Shanxi; (2) Middle Plain: Anhui, Henan, Shandong; (3) Northwest: Gansu, Xinjiang; and (4) Southern: Guizhou, Chongqing, Sichuan, Hubei. We find that most of the parameter estimators are consistent across regions and the total sample. We continue to observe the Mixed logit model results are more extreme than the conditional Logit model, indicating some level of heterogeneous preferences.

Table 4.4: Conditional Logit Results by 4 Regions

Conditional Logit Models	(0)	(1)	(2)	(3)	(4)
VARIABLES	All 4 Regions	Northern	Middle Plain	Northwest	Southern
	Total 211 agents	81 Agents	53 Agents	28 Agents	49 Agents
Claim Starting Point	-0.0875	1.052	-1.644*	1.291	-0.468
	(0.476)	(0.821)	(0.938)	(1.290)	(1.100)
Risk Environment	-0.348***	-0.307***	-0.507***	-0.522***	-0.309***
Linear Risk Order: Low Mid High = 1 2 3	(0.0466)	(0.0777)	(0.105)	(0.153)	(0.0959)
Market Disaster Pay	0.00312	-0.0439	0.0875	0.405**	-0.145
	(0.0621)	(0.105)	(0.123)	(0.205)	(0.133)
Premium actuarial no subsidy	1.123**	0.827	0.755	1.630	2.693**
	(0.446)	(0.709)	(0.805)	(1.246)	(1.332)
Premium subsidy	3.166***	2.980***	2.530***	4.282***	3.663***
	(0.254)	(0.421)	(0.496)	(0.950)	(0.521)
Type Index IBLI	-1.027***	-1.014***	-0.819***	-1.650***	-1.260***
	(0.0910)	(0.140)	(0.176)	(0.325)	(0.212)
Type Revenue or Price	-0.202*	-0.190	0.0197	0.213	-0.422*
	(0.106)	(0.163)	(0.193)	(0.801)	(0.253)
Type Margin or Ratio	-0.454***	-0.632***	0.0593	-1.071***	-0.603**
	(0.110)	(0.176)	(0.250)	(0.324)	(0.241)

o. Type Mortality (Base Type)	-	-	-	-	-
Cata and PC Coverage	-0.208***	-0.221**	-0.0758	0.309	-0.462***
	(0.0679)	(0.110)	(0.140)	(0.237)	(0.157)
Loan Allowance	0.107	0.0538	0.431*	-0.190	-0.00696
	(0.112)	(0.189)	(0.235)	(0.376)	(0.222)
Observations	6,328	2,430	1,590	840	1,468

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.5: Mixed Logit Results by 4 Regions

Mixed Logit Models	(0)	(1)	(2)	(3)	(4)
VARIABLES	All 4 Regions	Northern	Middle Plain	Northwest	Southern
	Total 211 agents	81 Agents	53 Agents	28 Agents	49 Agents
Claim Starting Point	0.0519	1.925	-2.548**	2.161	0.460
	(0.693)	(1.300)	(1.224)	(1.888)	(1.850)
Risk Environment	-0.451***	-0.508***	-0.645***	-0.536**	-0.392***
Linear Risk Order: Low Mid High = 1 2 3	(0.0654)	(0.137)	(0.160)	(0.216)	(0.124)
Market Disaster Pay	0.0112	0.0237	0.0956	0.916**	-0.188
	(0.0834)	(0.164)	(0.160)	(0.403)	(0.204)
Premium actuarial no subsidy	1.637***	1.523	1.158	1.517	3.880
	(0.568)	(1.039)	(1.222)	(1.580)	(2.721)
Premium subsidy	4.626***	4.423***	4.459***	7.295***	6.029***
	(0.471)	(0.763)	(1.115)	(2.072)	(1.115)
Type Index IBLI	-1.396***	-1.358***	-1.191***	-3.717***	-1.992***
	(0.132)	(0.203)	(0.299)	(1.020)	(0.379)
Type Revenue or Price	-0.235	-0.126	-0.0679	1.330	-0.697*
	(0.145)	(0.245)	(0.304)	(1.466)	(0.358)
Type Margin or Ratio	-0.685***	-0.889***	-0.166	-1.991***	-1.135***
	(0.142)	(0.219)	(0.313)	(0.671)	(0.391)
o. Type Mortality (Base Type)	-	-	-	-	-
Cata and PC Coverage	-0.263**	-0.339*	-0.110	0.173	-0.750**
	(0.104)	(0.193)	(0.210)	(0.343)	(0.305)
Loan Allowance	0.244	0.302	0.983**	-0.234	-0.0906
	(0.158)	(0.266)	(0.450)	(0.516)	(0.306)
Observations	6,328	2,430	1,590	840	1,468

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

4.3.3. Odds Ratio

We can also interpret agent preference using the odds ratio, which is derived by estimators from the conditional and mixed Logit model. The Logit model's underlying utility can be expressed using attribute selection probabilities below, showing the log of the odds ratio of selecting (P_{ij}), or not selecting $(1 - P_{ij})$, an attribute j.

$$U_{ij} = LN\left(\frac{P_{ij}}{1 - P_{ij}}\right),\ j = 1,\ \ldots,\ 8 \text{ attributes} \tag{4.11}$$

Subsequently, the marginal utility for a particular attribute is denoted as

$$\frac{\partial U_{ij}}{\partial x_j} = e^{\beta_j} \tag{4.12}$$

It indicates that marginal changes of the utility toward a change of the attributes, are given by the logarithm percentage change in the odds that attribute j will be preferred, with all other things held constant. The Logit estimators reveal utility gain or losses depending on the attribute level changes. One level change in the attribute j would shift agents' utility by β_j. If $\beta_j > 0$, it means the one level increase of attribute level makes agents e^{β_j} % more likely to offer this insurance product. If $\beta_j < 0$, it reads as one level increase of attribute level led agents to be $-\left(1 - e^{\beta_j}\right)$ % less likely to offer this insurance. Table 4.6 presents the odds ratio derived from conditional and mixed Logit model results of Table 4.4 and 4.5, explaining agents' likelihood to offer insurance based on the attribute changes.

Table 4.6: Odds Ratios for Conditional Logit (CL) and Mixed Logit (ML) Models

Odds Ratio	All 4 Regions		Northern		Middle Plain		Northwest		Southern	
VARIABLES	CL	ML	CL	ML	CL	ML	CL	ML	CL	ML
Claim Starting Point	-0.084	1.053	2.863	6.855	-0.807	-0.922	3.636	8.680	-0.374	1.584
Risk Environment	-0.294	-0.363	-0.264	-0.398	-0.398	-0.475	-0.407	-0.415	-0.266	-0.324
Market Disaster Pay	1.003	1.011	-0.043	1.024	1.091	1.100	1.499	2.499	-0.135	-0.171
Premium actuarial	3.074	5.140	2.286	4.586	2.128	3.184	5.104	4.559	14.776	48.424
Premium subsidy	23.712	102.105	19.688	83.346	12.554	86.401	72.385	1472.917	38.978	415.300
Type Index IBLI	-0.642	-0.752	-0.637	-0.743	-0.559	-0.696	-0.808	-0.976	-0.716	-0.864
Type Revenue or Price	-0.183	-0.209	-0.173	-0.118	1.020	-0.066	1.237	3.781	-0.344	-0.502
Type Margin or Ratio	-0.365	-0.496	-0.468	-0.589	1.061	-0.153	-0.657	-0.863	-0.453	-0.679
Cata and PC Coverage	-0.188	-0.231	-0.198	-0.288	-0.073	-0.104	1.362	1.189	-0.370	-0.528
Loan Allowance	1.113	1.276	1.055	1.353	1.539	2.672	-0.173	-0.209	-0.007	-0.087

The derived odds-ratios are of similar scale between the conditional and mixed Logit models, but some differ across various regions, especially for the statistically insignificant

variables. Here we interpret the total effect of all 4 regions in the very left column. Risk environment shows a negative effect, with a one level increase of risk results to decrease in odds of offering insurance by 29.1% in conditional Logit and 36.2% in the mixed Logit model. The actuarial premium charge is favored by the insurance agents, with the odds ratio showing that one level of premium charge increases, agents are ranging 3.074 to 5.14 times more likely to offer insurance. What strikes us the most is the extraordinary impact the subsidy on premium has on agent's willingness to offer, with the odds ratio indicating one level increase of subsidy (10%) will lead to agents 23.71 to 102.11 times more likely to offer insurance. Northwest region has the strongest positive effect from higher subsidy on premium charges, showing a 10% increase of subsidy lead agents to be 72.385 to 1472.917 times more likely to offer insurance. Intuitively speaking with stories shared by the agents but without precise measurement, Northwest is the poorest region among the 4 with low affordability farmers. In addition, Northwest farmers generally are generally less educated and know less about the benefit of insurance. Therefore, agents find it very difficult to sell agricultural insurance in this region and subsidy is particularly more helpful in this region. The base insurance coverage type is the traditional mortality insurance, while what is included in the table are the rather newer experimented insurance types: Index IBLI, Revenue insurance, and Margin insurance are making agents less likely to offer insurance. The Index IBLI insurance is the most disliked, which is 64.2% to 75.2% less likely to be offered compared to the mortality insurance. Catastrophic and property-casualty coverage is not favored by agents, when included, agents are 18.8% to 23.1% less likely to offer insurance. Agents also favor higher loan allowance, as one level (10%) increase in loan allowance leads agents to be 1.113 to 1.276 times more likely to offer insurance; though this estimate is statistically insignificant so it should be interpreted with caution. Similarly, the interpretation for claim starting point and inclusion of market disaster pay is inconsistent among models and regions, thus should be read with less confidence. And we did explain earlier why we think these few attributes are not statistically significant.

4.3.4. Willingness to Offer (WTO)

The Willingness to Offer (WTO) measures are obtained by dividing each non-price-related estimator in Tables 4.4 and 4.5 by the negative of the estimator for the actuarial premium

attribute[42] (refer to Eq. 10). The unit measurement for premium is the percent of the protected value[43]. Here we illustrate mostly based on the total effect of all 4 regions sample and primarily on conditional logit model results shown in Table 4.7.

As discussed earlier the claim starting point for livestock insurance is not flexible in design by an insurance company due to regulation. The estimators are not statistically significant across most of the models and the WTO calculation is also varying a lot from model to model. Similarly, the estimators for attribute Market Disaster Pay and Loan Allowance are statistically insignificant and have reasonable background rationales. And the WTO varies a lot from model to model for these two variables as well.

The interpretation of WTO will be opposite from Willingness to Pay (WTP) in insurance studies that focused on demand, thus we need to be cautious about interpreting intuitively. As shown in the conditional logit results, the agents will want to charge 31% higher in premium if the risk environment increases by one level. In other words, one level increase of risk will lower suppliers' WTO by 31%. It has a very similar magnitude in the mixed logit results and consistent among different regions as well. Including Catastrophic and Property Casualty coverage will make agents want to charge 18.5% more in insurance premium price.

The premium subsidy and insurance type are the most important attributes and what we believe the main contribution of our study. We can see the subsidy WTO indicates that agents are 281.9% more willing to offer insurance by one level increase of subsidy. The base insurance type is mortality, and the WTO results suggest that agents are 91.5%, 18%, and 40.4% less willing to offer IBLI, revenue insurance, and margin insurance compared to mortality insurance respectively. In other words, for example, the agents would want to charge 18% higher in premium for revenue insurance than mortality insurance. The results are consistent

[42] Recall Table 1, the actuarial premium levels are 1%, 3%, 6%,12%, and 25% of the protected value. Therefore, one level increase of the premium charges is approximately doubling.

[43] Take pig insurance as an example, the protected value of one pig is around 1200 RMB and the premium is 6% and thus 72 RMB is the premium. For cattle it would be more than 7000 RMB and the premium rate is around 3% or 210 RMB. Because in this study we consider multiple animals, as some agents offer insurance for a variety of animals, talking in percentages are more widely applicable in choice card design.

with the mixed Logit model across different regions. The WTO results are very similar between conditional Logit and mixed Logit models indicating that the heterogeneous preferences among agents may not be statistically significant.

Table 4.7: Willingness to Offer (WTO) in Conditional Logit (CL) and Mixed Logit (ML)

WTO	All 4 Regions		Northern		Middle Plain		Northwest		Southern	
VARIABLES	CL	ML	CL	ML	CL	ML	CL	ML	CL	ML
Claim Starting Point	-0.078	0.032	1.272	1.264	-2.177	-2.200	0.792	1.425	-0.174	0.119
Risk Environment	-0.310	-0.276	-0.371	-0.334	-0.672	-0.557	-0.320	-0.353	-0.115	-0.101
Market Disaster Pay	0.003	0.007	-0.053	0.016	0.116	0.083	0.248	0.604	-0.054	-0.048
Premium subsidy	2.819	2.826	3.603	2.904	3.351	3.851	2.627	4.809	1.360	1.554
Type Index IBLI	-0.915	-0.853	-1.226	-0.892	-1.085	-1.028	-1.012	-2.450	-0.468	-0.513
Type Revenue or Price	-0.180	-0.144	-0.230	-0.083	0.026	-0.059	0.131	0.877	-0.157	-0.180
Type Margin or Ratio	-0.404	-0.418	-0.764	-0.584	0.079	-0.143	-0.657	-1.312	-0.224	-0.293
Cata and PC Coverage	-0.185	-0.161	-0.267	-0.223	-0.100	-0.095	0.190	0.114	-0.172	-0.193
Loan Allowance	0.095	0.149	0.065	0.198	0.571	0.849	-0.117	-0.154	-0.003	-0.023

4.4. Conclusion

Livestock insurance is relatively new compared to crop insurance in China and the product design, as well as the assisting policy, are still improving. In this paper we study the rarely analyzed supply side of livestock insurance, using a DCE approach to identify insurance agents' Willingness to Offer (WTO) agricultural insurance. Our interest is in determining which attributes insurance professional agents value the most in offering agricultural insurance to farmers. The DCE includes eight main insurance attributes and their various levels of variations to generate choice cards for agents to consider and pick, we reveal their utility and preferences through their thinking on tradeoffs of different attributes and levels of insurance products. The in-the-field experiment data was analyzed using conditional and mixed logit models closely following the literature. We also conducted a survey following the DCE, that was used to supplement the analysis of DCE.

We obtained 211 respondents in our study, from 15 companies and branches in 13 provinces of China. Most of our participants are frontline agents who sell and operate livestock

insurance day to day at the local level. We find most of our agents work with pig insurance and their customer base is rather evenly distributed among different sizes of farms. The model shows great statistical significance, and the estimator signs are as expected. The insurance agents prefer customers to have lower risk, and not including catastrophic and property-casualty coverage. Claim starting point, market disaster payment, and insurance collateralized loan allowance are not statistically significant; probably due to the lack of freedom the insurance agents have on these attributes in their product offering considerations, as these attributes are more ruled by the regulatory body.

Insurance premium subsidy as the key policy to support agricultural development and poverty reduction is one important attribute we highlight, especially its role in the endogenous structure between quantity and price. We observe the limited offering of insurance product types and see the challenges in piloting some more advanced insurance types. Therefore, the preference of different insurance coverage subject types is also our focus. We find significant and positive effects of the subsidy on WTO with a one level increase of subsidy leading to 3.166 times higher probability to offer. The model also reveals that mortality insurance is the most preferred by insurance agents due to their familiarity with it and the mature development in the product offering. The newer introduced revenue insurance and feeding profit insurance are a bit less preferred, while the Index-Based Livestock Insurance (IBLI) is a lot less preferred to offer by the insurance agents. Through interaction term models we find that a higher knowledge level of the insurance types will raise their willingness to offer such a product.

This study thus suggests raising subsidies for livestock insurance as the budget allows. We also recommend the policymakers be aware of the agents' unwillingness to offer newer types of livestock insurance and to conduct more education and research for implementing these products. Methodologically, our study contributes to advancing the use of DCE to evaluate supply-side WTO, to complement demand-side applications in the existing literature. Future work in evaluating both WTO and WTP will be helpful in evaluating the equilibrium of insurance premium price.

4.5. Reference

Adamowicz, W., P. Boxall, M. Williams, and J. Louviere. 1998. Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics* 80(1): 64-75.

Belasco, E., Cheng, Y., & Schroeder, T. (2015). "The Impact of Extreme Weather on Cattle Feeding Profits". *Journal of Agricultural and Resource Economics,* 40(2), 285-305. http://www.jstor.org/stable/44131862

Belasco, E., Cooper, J., Smith, V. H., (2019), "The Development of a Weather-based Crop Disaster Program", *American Journal of Agricultural Economics,* 102: 240-258. https://doi.org/10.1093/ajae/aaz021

Belasco, E., Ghosh, S. K., Goodwin B. K., (2009) "A Multivariate Evaluation of Ex ante Risks Associated with Fed Cattle Production", *American Journal of Agricultural Economics,* 91(2), 431-443. http://www.jstor.org/stable/20492443

Cai, Jing. 2016. "The Impact of Insurance Provision on Household Production and Financial Decisions." *American Economic Journal: Economic Policy*, 8 (2): 44-88.

Cai, Jing, Alain de Janvry, and Elisabeth Sadoulet. 2020. "Subsidy Policies and Insurance Demand." *American Economic Review*, 110 (8): 2422-53. DOI: 10.1257/aer.20190661

Cai, Q., Ding, Y., Tuvey, C. and Zhang, Y., (2021). "The influence of past experience on farmers' preferences for hog insurance products: a natural experiment and choice experiment in China". The Geneva Papers on Risk and Insurance - Issues and Practice. 46, 399–421. https://doi.org/10.1057/s41288-021-00209-7

Carter, M., Cheng, L., & Sarris, A. (2011). The Impact of Interlinked Index Insurance and

Credit Contracts On Financial Market Deepening and Small Farm Productivity. *in Organized Symposium Paper, Agricultural and Applied Economics Association Annual Meeting*, Pittsburgh, PA, 2011

Carlsson, Fredrik and Martinsson, Peter, (2001), "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments?: Application to the Valuation of the Environment", *Journal of Environmental Economics and Management*, 41, issue 2, p. 179-192, https://EconPapers.repec.org/RePEc:eee:jeeman:v:41:y:2001:i:2:p:179-192

Chantarat, S., Barrett, C., Mude, A., & Turvey, C. (2007). "Using Weather Index Insurance to Improve Drought Response for Famine Prevention". *American Journal of Agricultural Economics,* 89(5), 1262-1268. http://www.jstor.org/stable/30139471

Chantarat, S., Mude, A.G., Barrett, C.B. and Carter, M.R. (2013), "Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya". *Journal of Risk and Insurance,* 80: 205-237. https://doi.org/10.1111/j.1539-6975.2012.01463.x

China Government, 2019-#102. "Guiding Opinions on Accelerating the High-Quality Development Of Agricultural Insurance", Ministry of Finance Website. http://www.gov.cn/xinwen/2019-10/12/content_5438771.htm Released on 2019-10-12, Accessed on 2021-7-7.

Coble, K., & Barnett, B. (2013). Why Do We Subsidize Crop Insurance? *American Journal of Agricultural Economics*, 95(2), 498-504. *http://www.jstor.org/stable/23358423*

Cole, S., Giné, X., Vickery, J., (2017). How Does Risk Management Influence Production Decisions? Evidence from a Field Experiment, The Review of Financial Studies, Volume 30, Issue 6, June 2017, Pages 1935–1970, https://doi.org/10.1093/rfs/hhw080

Global Facility for Disaster Reduction and Recovery, 2020. Web. June 29, 2021. https://www.gfdrr.org/en/feature-story/fostering-rural-resilience-closer-look-chinas-agricultural-insurance-pilot

Giné, X. and Yang, D. “Insurance, credit, and technology adoption: Field experimental evidence from Malawi”, *Journal of Development Economics*, Vol. 89, (2009) pp. 1–11.

Hart, C.E., Babcock, B.A. and Hayes, D.J. (2001), Livestock Revenue Insurance. *J. Fut. Mark.*, 21: 553-580. https://doi.org/10.1002/fut.1603

Hennessy, D.A., Babcock, B.A. and Hayes, D.J. (1997), Budgetary and Producer Welfare Effects of Revenue Insurance. *American Journal of Agricultural Economics*, 79: 1024-1034. https://doi.org/10.2307/1244441

Jensen, N.D., Barrett, C.B. and Mude, A.G. (2016), Index Insurance Quality and Basis Risk: Evidence from Northern Kenya. *American Journal of Agricultural Economics*, 98: 1450-1469. https://doi.org/10.1093/ajae/aaw046

Johnson, R. and B. Orme. 2003. “Getting the most from CBC”, Sawtooth Software Research Paper Series. *Sawtooth Software, Sequim*.

Kong, R., Turvey, C.G., He, G., Ma, J. and Meagher, P. (2011), "Factors influencing Shaanxi and Gansu farmers' willingness to purchase weather insurance", *China Agricultural Economic Review,* Vol. 3 No. 4, pp. 423-440. https://doi.org/10.1108/17561371111192293

Lancaster K. (1966), “A new approach to consumer theory”, *The Journal of Political Economy*, Vol. 74, pp. 132–157. http://www.jstor.org/stable/1828835

Long, Wenjun (2021). Agricultural Insurance Trend Judgment, Overall Thinking and Guarantee Mechanism in the Period of the "14th Five-Year Plan". *China Insurance Magazine*

Lusk, J.L. and Schroeder, T.C. (2004), “Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks”. *American Journal of Agricultural Economics*, 86: 467-482. https://doi.org/10.1111/j.0092-5853.2004.00592.x

Mahul, Olivier; Skees, Jerry. 2007. “Managing Agricultural Risk at the Country Level : The

Case of Index-Based Livestock Insurance in Mongolia". *Policy Research Working Paper*; No. 4325. World Bank, Washington, DC.

McFadden, D. (1974), "Conditional logit analysis of qualitative choice behavior", in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York, pp. 105– 142.

McFadden, D., W. B. Tye, and K. Train. (1977). "An application of diagnostic tests for the independence from irrelevant alternatives property of the multinomial logit model". *Institute of Transportation Studies*, University of California: 39-45.

McFadden, D. and Train, K. (2000). "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, Vol. 15, No.5, pp. 447-470.

McFadden, Daniel. (2001). "Economic Choices." *American Economic Review*, 91 (3): 351-378.

Miranda, M.J. and Farrin, K. (2012), Index Insurance for Developing Countries. *Applied Economic Perspectives and Policy*, 34: 391-427. *https://doi.org/10.1093/aepp/pps031*

Orme, B. 1998. "Sample size issues for conjoint analysis studies". *Sawtooth Software Technical Paper, Sequim*. *https://sawtoothsoftware.com/resources/technical-papers*

Rose, J.M. and M.C.J. Bliemer. 2013. "Sample size requirements for stated choice experiments". *Transportation* 40: 1021-1041. https://doi.org/10.1007/s11116-013-9451-z

Shee, Turvey, and Marr. (2021) "Heterogeneous Demand and Supply for an Insurance-Linked Credit Product in Kenya: A Stated Choice Experiment Approach", *Journal of Agricultural Economics*, 72: 244-267. https://doi.org/10.1111/1477-9552.12401

Skees, J. R. and Barnett B. J. (2006) "Enhancing Microfinance Using Index-based Risk-Transfer Products". *Agricultural Finance Review,* Vol. 66 No. 2, pp. 235-250. https://doi.org/10.1108/00214660680001189

Skelton, C. and Turvey, C.G. (1994), "On the Inclusion of Hay in Ontario's Gross Revenue

Insurance plan: A Target Semi-Variance Approach". *Applied Economic Perspectives and Policy,* 16: 321-331. https://doi.org/10.2307/1349472

Turvey, C. (2001). Weather Derivatives for Specific Event Risks in Agriculture. *Review of Agricultural Economics*, 23(2), 333-351. http://www.jstor.org/stable/1349952

Turvey, C. G., A. Shee, and A. Marr, (2019), "Addressing Fractional Dimensionality in the Application of Weather Index Insurance and Climate Risk Financing in Agricultural Development: A Dynamic Triggering Approach". *Weather, Climate, and Society,* 11(4), 901-915. https://journals.ametsoc.org/view/journals/wcas/11/4/wcas-d-19-0014_1.xml

Turvey, C.G. (1992), "An Economic Analysis of Alternative Farm Revenue Insurance Policies". *Canadian Journal of Agricultural Economics,* 40: 403-426. https://doi.org/10.1111/j.1744-7976.1992.tb03704.x

Turvey, Calum G. & Kong, Rong & Belltawn, Burgen, (2009). "Weather Risk and the Viability of Weather Insurance In Western China," 2009 Annual Meeting, July 26-28, 2009, Milwaukee, Wisconsin 49362, *Agricultural and Applied Economics Association*

Train, K. E. 2009. Discrete choice methods with simulation (2nd ed.). *Cambridge University Press.*

Wang, H. H., Liu, L., Ortega, D. L.., Jiang, Y., and Zheng Q., 2020. "Are smallholder farmers willing to pay for different types of crop insurance? An application of labelled choice experiments to Chinese corn growers," The Geneva Papers on Risk and Insurance - Issues and Practice, Palgrave Macmillan; The Geneva Association, vol. 45(1), pages 86-110, January. https://doi.org/10.1057/s41288-019-00153-7

Woodard, J.D. and Garcia, P. (2008), "Basis risk and weather hedging effectiveness", *Agricultural Finance Review*, Vol. 68 No. 1, pp. 99-117. https://doi.org/10.1108/00214660880001221

Woodard, J.D., and Yi, J. (2020), “Estimation of Insurance Deductible Demand Under Endogenous Premium Rates”. *Journal of Risk and Insurance,* 87: 477-500. https://doi.org/10.1111/jori.12260

Ye, T., Li, Y., Gao, Y., Wang, J., Yi, M., (2017) “Designing index-based livestock insurance for managing snow disaster risk in Eastern Inner Mongolia, China”, *International Journal of Disaster Risk Reduction*, 23, 160-168.

Yu, J. and Sumner, D.A. (2018), “Effects of subsidized crop insurance on crop choices”. *Agricultural Economics*, 49: 533-545. https://doi.org/10.1111/agec.12434

Zhang, Y., Zhu, X. & Turvey, C. (2016) On the Impact of Agricultural Livestock Microinsurance on Death-Loss, Production and Vaccine Use: Observations from a Quasi-Natural Experiment in China. *The Geneva Papers on Risk and Insurance - Issues and Practice.* 41, 225–243. https://doi.org/10.1057/gpp.2016.3

5. CHAPTER 5

Conclusion – Investigating Emerging Issues in Agricultural Finance

The above chapters investigate the issues and advancement of three different emerging aspects of agricultural finance: financial technology, farmland markets, and livestock insurance products. The first chapter discusses the importance of being aware of these areas of agricultural finance strategies to support the development of both developed and developing agricultural economies. We provide background domain knowledge on Agricultural FinTech with an emphasis to blockchain and cryptocurrencies, elaborating the applications of blockchain in food and agriculture value chain and supply chain, that improves the transparency and efficiency of tracing and trading. By qualitatively evaluating the benefits and challenges of blockchain based Agricultural FinTech solutions, we propose a blockchain token system that is linked to warehouse receipt and enables the holding and trading grains electronically. This solution will primarily be useful to reduce the supply shock upon harvest that usually result low prices during that time and thus less profits of the farmers, and we can help facilitate the storage, insurance and other risk management tools through this blockchain token system. Understanding the relationship between blockchain and cryptocurrencies is crucial to this type of application, thus we relay quantitative research to the more practically operated cryptocurrencies to understand its market property.

Chapter 2 examines the financial property of 3,351 cryptocurrency price series, particularly looking at the long or short memory within its price series and evaluate if they follow a fractional Brownian motion through Hurst estimators and Stepwise autoregression. We investigate thoroughly the releasing mechanism of Bitcoin, whose supply is released in a decreasing rate and has a capped total. We thus use top 105 coins' supply structure to explain the memory in their price series, for which we find the % of total coins issued is explanatory. Due to concerns of built-in memory within the price series, we propose a de-mean method to decompose the price series with the time varying mean and the variations, and we found some of the previously found long memory was spurious. The gradually increasing (time-varying) mean can explain this spuriousness, and we can use the market structure such as the 4-year

bull circle and the deflationary releasing mechanism to explain this time-varying phenomenon. The findings of this chapter confirm the fractionality of crypto markets, and provided clarity of true or spurious long memory using both traditional and proposed method, which enhance our understanding to this emerging financial assets that could potentially be applied to agricultural value chain. Chapter 2's quantitative research suggests the inefficiency in cryptocurrency market primarily come from the supply structure, and thus we believe it is important to understand the tokenomics of digital assets when applying token system in any settings, not only in agriculture. Network effects and sentiment hype could also be the factors of market performance of cryptocurrencies that we leave for future studies to address. We hope as our collaborative efforts in blockchain token systems of warehouse receipt develop and establish participants, we will eventually have data for us or future scholars to empirically analyze its value chain more besides purely the financial aspect of it that we have conducted here. The cryptocurrency market has been evolving fast and continues to disrupt in all kinds of businesses, our findings here maybe only subject to the current 10-year history that we study. It is not our intention to encourage any illegal use or irrational investment of cryptocurrencies, nor we accuse any previous studies that conclude different directions than we do. We hope to remain an open and objective mind to study new emerging technologies and their potential benefits / challenges in practical adoption. We hold optimism for more mature, efficient, and legitimate development in the crypto space and its applications in various industries. Particularly, we call for agricultural economists' attention to this line of potentially disruptive agricultural financial technology that may contribute to the development of digital agriculture and agriculture finance in general.

The third chapter looks at Amish population growth and how that could affect farmland prices when they settle in New York state and buy land to start their family farms. Amish people have unique religious belief which encourages them to have strong family ties and human-nature interactions, thus they generally use less modern technology including those applied to agricultural production. We rationalize the farming style and living habits of Amish farmers and compare that with conventional farmers first in a conceptual framework,

hoping to explain how the less productive Amish farmers can stay competitive and even keep growing and prospering in various states. The main idea is that the cheaper labor costs of Amish due to their larger family members can compensate the disadvantage from less output due to their limited use of farming technologies. Simply speaking, Amish may have lower revenue, though their costs are low as well, so they could maintain the similar level of profits to their non-Amish neighbors. Because farmland is a crucial input for farming, we use transaction records and population trends to empirically to study the impact of Amish population growth on farmland prices. We use a standard hedonic approach and unique shift-share like instrumental variable in this empirical model and find there's no statistically significant relationship between Amish population growth and farmland prices. Therefore, we infer that this may confirm the conceptual model that Amish compete on the farmland markets similar as the conventional farmers. This chapter does not only offer insights to living and farming styles of Amish farmers that result a coexistence of multiple farming systems in one market, represented in the farmland market; but also methodologically showcases how an innovative identification strategy from the labor literature can be applied to agriculture finance issues.

The fourth chapter presents our work in investigating agricultural insurance agents' willingness to offer (WTO) livestock insurance in China, through an in-the-field discrete choice experiment (DCE). We primarily focus on livestock insurance as we see it is less developed than crop insurance and particularly got attention due to animal disease outbreaks in the recent years. We include eight main attributes of livestock insurance and contain various combinations of different levels of them on the choice cards. We implemented the analysis in 6 blocks, with 35 insurance agents in each block. We ask each of them to conduct the choice experiment on 15 cards with each card has two choices of hypothetically designed livestock insurance. The card choice combinations of various attributes and levels are designed using JMP software through a D-optimal approach, which use limited sets of choices to maximize the choice exposures to participants and reveal their utility changes when they decide the trade-offs between two choices on the cards based on the different attribute levels included.

Premium subsidy and insurance types are two out of eight attributes we primarily study and also find strong clear evidence on. We find that one level (10%) increase of subsidy lead agents' probability to offer be 3.166 times higher. We also find the more traditional type, mortality insurance is still strongly preferred than the newer introduced insurance, with index insurance being the least preferred, because of farmers' difficulty in understanding and conflicts when basis risk occurs. Through using survey question to generate interaction term model, we also find knowledge to the newer type of the insurance products improves the WTO on that particular insurance type. This chapter is among the very first to study the supply side of agriculture insurance, and the DCE method we use is a recently popularized method in evaluating insurance products and participants preferences with designs of attributes variations. Our research provides important policy implications as the government and insurance companies work out the details in enlarging the take-up of insurance via interventions of subsidy, education, and innovative /diverse product offering.

The key message I wish to convey through this series of essays is that many areas of emerging questions in agriculture finance deserve attention to explore, understand, and develop further. We picked the areas in Agri FinTech, Farmland Markets, and livestock insurance supply in China due to prior knowledge background that fits into the progressing issues that we observe that are mostly interesting to us. Our research does not only provide findings and policy implications in the related specific context, but also contribute to the methodological advancement in addressing issues and research questions in agricultural finance. As these and other topics in emerging issues of agriculture finance continues to evolve, we hope to see more research to be developed that verifies the benefits of innovative technologies and products, and evaluate the disruptions toward the traditional agricultural system, as well as quantitative investigate the progresses to support the future development of agricultural economy.

APPENDIX

Appendix For Chapter 3

Table 3.A1: Descriptive Statistics for Variables Used in Period 2007-2015 (N=6806)

Variable	Mean	Std. Dev.	Min	Max
Deflated per acre farmland price	4314.251	6430.82	103.516	63607.199
Log deflated per acre farmland price	7.834	.969	4.64	11.06
Amish trend (church districts)	1.078	2.158	0	16.824
Number of parcels in sale	1.398	3.31	0	266
Soil characteristics	5.036	1.555	0	10
Distance to large farms	15380.294	18065.697	0	139452.17
Large farms within 10 miles radius	1.333	2.592	0	18
Distance to New York City(ft)	337480.34	94466.349	61957.836	510181.34
Distance to metropolitan statistical area	7849.125	6326.017	0	41686.566
Distance to town of at least 2500 people	11815.957	6814.129	292.482	60426.629
Distance to the nearest hospital(ft)	14540.615	6694.276	429.322	43542.813
Distance to the nearest college(ft)	21356.021	10849.984	1128.772	63897.008
Distance to a school	5373.231	2667.798	200.176	17104.004
Distance to a park	10273.327	6584.915	0	37614.531
Distance to the nearest golf course	8284.711	6530.503	229.477	54367.656
Distance to EPA site	5750.891	3681.987	93.359	28131.814
Distance to boat launch (recreation)	13946.397	8226.007	321.046	42538.699
Distance to ethanol plant	108128.9	66615.058	2668.708	303815.09
Distance to rail terminal	6990.207	5559.318	1.685	35796.227
Distance to an oilwell	56160.678	64633.722	34.047	285995.72
Distance to an inactive oil well	62392.07	66387.124	122.605	263753.91
Distance to highway on-ramp	13852.355	9599.655	180.89	60081.754
Slope of parcel	9.663	77.804	.002	1000
Tree cover on parcel	23.856	20.423	0	88.968
Available Water Storage (mm)	163.295	78.852	0	868.037
Thickness(cm) in soil organic carbon calc	139.561	34.725	0	331.197
Soil organic carbon (g C per m2)	14179.074	13448.103	0	170327
National commodity crop productivity index	.445	.147	0	.901
Percent earthy major components	80.467	5.823	0	95
Maximum rooting depths	110.861	36.79	0	150
Water storage within crop root zone depths	154.036	69.875	0	675
Droughty Soil Landscapes	.538	.377	0	1
Potential wetland soil landscapes	33.715	70.542	0	999

Table 3.A2: OLS Quantile regression results

2007-2015	OLS Reg	Quantile regression		
Variable	Mean	At 0.25 quantile	At 0.50 quantile	At 0.75 quantile
Amish trend	-0.0192***	-0.0195***	-0.0143***	-0.0182***
	(0.0055)	(0.0058)	(0.005)	(0.0041)
Constant	9.5100***	8.5866***	9.6251***	10.2486***
	(0.2433)	(0.2019)	(0.2586)	(0.2671)
Observations	6,806	6,806	6,806	6,806

Note: none of the 0.25, 0.50, and 0.75 coefficient is statistically significant from OLS.

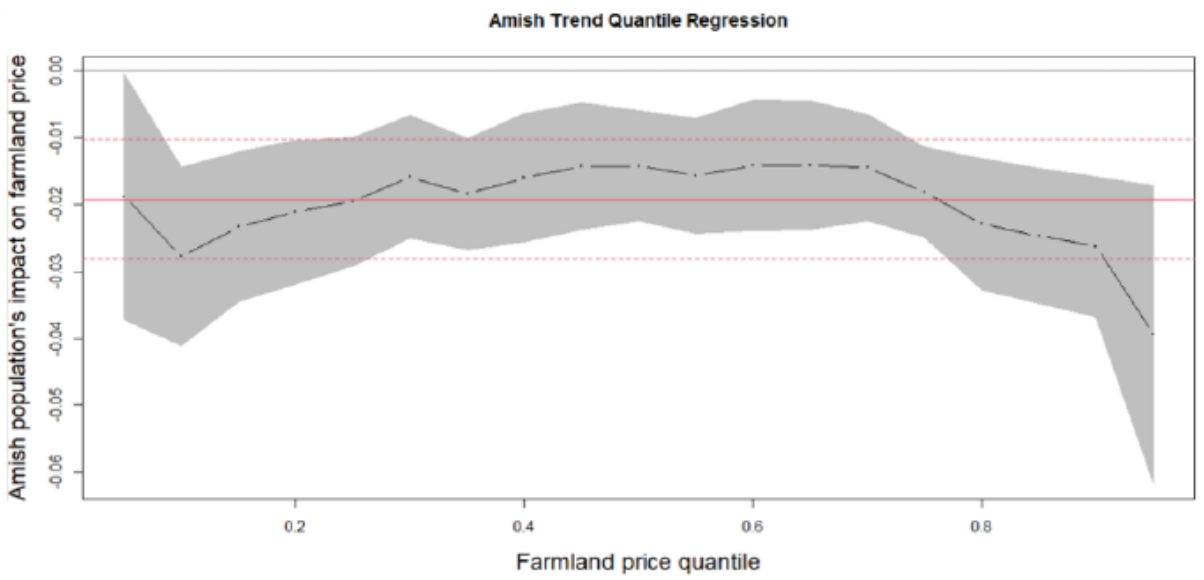


Figure 3.A1: OLS Quantile Regression Analysis: Amish Trend Impact on Farmland Prices

Table 3.A3: Full results of specifications with all control variables

2007-2015	OLS		Enclave IV	
Variable	Standard	IV	+CAFOs	Drop<10 ac.
Amish trend (church districts)	-0.0129	-0.0418	-0.0545	-0.0491
	(0.0126)	(0.0431)	(0.0437)	(0.0375)
Number of parcels in sale	0.00737**	0.00722**	0.00718***	0.00752***
	(0.00298)	(0.00280)	(0.00272)	(0.00289)
Soil characteristics	-0.0181	-0.0163	-0.0161	-0.0157
	(0.0133)	(0.0139)	(0.0137)	(0.0139)
Distance to New York City(ft)	-2.95e-06***	-2.87e-06***	-2.88e-06***	-2.87e-06***
	(1.07e-06)	(1.05e-06)	(1.03e-06)	(1.05e-06)
Distance to metropolitan stats area	-3.23e-06	-1.12e-06	-1.25e-06	-1.31e-06
	(4.75e-06)	(6.52e-06)	(6.48e-06)	(6.57e-06)
Distance to town of ≥ 2500 people	-2.01e-06	-2.71e-06	-2.57e-06	-2.60e-06
	(4.11e-06)	(4.33e-06)	(4.37e-06)	(4.33e-06)
Distance to the nearest hospital(ft)	1.60e-06	2.98e-06	2.74e-06	3.14e-06
	(3.97e-06)	(4.42e-06)	(4.58e-06)	(4.39e-06)
Distance to the nearest college(ft)	-4.79e-06**	-4.81e-06**	-4.64e-06**	-4.85e-06**
	(2.10e-06)	(2.23e-06)	(2.33e-06)	(2.24e-06)
Distance to a school	1.79e-06	3.88e-06	3.94e-06	4.25e-06
	(6.51e-06)	(7.37e-06)	(7.28e-06)	(7.37e-06)
Distance to a park	-1.11e-06	3.41e-07	5.79e-07	5.29e-07
	(4.05e-06)	(4.77e-06)	(4.80e-06)	(4.77e-06)
Distance to the nearest golf course(ft)	-1.03e-05	-9.18e-06	-9.21e-06	-9.24e-06
	(8.19e-06)	(7.86e-06)	(7.82e-06)	(7.88e-06)
Distance to EPA site	-5.45e-06	-7.37e-06	-7.10e-06	-6.88e-06
	(7.20e-06)	(8.44e-06)	(8.55e-06)	(8.48e-06)
Distance to boat launch (recreation)	5.34e-06**	4.98e-06**	4.71e-06**	5.07e-06**
	(2.63e-06)	(2.45e-06)	(2.32e-06)	(2.48e-06)
Distance to ethanol plant	2.65e-06***	3.27e-06***	3.17e-06**	3.23e-06***
	(9.14e-07)	(1.23e-06)	(1.23e-06)	(1.22e-06)
Distance to rail terminal	6.74e-06	5.48e-06	5.45e-06	5.53e-06
	(6.95e-06)	(6.87e-06)	(6.82e-06)	(6.89e-06)
Distance to an oilwell	1.72e-06	1.45e-06	1.48e-06	1.55e-06
	(1.50e-06)	(1.55e-06)	(1.53e-06)	(1.54e-06)
Distance to an inactive oil well	-1.03e-06	-1.37e-06	-1.23e-06	-1.36e-06
	(1.13e-06)	(1.12e-06)	(1.11e-06)	(1.12e-06)
Distance to highway on-ramp	4.29e-06	4.19e-06	4.21e-06	4.29e-06
	(2.87e-06)	(2.91e-06)	(2.96e-06)	(2.90e-06)
Slope of parcel	0.000475***	0.000445**	0.000442**	0.000457**

	(0.000175)	(0.000191)	(0.000189)	(0.000193)
Tree cover on parcel	-0.00562***	-0.00604***	-0.00603***	-0.00596***
	(0.00138)	(0.00148)	(0.00147)	(0.00150)
Available Water Storage (mm)	-0.000216	3.86e-05	-4.52e-06	6.94e-05
	(0.00103)	(0.00103)	(0.00105)	(0.00105)
Thickness(cm) in soil organic carbon calc	0.000278	3.31e-06	-4.05e-05	-4.74e-05
	(0.00101)	(0.00115)	(0.00114)	(0.00115)
Soil organic carbon (g C per m^2)	8.57e-07	5.22e-07	5.78e-07	5.84e-07
	(4.36e-06)	(4.22e-06)	(4.26e-06)	(4.26e-06)
National commodity crop productivity index	0.412	0.434	0.446	0.419
	(0.319)	(0.315)	(0.313)	(0.321)
Percent earthy major components	-0.00963**	-0.0100**	-0.00974**	-0.00970**
	(0.00401)	(0.00432)	(0.00425)	(0.00428)
Maximum rooting depths	0.00255***	0.00239***	0.00239***	0.00229***
	(0.000873)	(0.000845)	(0.000843)	(0.000854)
Water storage within crop root zone depths	-0.00162	-0.00174	-0.00173	-0.00172
	(0.00108)	(0.00107)	(0.00106)	(0.00108)
Droughty Soil Landscapes	-0.120	-0.131	-0.134	-0.143
	(0.167)	(0.168)	(0.166)	(0.170)
Potential wetland soil landscapes	0.000184	0.000181	0.000179	0.000200
	(0.000277)	(0.000271)	(0.000272)	(0.000269)
Distance to large farms			7.24e-07	
			(2.02e-06)	
Large farms within 10 miles radius			0.00825	
			(0.0140)	
2008.year	0.0669	0.0630	0.0611	0.0670
	(0.0702)	(0.0694)	(0.0679)	(0.0696)
2009.year	0.0635	0.0639	0.0627	0.0687
	(0.0728)	(0.0710)	(0.0706)	(0.0714)
2010.year	0.0591	0.0605	0.0590	0.0642
	(0.0733)	(0.0719)	(0.0709)	(0.0730)
2011.year	0.0326	0.0337	0.0333	0.0374
	(0.0866)	(0.0846)	(0.0835)	(0.0855)
2012.year	-0.0423	-0.0336	-0.0351	-0.0290
	(0.0876)	(0.0862)	(0.0859)	(0.0873)
2013.year	0.0189	0.0311	0.0293	0.0349
	(0.0834)	(0.0804)	(0.0794)	(0.0809)
2014.year	-0.0773	-0.0605	-0.0350	-0.0557

	(0.125)	(0.122)	(0.139)	(0.122)
2015.year	0.0814	0.0918	0.119	0.0970
	(0.0850)	(0.0831)	(0.0989)	(0.0832)
2.regionnum	0.0380	0.0692	0.0740	0.0753
	(0.202)	(0.207)	(0.204)	(0.208)
3.regionnum	0.494*	0.506**	0.508**	0.520**
	(0.249)	(0.247)	(0.238)	(0.246)
4.regionnum	0.0853	0.0941	0.0781	0.0950
	(0.215)	(0.217)	(0.224)	(0.218)
5.regionnum	0.527	0.533	0.513	0.545
	(0.348)	(0.343)	(0.362)	(0.340)
6.regionnum	-0.132	-0.0628	-0.0582	-0.0556
	(0.163)	(0.192)	(0.189)	(0.193)
7.regionnum	-0.502**	-0.462**	-0.462**	-0.466**
	(0.192)	(0.197)	(0.194)	(0.196)
8.regionnum	-0.0110	0.00643	0.0186	0.0173
	(0.210)	(0.211)	(0.202)	(0.211)
9.regionnum	0.262	0.326	0.342	0.331
	(0.323)	(0.328)	(0.320)	(0.325)
Constant	9.143***	9.124***	9.051***	9.124***
	(0.472)	(0.481)	(0.486)	(0.412)
Observations	6806	6806	6806	6384
R-squared	0.309	0.305	0.306	0.305
Year FE	YES	YES	YES	YES
Region FE	YES	YES	YES	YES
Acre Weighted	YES	YES	YES	YES
County Cluster	YES	YES	YES	YES

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix For Chapter 4

Experiment Materials: Field Choice Experiment and Survey (Block 1)

Agricultural Insurance Agents' Willingness / Preference to Offer Livestock Insurance

NARRATIVE TO BE READ TO RESPONDENTS:

Thank you for taking the time to meet with us. This study includes two parts (choice cards and survey), and it should take approximately 30-40 minutes. The study is a joint work between China Agriculture University and Cornell University in the United States. The study mainly collects information relates to livestock insurance, it targets insurance practitioners, local administrators who assisted agricultural insurance companies in their work, and rural cooperative workers.

Your responses will be completely confidential and under no circumstances will your responses be identifiable. In addition, we understand that you may not have all the precise information available. In these cases, all we ask is that you provide us with your best estimates or best judgments. Finally, you have the right to refuse to answer any question we might ask.

Before we get into the main experiment, could you please answer two pre-questions:

Pre-1. What is the animal type that you are most familiar with or primarily insure for?

- ○ 1. Large animals (live cattle, dairy cow, yak, horse, camel, etc)
- ○ 2. Medium-sized livestock (capable of breeding sows, fattening pigs, sheep, etc.)
- ○ 3. Poultry and small animals (chicken, duck, rabbit, etc.)
- ○ 4. Aquatic animals (fish, shrimp, crabs, etc.)

Pre-2. What is the client type that you are most familiar with or primarily insure for?

○ 1. Small farms ○ 2. Large farms ○ 3. Feeding enterprises

Part 1: Choice Experiment (Livestock Insurance Products)

Introduction to experiment selection method

This section contains 15 choice cards about livestock insurance, each with 2 choices, combining different levels of eight core insurance attributes. The leftmost column of each card is the name and introduction of different attributes, the two right columns are two livestock insurance products (A and B), and each column in each card is an insurance product. At most four attributes out of eight in the two options of each card are different. Please read carefully and comprehensively evaluate the attributes of the insurance, and then choose the insurance that you are most willing to

provide among the two-column products A and B. Even if both options are not satisfactory, please choose the relatively satisfactory one.

For example, Livestock insurance A in the first card is a deductible of 20%, currently facing low risk, include market catastrophic loss compensation, the insurance premium rate is 1%, ..., can be used to mortgage loans up to 50% of the protected value. If you are familiar with the main insurance attributes, please go directly to card 1 to select. Otherwise, please read the explanations first:

Definition of main insurance attributes

1. Starting point for loss settlement (deductible): If the starting point for claims is 20%, that is, when the insured agricultural product loss rate exceeds 20% due to natural disasters, the compensation will be paid. The amount of compensation is the part that exceeds the deductible point. For example, when the loss reaches 25%, a 5% amount will be indemnified.

2. Risk environment: There are three levels: low, medium, and high. There is no specific definition here. You need to define your own judgment. For example, you may think that the recent combination of multiple disasters such as African swine fever, trade war, new crown epidemic, floods, etc. is a high-risk situation that is rare in a century. You may also think that it is a medium-risk situation. If you see greater risks than the current environment.

3. Compensation coverage for market risk losses: For example, when market changes such as trade wars, financial crises, and economic depressions cause sudden changes in livestock and feed prices, even if there are no natural disasters, farmers will face large economic losses.

4. Actuarial premium rate + loadings (premium rate before government subsidies): the original actuarial insurance premium rate calculated based on the risk losses, plus the operating expenses of the insurance company and the regular rate of return.

5. Insurance premium subsidy rate: the insurance premium subsidy provided by the government or institution, for example, the government at all levels provides a total of 80% of the subsidy, then the farmers need to pay the insurance company at a rate of 20% of the actuarial rate.

6. Forms of insurance: livestock mortality insurance, index-based livestock insurance, revenue or price insurance, livestock feeding net profit insurance.

7. Compensation coverage for uninsured disaster events (such as war, strike, riot, terrorism, etc.) and property damage (such as fire, theft, etc.).

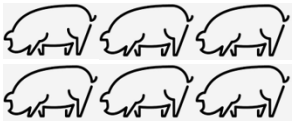
8. Loan allowance: Can the insured insurance be used as proof to apply for a loan? The farm can be granted a loan at 30% to 70% of the amount protected value of the insurance policy.

The following are 15 choice cards, please start to read and select, thank you for your careful consideration:

Choice Card 1 (Choose between A and B):

Block 1 Card 1	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	20% 	
Risk Environment	Low 	
Market Disaster Payment	included 	Not included 
Premium Actuarial + Loading (before subsidy)	1% of protect value 	12% of protect value 
Premium Subsidy (by gov and org)	80% (farmers pay 20%) 	
Coverage Types	Index IBLI 	Margin Or Ratio 
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	included 	
Loan Allowance	Can be allowed loan for **<u>50%</u>** of protected value 	

Choice Card 2 (Choose between A and B):

Block 1 Card 2	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	30% 	15% 
Risk Environment	Moderate 	Low 
Market Disaster Payment	Included 	
Premium Actuarial + Loading (before subsidy)	25% of protected value 	
Premium Subsidy (by gov and org)	80% (farmers pay 20%) 	50% (farmers pay 50%) 
Coverage Types	Mortality 	Index IBLI 
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	0%(Can **<u>not</u>** be allowed for a loan) 	

Choice Card 3 (Choose between A and B):

Block 1 Card 3	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	25% 	
Risk Environment	High 	Low 
Market Disaster Payment	Not included 	
Premium Actuarial + Loading (before subsidy)	6% of protected value 	
Premium Subsidy (by gov and org)	80% (farmers pay 20%) 	90% (farmers pay 10%) 
Coverage Types	Index IBLI 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Included 	
Loan Allowance	0%(Can **<u>not</u>** be allowed for a loan) 	Can be allowed loan for **<u>30%</u>** of protected value 

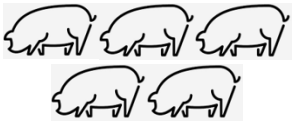
Choice Card 4 (Choose between A and B):

Block 1 Card 4	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	10% 	
Risk Environment	High 	
Market Disaster Payment	Not included 	
Premium Actuarial + Loading (before subsidy)	25% of protected value 	
Premium Subsidy (by gov and org)	80% (farmers pay 20%) 	50% (farmers pay 50%) 
Coverage Types	Index IBLI 	Mortality 
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Included 	
Loan Allowance	Can be allowed loan for **<u>90%</u>** of protected value 	

Choice Card 5 (Choose between A and B):

Block 1 Card 5	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	20% 	15% 
Risk Environment	Moderate 	
Market Disaster Payment	Included 	
Premium Actuarial + Loading (before subsidy)	25% of protected value 	
Premium Subsidy (by gov and org)	50% (farmers pay 50%) 	90% (farmers pay 10%) 
Coverage Types	Mortality 	Margin Or Ratio 
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	Can be allowed loan for **<u>90%</u>** of protected value 	0%(Can **<u>not</u>** be allowed for a loan) 

Choice Card 6 (Choose between A and B):

Block 1 Card 6	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	15% 	25% 
Risk Environment	Moderate 	
Market Disaster Payment	Not included 	Included 
Premium Actuarial + Loading (before subsidy)	6% of protected value 	
Premium Subsidy (by gov and org)	60% (farmers pay 40%) 	
Coverage Types	Revenue Or Price 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	0%(Can **<u>not</u>** be allowed for a loan) 	

Choice Card 7 (Choose between A and B):

Block 1 Card 7	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	10% 	
Risk Environment	Low 	
Market Disaster Payment	Included 	
Premium Actuarial + Loading (before subsidy)	1% of protected value 	
Premium Subsidy (by gov and org)	90% (farmers pay 10%) 	50% (farmers pay 50%) 
Coverage Types	Revenue Or Price 	Mortality 
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Included 	Not Included 
Loan Allowance	0%(Can **<u>not</u>** be allowed for a loan) 	Can be allowed loan for **<u>30%</u>** of protected value 

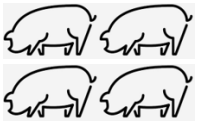
Choice Card 8 (Choose between A and B):

Block 1 Card 8	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	15% 	10% 
Risk Environment	Low 	
Market Disaster Payment	Not included 	
Premium Actuarial + Loading (before subsidy)	6% of protected value 	
Premium Subsidy (by gov and org)	60% (farmers pay 40%) 	50% (farmers pay 50%) 
Coverage Types	Mortality 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	Can be allowed loan for **<u>90%</u>** of protected value 	

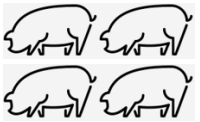
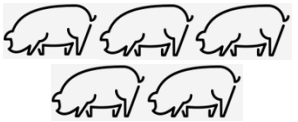
Choice Card 9 (Choose between A and B):

Block 1 Card 9	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	30% 	
Risk Environment	High 	
Market Disaster Payment	Not Included 	Included 
Premium Actuarial + Loading (before subsidy)	25% of protected value 	
Premium Subsidy (by gov and org)	70% (farmers pay 30%) 	
Coverage Types	Revenue Or Price 	Margin Or Ratio 
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	Can be allowed loan for **<u>70%</u>** of protected value 	0%(Can **<u>not</u>** be allowed for a loan) 

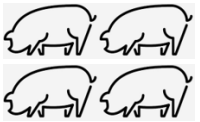
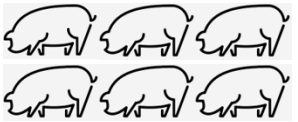
Choice Card 10 (Choose between A and B):

Block 1 Card 10	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	10% 	20% 
Risk Environment	Moderate 	
Market Disaster Payment	Not Included 	
Premium Actuarial + Loading (before subsidy)	6% of protected value 	
Premium Subsidy (by gov and org)	70% (farmers pay 30%) 	90% (farmers pay 10%) 
Coverage Types	Mortality 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Included 	
Loan Allowance	0%(Can **<u>not</u>** be allowed for a loan) 	

Choice Card 11 (Choose between A and B):

Block 1 Card 11	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	20% 	25% 
Risk Environment	High 	Low 
Market Disaster Payment	Included 	
Premium Actuarial + Loading (before subsidy)	3% of protected value 	
Premium Subsidy (by gov and org)	80% (farmers pay 20%) 	
Coverage Types	Mortality 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	Included 
Loan Allowance	Can be allowed loan for **<u>70%</u>** of protected value 	0% (Can **<u>not</u>** be allowed for a loan) 

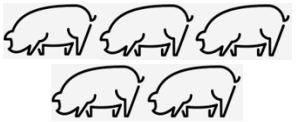
Choice Card 12 (Choose between A and B):

Block 1 Card 12	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	20% 	30% 
Risk Environment	High 	Low 
Market Disaster Payment	Not Included 	Included 
Premium Actuarial + Loading (before subsidy)	1% of protected value 	3% of protected value 
Premium Subsidy (by gov and org)	80% (farmers pay 20%) 	
Coverage Types	Index IBLI 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	Can be allowed loan for **<u>30%</u>** of protected value 	

Choice Card 13 (Choose between A and B):

Block 1 Card 13	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	20% 	
Risk Environment	Low 	
Market Disaster Payment	Not Included 	
Premium Actuarial + Loading (before subsidy)	12% of protected value 	
Premium Subsidy (by gov and org)	60% (farmers pay 40%) 	80% (farmers pay 20%) 
Coverage Types	Margin Or Ratio 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	
Loan Allowance	Can be allowed loan for **<u>90%</u>** of protected value 	Can be allowed loan for **<u>70%</u>** of protected value 

Choice Card 14 (Choose between A and B):

Block 1 Card 14	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	25% 	10% 
Risk Environment	Low 	
Market Disaster Payment	Not Included 	Included 
Premium Actuarial + Loading (before subsidy)	6% of protected value 	
Premium Subsidy (by gov and org)	60% (farmers pay 40%) 	90% (farmers pay 10%) 
Coverage Types	Revenue Or Price 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Not Included 	Included 
Loan Allowance	Can be allowed loan for **<u>70%</u>** of protected value 	

Choice Card 15 (Choose between A and B):

Block 1 Card 15	Livestock Insurance A	Livestock Insurance B
Claim Starting Point (Deductible, or coverage level = 1 – deductible)	15% 	
Risk Environment	Low 	
Market Disaster Payment	Included 	
Premium Actuarial + Loading (before subsidy)	3% of protected value 	1% of protected value 
Premium Subsidy (by gov and org)	70% (farmers pay 30%) 	
Coverage Types	Mortality 	
Uninsured Catastrophic and (e.g. war, terrorism, etc.) **Property Casualty Coverage** (e.g. theft, fire, etc.)	Included 	
Loan Allowance	Can be allowed loan for **50%** of protected value 	Can be allowed loan for **70%** of protected value 

Part 2: Survey Questions (four groups A, B, C, D)

Number	Category	Question	Unit	Answer
A1	A. Basic information	Gender	1=Female, 2=Male	
A2		Age	Years Old (E.g. 26)	
A3		Education	1=High school 2=College 3=MS, MBA 4=PhD	
A4		Affiliated company level	1=City / county subdivision 2=Provincial branch 3=Central national headquarter 4=Others (Please explain)	
A5		Position level within the affiliated company, branch, subdivision	1=Entry level associate (working less than 5 years) 2=Middle level (manager/ director/ specialist) 3=President, VP, CXO, CFO Acturaial, claim adjuster, … 4=Others	
A6		Agriculture / Farm / Ranch background growing up or had systematic training (college major in ag etc.)	1=None 2=Some 3=Extensive	
B1	B. Risk perception	Are you worried about natural risks (weather or disease) in livestock production?	1=A little worried 2=Somewhat worried 3=Very much worried	
B2		Are you worried about market risks (price) in livestock production?	1=A little worried 2=Somewhat worried 3=Very much worried	

B3		Regarding agricultural risks, what are you most worried about? (Your insurance product caused the largest or most frequent payments) [Please sort from the most worried to the least, 6 is the most, 1 is the relatively least worried. Please fill the ranking number in the brackets]	[ ] Weather (drought, flood, frost, etc.) [ ] Grain pests or livestock diseases [ ]Market price [ ] Supply Chain Sales Channel [ ] Natural disasters (typhoons, earthquakes, etc.) [ ] Property damage (theft, fire, riot, etc.)	
B4		In general, are you a risk-taker (take bets that may lose)	1=No 2=Yes	
B5		From what you have observed: in addition to purchasing agricultural insurance, the other main way for farmers to diversify risks and to stabilize income? (Choose one)	1=Employed in a wage work 2=Low cost farm operation 3=Multiple crops/ feeding to be comprehensive (diversification) 4=Invest in technologies 5=Invest in irrigation system 6=Do not know or Others	
B6		Does your company regularly hold training on risk estimation and management?	1=No 2=Yes	
C1	Agricultural insurance offering experience	Years working in all types of the insurance industry (health, home, auto, etc)	Years	
C2		Years working specifically in Agricultural insurance	Years	
C3		What agricultural insurance are you mainly engaged in? (You can describe the major categories of crops, livestock, and aquatic products, or it can be specific to specific crops or	Livestock Crop Livestock and Crop Aquaculture	

		animal species)	Agricultural property casualty	
C4		Does your company help insured farmers conduct risk management, disaster and epidemic prevention?	1=No 2=Yes	
C5		Have you experienced conflict in assessing indemnity / loss?	1=No 2=Yes	
C6		Your company's Input in Ag insurance product design?	1=No Input (purely listen to gov's design in all) 2=Some Input 3=Extensive (only listen to the main structure of gov advice) 4=Total Freedom 5= Do not know	
C7		How flexible is your company in claim payment settlement for ag insurance products?	1=Not flexible at all 2=Somewhat flexible 3=Very flexible 4=Others	
C8		Your organization has experience in overlapping indemnity and government disaster assistance programs.	1=No 2=Yes 3=Do not know	
C9		Your company penalizes agents' salary for overpayment of polices they sold	1=No 2=Yes 3=Other punishment methods	
C10		Your company rewards a salary bonus for more policies that agents sell	1=No 2=Yes 3=Others	
C11		Does your company offer villages or coops group	1=No 2=Yes	

		insurance?	3=Do not know	
C12		Does your company purchase reinsurance?	1=No 2=Yes 3=Do not know	
C13		How does your company deal with the overpayment situation?	[fill in the blanks] (more opinion based)	
C14		Problems you have encountered in designing and selling agricultural insurance	[fill in the blanks] (more opinion based)	
C15		Is your company deeply involved in bank loan-insurance linkage (after farmers purchase breeding insurance and loan guarantee insurance, banks can provide low-interest, low-fee, collateral needless pure credit loans, etc. according to their credit status)?	1=No 2=Yes	
C16		What do you know about the payment capacity of the local gov finance dept?	1=Poor 2=General 3=Good	
D1	Ideal and feasible attribute levels of insurance products	Your knowledge about Index-Based Livestock Insurance?	1=A little (aware of) 2=Some (trained but not yet offered) 3=Extensive (has offered for some years or researched deeply)	
D2		Your knowledge about Livestock Mortality Insurance?	1=A little (aware of) 2=Some (trained but not yet offered) 3=Extensive (has offered for some years or researched	

			deeply)	
D3		Your knowledge about Livestock Gross revenue (Price*Live Weight) Insurance?	1=A little (aware of) 2=Some (trained but not yet offered) 3=Extensive (has offered for some years or researched deeply)	
D4		Your knowledge about Livestock Feeding Margin Insurance (Price*Live Wgt – Feed Cost), or Hog to corn price ratio Insurance (e.g. mainly 6:1 and 2:1 for extreme case lumpsum payment)?	1=A little (aware of) 2=Some (trained but not yet offered) 3=Extensive (has offered for some years or researched deeply)	
D5		What type of livestock (mainly pigs) insurance are you most optimistic about (can be widely promoted, efficient, trusted by farmers, etc.)? [Please sort from the most liked to the least, 1 is the most, 5 is the least. Please fill the ranking number in the brackets]	[ ] Feeding Cost Insurance [ ] Mortality Insurance [ ] Weather Index Insurance (IBLI, vegetation, etc.) [ ] Livestock Price or Revenue Index Insurance [ ] Animal Feeding Net Margin (or Hog/Corn) Insurance	
D6		Regarding pig insurance, what do you think is the relative balance point of the most reasonable (acceptable by farmers, profitable insurance company) insurance premium rate (after government subsidies)?	1=premium rate of 4% 2=premium rate of 5% 3=premium rate of 6% 4=premium rate of 7% 5=premium rate of 8% 6=premium rate of 9% 7=Other (please explain) _____ *	

D7		For pig insurance, the relative balance point of the deductible (starting point of loss settlement) that you think is the most reasonable (acceptable by farmers and the insurance company's risk is controllable)	1=Deductible of 10% 2=Deductible of 15% 3=Deductible of 20% 4=Deductible of 25% 5=Deductible of 30% 6=Deductible of 40% 7=Other (please explain) _____ *	
D8		Regarding pig insurance, what do you think is the relative balance point of the most reasonable premium subsidy rate (total financial subsidies at all levels) (financial expenditures can be afforded, insurance companies benefit, farmers accept, and can be widely promoted)?	1=Premium subsidy of 40% 2=Premium subsidy of 50% 3=Premium subsidy of 60% 4=Premium subsidy of 70% 5=Premium subsidy of 80% 6=Premium subsidy of 90% 7=Other (please explain) _____ *	
D9		Regarding pig insurance, what do you think is the relative balance point of the most reasonable credit for mortgage loans?	1=Mortgage credit of 0% 2=Mortgage credit of 30% 3=Mortgage credit of 50% 4=Mortgage credit of 70% 5=Mortgage credit of 90% 6=Other (please explain) _____ *	
D10		Do you think livestock insurance should include coverage on uninsured disaster events (e.g. war, strike, riot, terrorism, etc.) and property damage (e.g. fire, theft, etc.) in your pig insurance	1=No, should not 2=Yes, should	

D11		Do you think livestock insurance should include coverage on major animal diseases (African swine fever, blue ear disease, foot-and-mouth disease, pseudorabies, etc.)?	1=No, should not 2=Yes, should	
D12		Do you think livestock insurance should include market risk loss (trade war, financial crisis, etc.) in your pig insurance?	1=No, should not 2=Yes, should	

The survey is completed, thank you very much for your cooperation!