THREE ESSAYS ON MARKET EFFICIENCY: GLOBAL PRICE LEADERSHIP, INFORMAL PARALLEL MARKETS, AND MARKET MICROSTRUCTURE

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

Thanasin Tanompongphandh

January 2011

THREE ESSAYS ON MARKET EFFICIENCY:
GLOBAL PRICE LEADERSHIP, INFORMAL PARALLEL MARKETS, AND MARKET MICROSTRUCTURE

Thanasin Tanompongphandh, Ph.D.
Cornell University 2011

This dissertation tackles the concept of market efficiency from three distinct topics in applied economics, from microfinance, to agriculture commodity market, and further to market microstructure of the most advanced economy.

The first essay, entitled "Market Efficiency and Price Discovery Among Leading Rice Exporting Countries", focuses on the issue of rice market efficiency. The study establishes, under Johansen's procedure, that there are long-run price co-movements existing among the three major rice-exporting countries, and within the United States domestic markets, the long-run efficient linkage between spot and future prices of rough rice, as Chicago Board of Trade rough rice futures converge to United States Department of Agriculture rough rice prices in a cash market. Regarding the efficiency among the export market prices, results show that the hypothesis of market efficiency are rejected in two of the three pairs, namely Thai-Vietnam and Thai-US(Arkansas). The Gonzalo & Granger (1995) decomposition method finds that the Thai and United States rice are dominant in the price discovery process. Within the United States domestic markets, the dominant is the futures market followed by the cash market of the rough rice and then the milled rice export price.

The second essay, entitled "Determinants for Formal Credit and Informal Credit Access: The Case of Thai Farm Households", examines determinants for Thai agricultural households' participation in formal and its informal parallel credit

markets. The study follows Heckman's two-stage selection model (1979) approach to determine the informal loan participation of Thai agricultural households. Results reveal that households tend to 'stick' to the credit market in which they were previously engaged. This finding reinforces the vicious cycle which makes it more difficult for farmers to get out of debt. Secondly, the study finds that wealthier households are less likely to access credit, and are more likely to participate in formal credits than their less wealthy peers. Results also show less probability of credit access between May and December coinciding with the planting and harvesting season accentuating the nature of loans as working-capital rather than consumption loans. Finally, the study discovers that households with owned farmland are more likely to participate in the formal credit market, while households with rented farmland are more likely to participate in the informal credit market stressing the use of owned land as collateral to participate in the former.

The final essay, entitled "On the Challenge of Testing Weak-Form Market Efficiency using High Frequency Data", explores the issue of efficiency in microstructure of the Exchange-Traded-Fund (ETF). This essay shows that the profitability of a simple technical trading strategy hinges heavily on the way the Trades And Quotes (TAQ) dataset is filtered for mistakes and outliers. This paper uses ultra-high-frequency TAQ data that cover the time-span since the inception of the S&P 500 ETF from January 1993 to December 2006. First, a widely used filtering methodology proposed by Hasbrouck (2003) is adopted. Under this methodology, the technical trading strategy clearly outperforms the buy-and-hold benchmark. However, when a more appropriate (stringent) filtering methodology is used, the technical trading strategy clearly *underperforms* the buy-and-hold benchmark. This evidence suggests that studies that based their methodology on Hasbrouck's (2003) less stringent filtering criterion could produce misleading results.

BIOGRAPHICAL SKETCH

Thanasin Tanompongphandh was born in March 1981 in Bangkok, Thailand. He completed the Bachelor of Arts Degree in Economics with *First Class Honors* from Thammasat University in 2003. Under the Royal Thai Government Scholarship he earned his Master's degree and Doctor of Philosophy's from Dyson School of Applied Economics and Management, Cornell University in 2008 and 2011 respectively.

To my beloved family

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Professor Nancy Chau, for her invaluable advice, motivation, guidance, patience, and unfailing support throughout the entire length of this research. I also would like to thank Professor Timothy Mount, Hazem Daouk, and Jeffrey Prince, for a direction and for the excellence suggestions that shape my dissertation in many ways.

I would like to thank my friend, roommate and mentor, Nattapong Puttanapong, who has shown me the life of a well-respected PhD student, and who has inspired me tremendously.

I would like to extend my appreciation to all my classmates, especially Rahul Annand, Chandra Kumars, Wei Cen, Masato Nakane and Samuel Bell for help with constructive ideas and useful comments. I am also grateful for the Thai community in the New York area, where I have made life-long friendships, and Orn Banternghansa whose caring and compassionate love always uplifted me when I needed it most. Last but not least I would like thank my beloved family who provided love and endless emotional support throughout the years of my graduate student life. Without them, none of this would been possible.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE:

CHAPTER TWO:

CHAPTER THREE:

LIST OF FIGURES

CHAPTER ONE:

CHAPTER TWO:

CHAPTER THREE:

CHAPTER FOUR:

CHAPTER ONE

INTRODUCTION

This dissertation tackles the concept of market efficiency under three distinct topics in applied economics: microfinance, the agricultural commodity market, and market microstructure of the most advanced economy.

The first essay, entitled "Market Efficiency and Price Discovery Among Leading Rice Exporting Countries", focuses on the issue of rice market efficiency. The international rice market is best characterized as a segmented market with imperfect substitution. The existence of non-competitive pricing and absence of a centralized trading system allows countries with market power to exert a degree of price leadership. The key to understanding the behavior of the rice market, hence, lies with an ability to assess the degree of price discovery among main participants. To this end, the study sets up two main objectives. First it seeks to establish, under Johansen's procedure, that there are linkages, represented by price transmission elasticity, where price innovations from one market transmit to another among the top three rice exporting countries, namely Thailand, Vietnam, and the United States(Arkansas market), and, within the United States domestic markets, including the Chicago agricultural futures market. Secondly, it will utilize the Gonzalo & Granger decomposition method to uncover long-run price discovery among each pair of markets.

Consequently, through three bivariate cointegrated systems, the study shows that long-run price co-movements exist among the three major rice-exporting countries. In addition to this finding, the hypothesis of market efficiency, captured by the price transmission elasticity, is rejected in two of the three pairs, namely Thai-Vietnam and Thai-US (Arkansas). The magnitudes of the estimated price transmission elasticities reveal that Thai rice price shows higher acceleration in the long-run co-movement than that of US (Arkansas) and Vietnamese rice. To examine this further, the Gonzalo & Granger common factor decomposition technique finds that the Thai rice price is a dominant force in the price discovery process among the three. This finding is sensible as Thailand is the world's largest rice exporter.

To evaluate the performance of the United State Department of Agriculture (USDA) adjusted world price relative to other major rice exporting countries, this study looks for evidence of long-run relationship of the adjusted world price with Thailand, Vietnam, and United States (Arkansas). In this regard, the USDA adjusted world price shows only a long-run linkage with Vietnamese market. The size of the estimated coefficient far from 1 also pinpoints that there could be other factors influencing the USDA world price. This, combined with the fact that there is a failure to establish long-run price relationships between the USDA world rice price with the other two leading rice exporting markets, underscores a weakened argument to support validity and transparency of the USDA's methodology. This conclusion reinforces Taylor, Bessler, Waller, and Rister's (1996)'s skepticism that other considerations besides market force are in place for formulation of the USDA adjusted world price.

For the US domestic rice markets, this study finds the efficient linkage between spot and future prices of rough rice, as the Chicago Board of Trade (CBOT) rough rice futures converge to USDA rough rice prices in a cash market. In the vertical markets, the result also reveals the significant price linkage between two marketing levels of rough and milled rice in the United States. Building on these linkages, the Gonzalo & Granger's decomposition characterizes the dominant source of the price discovery process in the futures market followed by the cash market of the rough rice and then the milled rice export price.

The second essay, entitled "Determinants for Formal Credit and Informal Credit Access: The Case of Thai Farm Households", examines determinants for Thai agricultural households' participation in formal and informal credit markets. Understanding the workings of the parallel market is key to boosting the efficiency of the formal credit channel. Using the 2006 Socio-Economic Household Survey of Thailand, the study follows Heckman's two-stage selection model approach to determine the informal loan participation of Thai agricultural households. Since one can observe the behavior of formal and informal credit participants only if they participate in new loans, one might view this pool of samples as self-selected. The dire consequence of self-selection bias is first illuminated in Heckman (1979). To demonstrate within this context, unobservable variables such as business acumen may affect both the probability of success in obtaining credit as well as probability of success in participating in formal loan. Higher levels of business acumen lead to higher success in obtaining credit, and thus successful borrowers are likely to have higher levels of business acumen than non-borrowers. As a consequence of this, the probability of success in obtaining formal credit is overestimated from the influence of selected samples. Hence, excluding the selection stage to engage in new loans from the full estimation induces a selection bias due to a non-random sample. See Achen (1986) and Sartori (2003) for more discussion of selection bias.

Results from the two-stage selection model provide supporting evidence that, facing credit rationing, households are more likely to 'stick' to the credit market they were previously engaged, whether the formal or informal credit market. This finding reinforces the vicious cycle which makes it more difficult for farmers to get out of debt. Secondly, the study finds supporting evidence that wealthier households are less likely to access credit, but, if they do, they are more likely to participate in formal credits, and the opposite is true of less wealthy households. Thirdly, the study finds supporting evidence for less credit participation during May and December coinciding with the planting and harvesting season. Specifically, the probability for accessing new credit starts off high in the beginning of the year then slowly declines while participation in informal credit, conditional on participating in new credit, rises toward the end of the year. Alarmingly, this indicates that, while overall access to new credit declines, people who need credit tend to participate in informal credit more as months pass. This piece of evidence pinpoints the limited role of Bank for Agriculture and Agricultural Co-operatives (BAAC) in providing working capital only and not consumption-based loans underlining the importance of the lack of formal credit facility for farmers, and should be addressed quickly. Finally, the model finds that households with owned farmland are more likely to participate in the formal credit market, while households with rented farmland are more likely to participate in the informal credit market stressing the use of owned land as collateral to participate in formal credit channels.

The final essay, entitled “On the Challenge of Testing Weak-Form Market Efficiency using High Frequency Data” reveals that inherent data errors and mistakes in high frequency Trades And Quotes (TAQ) data complicate microstructure research. It demonstrates that the profitability of a simple technical trading algorithm hinges heavily on the way the TAQ dataset is filtered for mistakes and outliers. Hasbrouck’s (2003) technique of the 50 cent filter is insufficient to remove outliers that could potentially misguide the result on microstructure research. Using high frequency price data for 1993-2006 from the S&P 500 ETF, this paper shows that by running a simple trading algorithm on price after the Hasbrouck's filter has been applied, excess returns above those from the buy-and-hold strategy are large, significant, and robust. However, when a more appropriate (stringent) filtering treatment is used, the same algorithm clearly *underperforms* the buy-and-hold benchmark. It is clear from this analysis that Hasbrouck's (2003) filter allows many erroneous prices to remain in the data and therefore biases the profitability of the technical trading strategy. This study also recommends a more appropriate filter level of 0.1% as a mean to clean this specific dataset. Finally, these results demonstrate that testing weak form efficiency of high frequency stock prices cannot be done until the issue of appropriate filtering is completely resolved.

CHAPTER TWO

MARKET EFFICIENCY AND PRICE DISCOVERY AMONG LEADING RICE-EXPORTING COUNTRIES

1 INTRODUCTION

Rice is one of the most common commodities in global agricultural trade. In 2009, the USDA estimated that rice is critical to the diet of half of the global population. It also observed that countries with lower income per capita tend to spend a higher percentage of income on food, and a greater amount of caloric intake comes from rice. As income rises, per capita rice consumption declines and is replaced with increasing meat and dairy consumption (see Childs & Kiawu, 2009). This long term growth trend facilitates the view that stable global rice production is crucial to the welfare of low income populations. The challenge, however, lies within the market's infrastructure to bring the surplus of rice produced to international trade. Given the present low trade volume of rice compared to global production volume, the rice market is considered *thin* (see Slayton, 2009). Two major concerns arose; first the *thinness* of the rice market could potentially allow international rice prices to be determined by only a few major players. And secondly relying on this *thin* market the capacity of a rice exporting country to meet demand from rice importing countries becomes even more critical. Given the recent rice price peak, in early 2008, following government mandates for export restrictions, the question of steady supply of rice and dietary welfare of people at the margins becomes even more essential.

Before one can formulate a policy relevant to rice trade, the first task would be to understand the unique structure of the global rice trade and importance of linkages among major players. Accordingly, this study aims to understand the price behavior of key rice exporting countries, and once the linkages between rice exporting countries are established, to test whether each pair of markets is efficient. Secondly this study aims to measure the degree of price leadership through price discovery decomposition, specifically each market's long-term contribution to the permanent value of rice. This contribution of price discovery is the key to understanding the organizational behavior of the rice market as it assesses the degree of price leadership among its main participants.

To this end, the paper sets up two main objectives. First it seeks to establish, under Johansen's procedure, that there are linkages, represented by price transmission elasticity, where price innovations from one market transmit to another among the top three rice exporting countries, namely Thailand, Vietnam, and the United States(Arkansas market), and, within the United States domestic markets, including the Chicago agricultural futures market. Secondly, it utilizes the Gonzalo & Granger decomposition method to uncover long-run price discovery among each pair of markets. The possibility of linking the Chicago futures market with leading exporters is also considered, but a framework to directly compare the two seems lacking, since the Chicago future market trades rice futures in rough form while Thailand and Vietnam only export rice in milled form.

This paper follows a bivariate approach in time series in contrast to a multivariate approach due to a limited ability of the latter to identify and interpret coefficients involving market efficiency for a system larger than two markets (see Barret, 2001, Fackler & Goodwin, 2001). Consequently, through three bivariate cointegrated systems, this study shows that long-run price co-movements exist among the three major rice-exporting countries. In addition to this finding, the hypothesis of market efficiency, captured by the price transmission elasticity, are rejected in two of the three pairs, namely Thai-Vietnam and Thai-US (Arkansas). The magnitudes of the estimated price transmission elasticities reveal that the Thai rice price is more volatile in the long-run co-movement than that of US (Arkansas) and Vietnamese rice. To examine further, the Gonzalo & Granger common factor decomposition technique finds that the Thai rice price is a dominant force in the price discovery process, along with the United States. This finding is sensible as Thailand is the world's largest rice exporter while the United States dominates high-quality rice and high-income markets.

To evaluate the performance of the USDA adjusted world price relative to other major rice exporting countries, this study looks for evidence of a long-run relationship of the adjusted world price with Thailand, Vietnam, and United States (Arkansas). In this regard, the USDA adjusted world price shows only a long-run linkage with the Vietnamese market. The size of the estimated coefficient, being far from 1, also pinpoints that there could be other factors influencing the USDA world price. In combination with the fact that we fail to establish long-run price relationships between the USDA world rice price with the other two leading rice exporting markets, underscores a weakened ground to support validity and transparency of the USDA's methodology. This conclusion reinforces Taylor, Bessler, Waller, and Rister's (1996)'s skepticism that other considerations besides market force are in place for formulation of the USDA adjusted world price.

For the US domestic rice markets, this study finds the efficient linkage between spot and future prices of rough rice, as the Chicago Board of Trade rough rice futures converge to USDA rough rice prices in a cash market as the price transmission elasticity does not statistically differ from 1. In the vertical markets, the result also reveals the significant price linkage between two marketing level of rough and milled rice in the United States. Building on these linkages, the Gonzalo & Granger's decomposition characterizes the dominant source of the price discovery process in the futures market followed by the cash market of the rough rice and then the milled rice export price.

For the remainder of the paper, Section 1 introduces and provides the background for rice markets. Section 2 provides a survey of literature on spatial markets. Section 3 briefly visits the main hypotheses. Section 4 discusses empirical strategy. Section 5 discusses sources and time series properties of the data. Results are given in section 6. Section 7 presents policy implications, and Section 8 concludes the paper.

1.1 CHARACTERISTICS OF RICE MARKETS

This section provides brief background on rice markets, including a global supply perspective, consumption, and characteristics that define the international rice market.

1.1.1 Global supply and consumption

Global rice production has increased steadily due to both expansion of harvesting areas and yield improvement through technology and modern irrigation systems. According to the USDA grain database, the global harvesting area slowly increased from 147 million hectares in 1990/91 to over 160 million hectares in

2009/10, or a compound average growth rate of 0.43% over 20 years. Yield improvement of rough rice also saw gradual increases from 3.54 metric tons per hectare in 1990/91 to 4.3 metric tons per hectare in 2009/10, or a compound average growth rate of 1.03%. Milled rice production increased to 459 million tons in 2009/10 from 353 million tons in 1990/91, or a compound average growth rate of 1.36%. Over the same period, rice consumption also rose to 449 million tons from 351 million tons, or a compound average growth rate of 1.31%. Faster production growth allows excess rice production each year to be accumulated for next year's stocks.

1.1.2 Global rice stock-to-use ratio

The gauge to monitor the health of available rice supply is the stock-to-use ratio (figure 3). This ratio compares beginning stock to a year's consumption. While a slight price decrease during 1997 and 2000 corresponded with the general increase in global stock, and a price increase from 2001 to 2005 corresponded with the decrease in global stock of the same period, the peak in 2008 is not associated with a meaningful decrease in global rice stock. This poses a serious issue concerning the rice market infrastructure as a means to allocate rice efficiently.

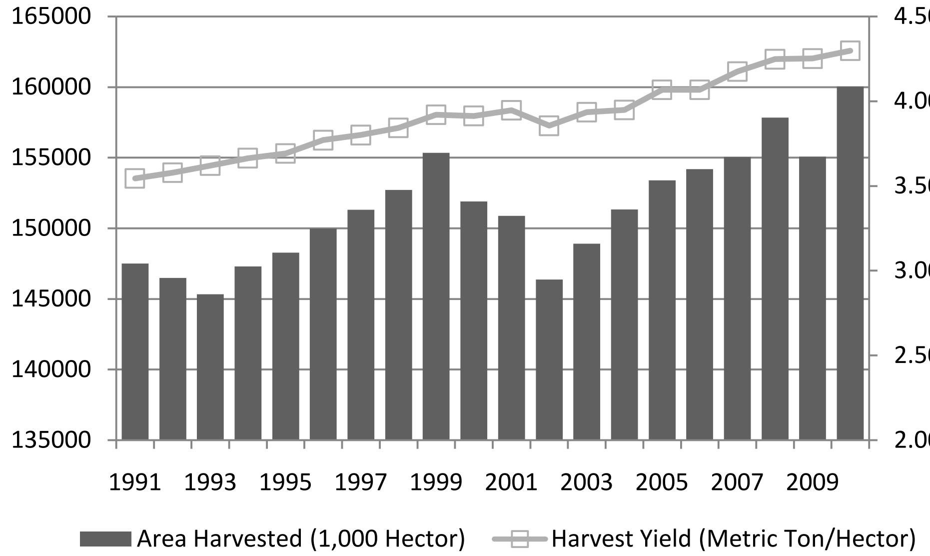


Figure 1 Area Harvested and Rough Rice Yield (rough basis)

Source: USDA Estimates. Foreign Agricultural Service. Author's calculation.

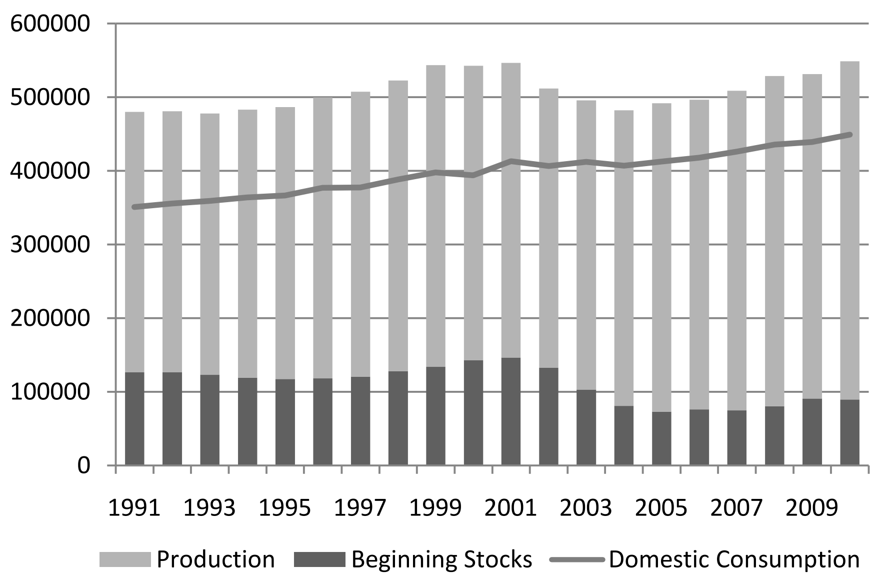


Figure 2 Rice Stocks and Consumption(milled basis)

Source: USDA Estimates. Foreign Agricultural Service. Author's calculation.

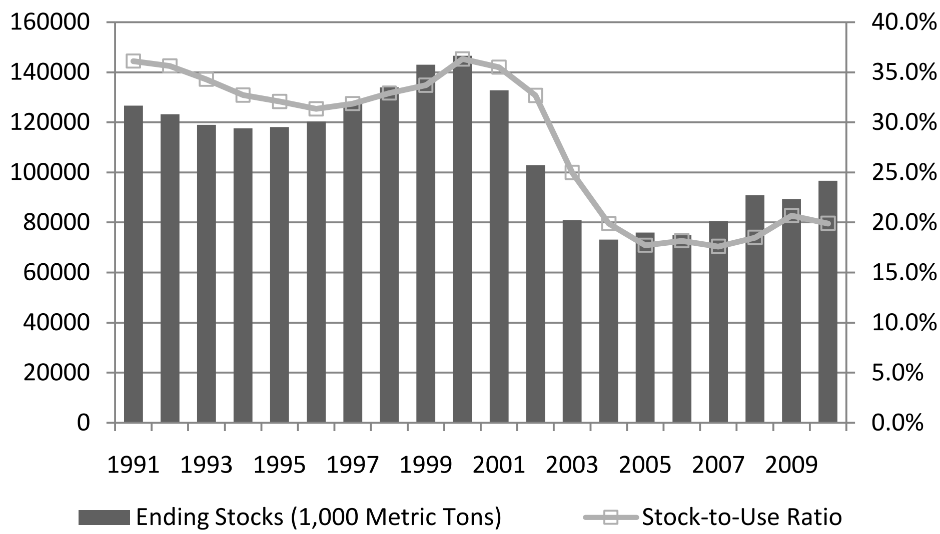


Figure 3 World Rice Ending Stocks and Stock-to-Use Ratio (milled basis)

Source: USDA Estimates. Foreign Agricultural Service. Author's calculation.

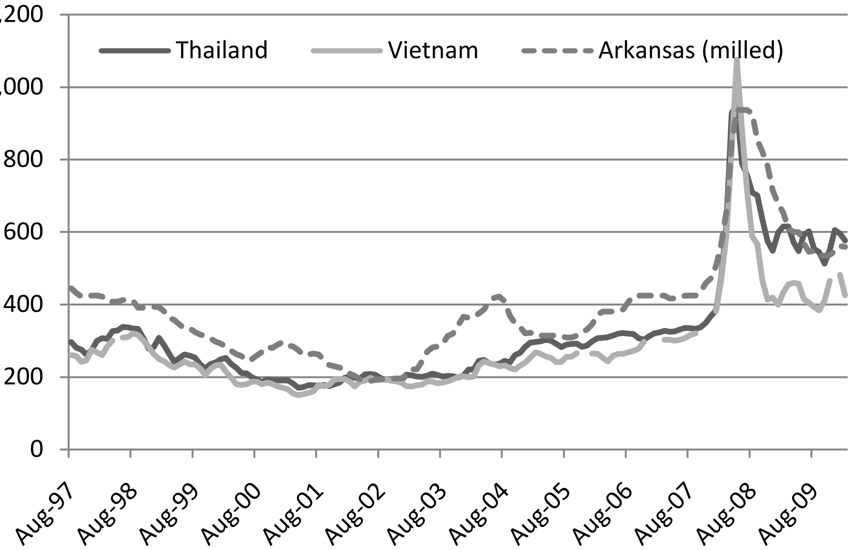


Figure 4 Milled Rice Prices (Unit in US$/Metric Ton)

Source: USDA Rice briefing room.

Table 1 Milled rice production by country (Unit in Thousand Metric Tons)

Production	**2006**	**2007**	**2008**	**2009**	**2010***
China	127200	130224	134330	137000	137500
India	93350	96690	99180	87500	99000
Indonesia	35300	37000	38300	38800	40000
Bangladesh	29000	28800	31000	30500	32300
Vietnam	22922	24375	24393	24380	24750
Thailand	18250	19800	19850	20300	20600
Burma	10600	10730	10150	10597	11000
Philippines	9775	10479	10753	9757	10800
Brazil	7695	8199	8569	7820	8400
Japan	7786	7930	8029	7711	7850
United States	6088	6149	6400	6917	7809
Pakistan	5450	5700	6700	6500	6500
Cambodia	3946	4238	4520	4780	4800
Korea, South	4680	4408	4843	4916	4600
Egypt	4383	4385	4402	4300	3900
Nigeria	2900	3000	3200	3400	3600
Nepal	2804	2810	2850	2900	2900
Madagascar	2240	2304	2505	2688	2688
Sri Lanka	2145	2200	2227	2566	2594
Iran	1724	1850	1500	2000	2050
Other	22117	22338	24326	25253	25638
World Total	420355	433609	448027	440585	459279

Source: USDA Estimates. Foreign Agricultural Service.
*2010 number is USDA Forecast.

Table 2 Rice consumption by country (Unit in Thousand Metric Tons)

Domestic Consumption	2006	2007	2008	2009	2010*
China	127200	127450	133000	134500	135500
India	86700	90466	91090	89300	93500
Indonesia	35900	36350	37090	38100	39500
Bangladesh	29764	30747	31000	31300	32500
Vietnam	18775	19400	19000	19150	19500
Philippines	12000	13499	13650	13614	13700
Burma	10670	10249	9648	10000	10100
Thailand	9780	9600	9500	9600	9800
Brazil	7925	8254	8530	8600	8650
Japan	8250	8177	8326	8200	8125
Nigeria	4400	4500	5150	5300	5500
Korea, South	4887	4670	4788	4750	4740
United States	3959	3919	3957	4001	4050
Cambodia	3646	3788	3770	3960	3970
Egypt	3276	3340	4000	4000	3850
Iran	3294	3297	3350	3400	3500
EU-27	2911	3185	2925	3100	3150
Pakistan	2207	2700	3400	3000	3000
Nepal	2993	3016	2909	2912	2925
Madagascar	2400	2499	2615	2838	2838
Other	37159	37059	38171	39745	40892
World Total	418096	426165	435869	439370	449290

Source: USDA Estimates. Foreign Agricultural Service.
*2010 number is USDA Forecast.

1.1.3 Rice as a differentiated product

In terms of product differentiation, rice is categorized into four main types, namely *Indica*, *Aromatic*, *Japonica*, and *Glutinous*. In 2009, long-grain *Indica* rice dominated global rice trade with a 75% share. Long-grain *Aromatic*, such as Thai's Jasmine and India's Basmati, accounts for a 12%-13% share. Short and medium *Japonica* accounts for around a 10% share of global trade. *Glutinous* accounts for the small remainder. The *Aromatic* rice usually sells at a premium to *Indica* to reflect demand from higher income countries. *Indica* rice is popular in the Middle East, East and Southeast Asia, and Africa. *Japonica* rice is preferred in East Asian countries, the Caribbean, and Eastern European (see Rice:Background, 2009, Bashir, 2002).

1.1.4 Thin market

Although rice is traded globally, rice trade is traditionally viewed as a *"thin market"* due to low trade volume compared to production volume. This is an inherent characteristic of rice markets when compared to other commodity markets (see Siamwalla & Haykin, 1983, Slayton, 2009). Compared to other commodities, in the 2009/10 marketing year, the average volume of trade per production is only 7% for rice, compared to 19% for corn, and 28% for wheat (see figure 6). Not only is the volume of rice trade thin, but only a few countries have export capacity to impact rice trade. Table 1 shows that the largest rice producer in the world by volume is China, but the equally large share of domestic consumption leaves little to export (see Table 2). Indonesia and Bangladesh are a similar story. Based on 2009/10 data, the top five exporting countries by volume are Thailand, Vietnam, Pakistan, the United States, and India (see Table 3). Altogether, five countries export 81% of total international rice trade volume. The fact that few players control so much trade volume makes a case of collusion seem plausible. Nonetheless, studies of market power in the global rice trade seem to suggest otherwise. Karp & Perloff (1989) found that the oligopoly model fits well with the data, but that the degree of market power is low, and closer to competitive equilibrium than to collusion. More recent literature, accounting for rice differentiation (Bashir, 2002), also found that modeling global rice in a Bertrand equilibrium model is the better representation compared to a collusion model, as the conjecture variation suggests that market behavior is closer to competitive than collusive [1].

A low degree of market power does not imply that the market functions efficiently. A thin market by itself could pose a big challenge, since we have already witnessed that an abrupt change in tradable supply conditions could create an upswing of price volatility. India's export restrictions on non-basmati rice and Vietnam's ban on rice exports to calm domestic food price pressure in 2008 coincided with triple-fold increases in international rice prices leading to a record-setting $1,000 U.S dollars in all major export markets, (see Appendix). Surging rice prices will increase nations' food price inflation, as they filter through to consumers.

[1] Recall that the Bertrand equilibrium represents the case of low market power. For a perfectly substitute product with two players, both would use marginal cost pricing.

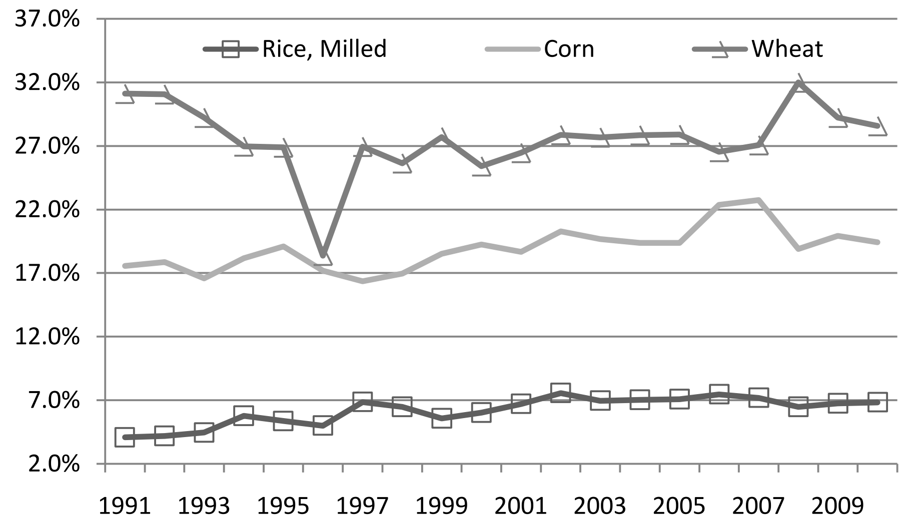


Figure 5 Volume of trade per production (based on weight)

Source: Data from Foreign Agriculture Service and Author's calculation.

Table 3 Rice Exports ranked by country on a milled basis (Unit in Thousand Metric Tons)

Exports	2006	2007	2008	2009	2010*
Thailand	9557	10011	8570	9500	10000
Vietnam	4522	4649	5950	5750	5800
Pakistan	2839	3000	3000	3300	3600
United States	2886	3305	3011	3303	3529
India	5740	4654	2090	2200	2500
China	1340	969	783	850	900
Cambodia	450	500	800	850	850
Uruguay	734	778	987	715	815
Argentina	452	443	554	625	700
Burma	31	541	1052	600	700
Brazil	242	550	569	300	500
Egypt	1203	750	550	600	300
Guyana	250	210	210	250	250
Japan	200	200	200	200	200
EU-27	148	152	140	140	140
Paraguay	67	79	125	163	127
Russia	12	21	90	100	110
Ecuador	161	90	15	40	100
Peru	20	20	80	50	100
Australia	166	36	15	40	65
Other	435	253	273	241	197
World Total	31455	31211	29064	29817	31483

Source: USDA Estimates. Foreign Agricultural Service
*2010 number is USDA Forecast.

Table 4 Rice Imports ranked by country on a milled basis (Unit in Thousand Metric Tons)

Imports	2006	2007	2008	2009	2010*
Philippines	1800	2570	2600	2200	2500
Nigeria	1500	1800	1750	1800	1900
Iran	1500	1550	1470	1300	1500
EU-27	1340	1568	1339	1350	1350
Iraq	613	975	1089	1100	1150
Saudi Arabia	958	961	1166	1049	1100
Malaysia	886	799	1039	1070	1020
Cote d'Ivoire	920	845	800	860	900
South Africa	795	1030	580	800	900
Japan	675	597	656	700	700
Senegal	675	820	683	700	700
United States	653	759	610	635	667
Brazil	732	422	675	950	650
Mexico	594	582	588	600	650
Cuba	574	652	463	565	550
Vietnam	450	300	500	500	500
Hong Kong	348	399	395	400	410
Bangladesh	769	2047	602	90	400
Guinea	230	175	150	330	375
Mozambique	335	450	270	350	375
Other	11883	10083	9734	9831	10370
Unaccounted	3225	1827	1905	2637	2816
World Total	31455	31211	29064	29817	31483

Source: USDA Estimates. Foreign Agricultural Service.
*2010 number is USDA Forecast.

1.2 COUNTRIES AND PRODUCTS

This section begins with a brief background on the market destination of key exporting countries, concluding with the largest rice importing country, the Philippines. Although the import side of the market is not considered in this research, a background on the Philippines as a key player can help the reader to understand the landscape of global rice trade.

1.2.1 Thailand

Thailand is the largest rice exporter in the world in terms of trade volume. Between 2005/06 and 2009/10, the USDA estimates that Thailand transferred around 48% of its rice production to exports. Based on information compiled from the Thai Rice Exporters Association (Statistics 2008, 2010), in 2008, Thailand shipped 25% of its exportable rice, around 10 million tons, to destinations in Asia, 46% to Africa, 14% to the Middle East, 8% to Europe, 6% to the Americas, and 2% to Oceanic countries. Thai rice export volume is comprised of 25% high quality rice, mostly long-grain *Aromatic* “Hom Mali Rice” or Thai Jasmine rice yet this generates about 30% of export revenue, and 75% long-grain *Indica* white rice, making up 70% of export revenue. Premium Thai Jasmine rice is popular in high income markets such as Hong Kong, Malaysia, Singapore, the United States, the EU, and the Middle East. Thai white rice finds markets in African countries, the Philippines, Indonesia, and China (see Prasertsri, 2009).

1.2.2 Vietnam

Vietnam is the second largest rice exporter in the world in terms of volume. From 2005/06 to 2009/10, the USDA estimates that Vietnam transferred 22% of its production to exports. In 2008, Vietnam shipped 52% of its exportable 4.6 million tons of long-grain *Indica* white rice to destinations in Asia, 27% to Africa, 5% to the Middle East, 3% to Europe, and 11% to the Americas. The Vietnamese rice price is traditionally lower than that for Thai rice, making Vietnam rice popular among lower income markets, such as African countries, the Philippines, Malaysia, Bangladesh, and Cuba. Vietnam also has a variant of fragrant rice, but the export volume of Vietnamese Jasmine rice is not significant to its share of total rice export volume (see Quan, 2009)

1.2.3 United States[2]

The United States is the fourth largest rice exporter in terms of volume, after Pakistan. From 2005/06 to 2009/10, the USDA estimates that the United States transferred 48% of its production to exports. The United States produces long-grain *Indica*, medium and short-grain *Japonica* rice with 60%, 35%, and 5% of weight-production share respectively. Each southern states, namely Arkansas, Louisiana, Mississippi, Missouri, and Texas, produced between 11.7 and 86 million *cwt* of long-grain rice in 2008/09, with Arkansas alone producing more than 50% of total long-grain rice in the United States. California is the largest state in terms of producing *Japonica* medium and short-grain rice, with a more than 80% share of total medium-grain rice in the United States. While the yield of the United States rice farmers is highest among top exporters (see figure 6), exports of United States milled rice have been slowly declining due to fierce price competition from lower cost Asian exporters. Nonetheless, the decline in milled rice exports has been compensated for by the rise in exports of rough rice, which has allowed the United States to reduce its lost share of rice exports (see Childs & Burdett, 2000). While Thailand and Vietnam only export rice in a milled-form, but in a rough rice arena, the United States is a

---

[2]Due to the limited data availability, India as the third largest rice exporter is excluded from the study.

dominant player. Latin America countries prefer to import rough rice from the United States to utilize their excess milling capacities since few other countries allow exports of rough rice, and no other countries can export significance volume of rough rice to match that of the United States (see Cramer, Wailes, Chavez, & Hanzen, 1998). Since 1998, rough rice has accounted for more than 30% of the US rice export mix, with Mexico and Central America as top export destinations. The United States was also able to export rough rice to Brazil, Ecuador, and Colombia when there was a need to boost the level of domestic stocks.

In response to lower cost Asian rice producers, the Marketing Loan program, part of the Commodity Loan program[3] was introduced for rice farmers in 1986 under provisions of the 1985 U.S. Farm Act. The Rice Marketing Loan program was designed to offer two loan repayment levels, based on market price above or below an adjusted world price. When market price is above the adjusted world price, the normal loan rate comes into effect, and when market price is below the adjusted world price, the lower loan repayment rate kicks in. Therefore, the benefits to U.S. rice farmers when market price is below the adjusted world price are 1) the difference between the loan rate and loan repayment, called a marketing loan gain, and 2) waiver of all accrued interest on the loan. (see Westcott & Price, 2001).

Consequently, this program requires USDA to calculate the fair world price. While the core component of the adjusted world price reflects major milled rice prices in major export markets, the other component also includes the effects of supply-

---

[3] Commodity loans may be settled in three ways: 1)Repaying at the loan rate plus interest costs (CCC interest cost of borrowing from the U.S. Treasury plus 1 percentage point), 2)Repaying at an alternative loan repayment rate, or 3) Forfeiting the pledged crop to the CCC at loan maturity.

demand changes, government-assisted sales, and other relevant price indicators, which adds significant element of judgment into the world price formula (see Westcott & Price, 2001). In this regard, Taylor, Bessler, Waller, and Rister (1996) argued that lack of transparency in the formula of the USDA adjusted world price makes it susceptible to error in judgment.

1.2.4 Philippines

The Philippines is the largest rice importer based on volume. This is despite the fact that rice production in the Philippines ranked eighth in the world. From 1995 onward, the Philippines' rice consumption continued outpacing domestic production increases. From 2005/06 to 2009/10, the USDA estimates that the Philippines imported on average 17% of its consumption. This crucial dependency on rice imports puts the Philippines at the top of list of rice importing countries, with an 8% share of the total international rice trade. Being the largest rice buyer in the international market, the Philippines regularly purchases rice through auctions or a government to government (G to G) tender offer. The rice auction operation is carried out by the rice bidding committee at the National Food Authority (NFA). The rice stock level of Philippines is closely watched by rice traders as an indicator of the probability of imminent rice imports which would significantly affect international rice prices.

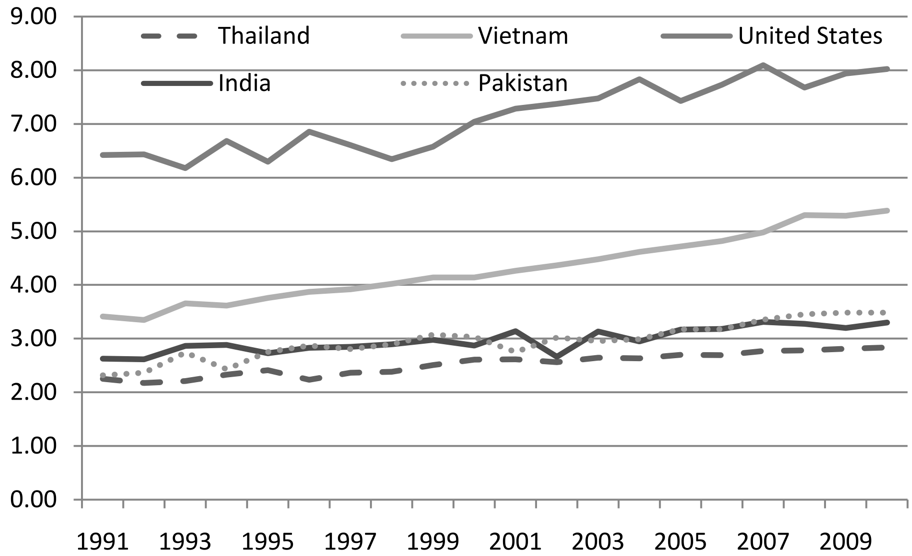


Figure 6 Rice Yield (rough basis) metric ton per hectare

Source: Data from the Foreign Agriculture Service and Author's calculation.

1.3 MARKET DYNAMICS

1.3.1 Domestic policy, trade policy

Out of the top five exporting countries, only the United States is an advanced economy while Thailand, Vietnam, Pakistan, and India are emerging economies. Levels of poverty and low income population within these countries certainly affect domestic policy where food price is concerned (see Timmer & Dawe, 2007). As flooding and drought, brought on by climate change, continue to put pressure on rice production, the threat of exporting bans on rice to calm domestic pressure is real. Together with increasingly scarcer land and water supplies to support rice production, the probability of increasing rice price volatility is high.

1.3.2 Genetically Modified rice

Presently, only the United States, Canada, Argentina, China and Brazil allow the widespread planting and commerce of the Genetically Modified (GM) rice. There is also speculation that the Philippines, India, and Bangladesh might soon follow China in accepting GM rice in Asia (see Serapio, 2010). Thailand remains cautious about GM rice and has completely banned GM rice research. Vietnam allowed GM rice research in 2001, but restricted it to laboratory-stage only (see Rice, 2010).

Proponents of GM rice argue that the GM rice improves yields and can solve malnutrition problems for the world's population (see The Golden Rice Project, 2010). Opponents argue that, with GM rice, there could be a hidden harmful component that, once unleashed in wide-scale planting, would be irreversible (see Hands Off Our Rice, 2010). Resistance from Europe and Middle Eastern countries toward GM rice will limit its growth in these high income markets. Despite this setback, GM rice continues to grow in lower income countries where malnutrition problems are greater (see Genetically Modified Rice, 2008).

2 SURVEY OF LITERATURE ON SPATIAL MARKET EFFICIENCY

This section describes a brief history of cointegration in a spatial market in the literature and ends with a definition of market efficiency. Based on the pioneering work of Engle and Granger (1987) on the concept of cointegration, which demonstrates that non-stationary variables are cointegrated if there exists a linear combination among them that is stationary, cointegration analysis finds popular use in many applications that seek the long-term equilibrium in dynamic settings, such as studies of securities price (see Hasbrouck, 2003), as well as convergence of prices in the future and the cash market (see Figuerola-Ferretti and Gonzalo, 2008). The cointegration concept also finds application in spatial market studies, in which its common use is to test the hypothesis of the Law of One Price in a spatial setting and to measure the degree of ‘integration’ or ‘efficiency’ (see Fackler and Goodwin, 2001 for more discussion).

Criticism of cointegration analysis spatial markets sees the test of market efficiency as indistinguishable from tests of the assumption underlying the model itself. In essence, the rejection of the model does not imply inefficiency of spatial markets, but may imply flawed assumptions underlying the error-correction model instead (see McNew and Fackler, 1997, Barrett, 2001). However, evidence for the efficient role of markets continues to emerge. Dawson and Dey (2002) analyzed rice market pairs in Bangladesh using an error-correction model and found results supportive of market integration. Rashid (2004) employed the same framework to analyze maize in Uganda and found evidence in favor of market efficiency. Thompson, Sul, and Bohl (2002) used the Seemingly Unrelated Regression error-correction model on wheat markets in the UK, France, and Germany and found evidence that conformed to the law of one price as well as efficient markets.

While earlier spatial market literature implies that price cointegration is evidence of market integration, Barrett (2001) argued that cointegration analysis based on price information alone cannot be used to infer market integration, that, in the absence of trade flow information, a price-based notion of market equilibrium or price cointegration can only suggest market efficiency. According to his definition, market efficiency is “the satisfaction of zero marginal benefit equilibrium conditions. It is therefore a statement about welfare, about whether there exist potential Pareto improvements in the (international) economy". In this sense, the market is highly efficient when there are few or no Pareto improvement opportunities.

The definition of market efficiency among rice-exporting countries in this study is based on observations of co-movement of the price in the cointegrating equation, specifically the price transmission elasticity. To the extent that two spatial prices show tight co-movement, that price shocks from one market are able to transmit to the other market efficiently, the windows to arbitrage would be minimal, and we can infer high efficiency between the two markets. In contrast, if two spatial prices demonstrate diverging trends, then it would suggest a possible Pareto improvement opportunity, and an indication of a low level of market efficiency as well as evidence in favor of market segmentation.

On a more cautious note, since this framework heavily relies on price information, but the lack of transportation and transaction costs, and more importantly, trade flow and the effect of domestic policies which can often be unpredictable, and difficult to distinguish, it is prudent that the interpretation of the estimated co-movement must be taken with consideration of these factors as well.

Finally, the main contributions of this paper are twofold. First, to the best of my knowledge, this is the first time the price discovery concept has been addressed in the context of the global rice trade. Secondly, policymakers often lack tools to gauge the impact of international rice price adjustment. The estimated price transmission elasticities from this research provide a preliminary assessment tool to both market participants and any governmental supervisory body.

3 HYPOTHESES

This section formalizes the questions this study attempts to answer.

Hypothesis 1: *There exists an efficient long-run price relationship among the Thai, Vietnamese, and US rice export market.*

The first hypothesis considers whether there exists a long-run price relationship among Thai, Vietnamese, and US (Arkansas) rice, and conditional on the existence of the linkages, the hypothesis of efficient market will be tested. The existence of a long-run price relationship is a prerequisite to Gonzalo and Granger decomposition which measures the degree of price discovery in each pair of markets (see Gonzalo and Granger, 1995). In contrast, if the existence of a long-run price relationship cannot be determined, one plausible explanation is a high degree of market segmentation, and that market-specific factors play a more dominant role in determining price movement within each markets.

Hypothesis 2: *The USDA adjusted world price reflects a long-run price relationship with Thai, Vietnamese, and US rice.*

This hypothesis considers whether the performance of the USDA world price truly reflects the actual rice prices in leading rice exporting countries. The verification of a long-run price relationship between the USDA adjusted world price and prices for other major rice exporting countries will ease policy observers' fear of USDA price manipulation through Marketing Loan Program and provides empirical justification to enhance its program's transparency (see discussions in Taylor, Bessler, Waller, and Rister, 1996).

Hypothesis 3: *There exists a convergence between spot and future markets of rough rice in the US market.*

The third hypothesis considers whether there exists an efficient long-run price relationship between cash and future markets for rough rice in US rice markets. Similar to the first hypothesis, the existence of a long-run price relationship is a precondition for market efficiency test. In the context of future and spot price, market efficiency implies that future market mechanism functions as a source of price discovery and that two prices converge in the long run. Conversely, if the result fails to comply with this prediction, the market mechanism maybe inefficient and its function should be examined further.

Hypothesis 4: *There exists a long-run price relationship between U.S. rough and milled rice.*

The final hypothesis considers whether there exists a price linkage between two marketing levels, rough and milled rice respectively. A proof of a cointegration signifies a price transmission channel in which price innovation travel from one to another marketing level. Conditional on the valid cointegration, the price discovery measurement reveals the source of information flow between the two marketing levels.

4 EMPIRICAL STRATEGY

The section describes the construction of stationary time series model and relevant statistical tests.

4.1 UNIT ROOT TESTS

Stationarity of variables is the most important concern in modern time series estimation. To determine if price and first differences are stationary I(0). In this study, the modified augmented Dicky-Fuller test (MADF) is considered instead of the augmented Dicky-Fuller (ADF). In essence, the MADF test applies the implementation of the ADF test on the price series that is transformed by a generalized least squares regression. The MADF test demonstrates greater power and more consistent results than the standard Augmented Dicky-Fuller (ADF) in rejecting the null hypothesis of unit root (see Elliott, Rothenberg, and Stock, 1996). The method for selecting lag is the minimum Schwarz information criterion (SIC).

4.2 ERROR CORECTION SPECIFICATION

Consider a price evolution in a vector autoregressive form

$$y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t \quad (1)$$

$$t = 1, \dots, T$$

where y_t is a $K \times 1$ vector of the logarithm of rice prices. v is a $K \times 1$ vector of constants, $A_1 \dots A_p$ are $K \times K$ matrices of parameters. And ε_t is a $K \times 1$ vector of error terms. ε_t has zero mean and variance-covariance matrix Σ, and is i.i.d. K is the number of markets in the consideration.

Rewriting (1) in an error correction representation yields the following

$$\Delta y_t = v + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \epsilon_t \quad (2)$$

$t = 1, \dots, T$

where $\Pi = \sum_{j=1}^{j=p} A_j - I_p$ and $\Gamma_i = -\sum_{j=i+1}^{j=p} A_j$.

Γ_i are $K \times K$ parameter matrices which contain short-run dynamics with Δy_{t-i}. Π is a $K \times K$ parameter matrix containing adjustment to the long-run economic relationships among the element of y_t, and is the main focus of our analysis.

If variables y_t are non-stationary, or $y_t \sim I(1)$, but stationary at a first difference, or $\Delta y_t \sim I(0)$, then, by the Granger representation theorem (Engle & Granger, 1987), Π has rank $0 \leq r < K$, where r is the number of linearly independent cointegrating vectors. The main assumption underpinning VECM is the conjecture that the system of variables should converge to their long term equilibrium, while long-term equilibrium is implied by cointegration.

There are three cases involving estimation of Π. First, if Π has a rank of zero, there is no long term equilibrium among variables in y_t and the estimation in the form of (2), but without the long-run parameters, is sufficient. And, if Π has rank of K, or full rank, then all elements of y_t are stationary and the estimation in the form of (1) is sufficient. Finally, if Π has reduced rank r where $0 < r < K$, then Π can be decomposed into a loading matrix α, and cointegrating equation β where $\Pi = \alpha\beta$. In this case, both α, β' are a $K \times r$ matrix of rank r.

In the case of two markets or a bivariate model $K = 2$, if rank $r = 1$, then there is a long-run price relationship between the two markets represented by cointegration. However, if rank $r = 0$, then there is no long-run relationship between the two prices. The procedure to determine rank r is given in the next section.

Rearranging terms in (2) gives

$$\Delta y_t = \alpha\beta y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \upsilon + \epsilon_t \qquad (3)$$

where υ is a $K \times 1$ vector of parameters for a linear time trend at y_t.

The deterministic trends υ above can be decomposed into the long-run cointegrating equation and a constant.

$$\upsilon = \alpha\mu + \gamma, \qquad (4)$$

where $$\gamma'\alpha\mu = 0 \qquad (5)$$

and μ is $r \times 1$ vectors of parameters.

Rearranging terms in (3) with (4) yields

$$\Delta y_t = \alpha(\beta y_{t-1} + \mu) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \gamma + \epsilon_t. \qquad (6)$$

4.3 TESTS FOR COINTEGRATION AND ERROR CORRECTION

The Johansen's procedure to determine whether there is a long run relationship among variables in y_t is followed. Under the maximum likelihood estimation, if the log likelihood of the unconstrained model that includes the cointegrating equations is significantly different from the log likelihood of the constrained model, which excludes cointegrating equations, the null hypothesis of no cointegration is rejected. The procedure always starts with the case of no cointegrating equations, or rank of zero, and accepts the first null hypothesis, which is non-rejection.

By design, this study focuses on bivariate systems; thus, there are always two cases for the null hypotheses for each system. The first null hypothesis is H(0) where there is no cointegrating equation. Second null hypothesis is H(1) where there is only one cointegrating equation. The lags' order (p) in each of the bivariate systems are chosen based on the Akaike Information Criterion (AIC). In small samples, the AIC may have better properties (choose the correct order more often) than the Hannan-Quinn Information Criterion (HQIC) and Schwartz-Baysian Criterion (SBIC), as they are designed for minimizing forecast error variance (see Lutkepohl, 2005).

4.4 COINTEGRATING EQUATIONS

The cointegrating vector β reveals the long-run relationship between the price pairs. Together with y_t and the constant term, the long-run price relationship is represented by $\beta y_t + \mu = \beta_1 lnP_1 + \beta_2 lnP_2 + \mu$. The cointegrating vector of each of markets' pair is displayed in the form of $(1, -\beta_2)$ where the coefficient of the first market is normalized to one, β_2 is the coefficient of the second corresponding market, and μ is a constant term that predicts variable costs related to the trade that would make the price in two markets arbitrage-free in the long-run. Since the model is based on the logarithm of prices, β_2 tell us the long-run percentage change in the rice price in the first market when the price in the second market changes by one percent. This is usually referred to as *"price transmission elasticity"* in agricultural market literature. To demonstrate a high level of market efficiency, β_2 should yield a number close to one, indicating that price shock from one market is able to transmit to the other market quickly, making the co-movement of prices of the same commodities close, albeit at different locations. Hence, the deviation from one of the $\beta_2$will allow us to better understand better the nature of rice market segmentation.

On the other hand, the cointegrating vector of the United States' spot-futures pair should also yield $(1, -1)'$ to demonstrate long-run convergence of spot and futures markets. While the conjecture on the cointegrating equation between milled and rough rice is less clear, the estimated parameter should reflect the level of marketing markups. Underlying these views is the strong assumption that transaction and transportation costs are stationary, as well as proportional to the price and would be captured within μ.

4.5 PRICE DISCOVERY

The concept of price discovery generally refers to the process of discovering an asset's full information or its permanent value (see Figuerola-Ferretti and Gonzalo, 2008). To observe the degree of price discovery process in each price series, since the fundamental value of the stock or commodity is unobservable, the observable price must be broken down into two pieces, the core fundamental value and its transitory component. To this end, the use of Gonzalo and Granger's (1995) permanent transitory decomposition technique has gained traction in mainstream price discovery literature. Among the first were Harris, McInish, and Wood (1997 and 2002). Among many desirable properties, this decomposition technique allows us to project both the permanent component $(W_t)$and the transitory component (Z_t) on a linear combination of the original variables, y_t, which is observable. Other attractive features of this method, as opposed to the competing price discovery technique (see Hasbrouck, 2003) for example, are that the result of decomposition is unique and valid even when β_2 in the cointegrating vector $(1, -\beta_2)$ is not 1.

Based on the Granger and Gonzalo definition of decomposition, the permanent component is fully identified as

$$W_t = \alpha_\perp y_t \tag{9}$$

where $\alpha_\perp$ is an orthogonal vector of the α vector, or $\alpha_\perp{}' \propto = 0$,

and the transitory component is

$$Z_t = \beta y_t, \tag{10}$$

where β is the cointegrating vector.

Then, the Gonzalo and Granger decomposition of y_t becomes

$$y_t = A_1 W_t + A_2 Z_t \tag{11}$$

$$y_t = A_1 \alpha_\perp y_{t_t} + A_2 \beta y_t \tag{12}$$

where $A_1 = \beta_\perp(\alpha_\perp{}'\beta_\perp)^{-1}$,

$A_2 = \alpha(\beta'\alpha)^{-1}$,

with $\alpha_\perp{}' \propto = 0$,

$\beta_\perp{}'\beta = 0$.

To estimate the proportion of permanent value for each market, or price discovery, the focus is on the $\alpha_\perp$ vector. Baillie, Booth, Tse, and Zabotina (2002) show that, when used in the context of a bivariate system, the permanent share of each market in the long term price equilibrium can be obtained from the formula

$$\alpha_\perp = \begin{bmatrix} \alpha_{\perp 1} \\ \alpha_{\perp 2} \end{bmatrix} = \begin{bmatrix} \left(\frac{\alpha_2}{\alpha_2 - \alpha_1}\right) \\ \left(\frac{-\alpha_1}{\alpha_2 - \alpha_1}\right) \end{bmatrix}, \tag{13}$$

where $\alpha_{\perp 1}$ represents the proportion of long-run price discovery associated with the first market and vice versa. Based on this formula, statistical testing on price discovery can be done via the test of significance on the α vector itself. For more discussion of comparing price discovery measures, please see Yan & Zivot (2007).

5 DATA

The data includes logarithms of monthly rice prices, where obtained from the United States Department of Agriculture (USDA Economic Research Service, 2010) and a privately obtained CEIC data subscription.[4] The USDA Rice Yearbook collects monthly rice prices on various markets both within the US and in other major rice exporting countries. The price series for CBOT rough rice futures is obtained from the CEIC data subscription. A reference price for the U.S. market is the *Indica* long-grain number two, based on the Arkansas market as production volume of long-grain rice in Arkansas is the largest in the United States. The U.S. long-grain number two is also the underlying commodity for rough rice futures for the Chicago Board of Trade. The technical specification of the U.S. long-grain number two is conveniently comparable to the Thai 100% grade B which is the global benchmark for rice prices. The rice price for the Vietnamese market is referred to as 5% broken double washed. Although Vietnamese 5% broken double washed is seen as slightly lower quality and not directly comparable to Thai rice, this is the highest quality of Vietnamese rice for which data is available. Both Thai 100% Grade B and U.S. long-grain number two are considered high standard, and are equivalent (see Bashir, 2002, Kaosa-ard and Juliano, 1990). All three are *Indica* long-grain rice.

Finally, the adjusted world price calculated by USDA, based on its internal formula, is also obtained through the USDA Rice Yearbook. Change in the USDA

[4] See www.ceic.com for information on the subscription

adjusted world price affects US rice farmers' benefits and hence their decisions on farms' operation.

For the remainder of this paper, Thai 100% Grade B will be referred to as 'Thai rice', Vietnamese 5% broken will be referred to as 'Vietnamese rice,' and Arkansas long-grain US number two broken not exceeding 4% will be referred to as 'Arkansas rice.'

The data covers price series from January 1998 to February 2010. The Thai, Vietnamese, and Arkansas prices are freight-on-board or F.O.B. pricing.
The US rough rice price series on the cash market is the average price received by US rice farmers. All prices are quoted in current US dollars.

Table 5 Summary statistics for the logarithm of rough and milled rice prices 1998:1 - 2010:2

(Logarithm of)	Obs.	Mean	Std. Dev	Min.	Max.
Thai price	146	5.67	0.406	5.13	6.85
Vietnamese price	146	5.56	0.371	5.01	6.98
Arkansas price	146	5.89	0.369	5.24	6.84
Adjusted world price	146	5.37	0.475	4.64	6.46
CBOT Rough Rice Futures	146	5.15	0.424	4.38	6.18
USDA Rough Rice price	146	5.16	0.416	4.43	6.08

Source: Author's calculation

5.1 UNIT ROOT TESTS

This section gives a standard stationarity analysis preceding the formulation of a vector error correction model. To begin, four of the price series are tested for unit root. A Modified Augmented Dicky-Fuller (MADF) test has been implemented. To control for the effect of the government-imposed export restriction in Vietnamese rice market, the original price series were regressed on the dummy variable that indicates the export restriction during March 2008 – June 2008. The predicted residuals of the series were then used in the MADF. The lag order of each price series is selected by Schwartz Criterion. Table 6 shows that the null hypotheses of unit roots could not be rejected on all price series at level but were rejected on the first difference on all series. Thus, based on the MADF test, the logarithm of prices is non-stationary at level or I(1) and stationary at the first difference or I(0).

5.2 JOHANSEN'S RANK TEST

The first step in estimating a bivariate system in equation (6) is to run a Johansen's cointegration rank test to determine whether there is any long-run relationship among variables of interest. To control for seasonality, a set of centered seasonal dummies is added. This set of dummies is summed to zero over a year so that linear terms from the dummies disappear and are replaced with a constant term.

Table 7 presents results with evidence for strong support for a long-run relationship between Thai and Vietnamese rice prices. Johansen's Rank Test rejects the null hypothesis of no cointegration and fails to reject the hypothesis of one cointegrating equation.

Table 6 Unit Root Tests

Time period (1998:1-2010:2)	Modified Augmented Dicky-Fuller Test Statistic	
	Level	First Difference
Thai	-1.08	-2.87**
Vietnam	-0.71	-2.11**
Arkansas	-1.44	-4.48**
World Milled Price	-0.65	-2.31*
CBOT Rough Rice Future	-0.97	-6.26**
USDA Rough Rice	-0.66	-4.81**

H0: Unit Root. *, and ** denote 5% and 1% significance Level. The MADF test applies the implementation of the ADF test on the price series that is transformed by a generalized least squares regression. To control for the effect of the government-imposed export restriction in Vietnamese rice market, the original prices were regressed on the dummy variable that indicates the export restriction during March 2008 – June 2008. The predicted residuals were then used in the MADF. Lag orders for each series are selected by Schwartz Criterion.

Table 7 Johansen Cointegration Tests

Time period (1998:1-2010:2)	Linear Trend	
	H(0)	H(1)
Thai - Vietnam	42.32**	0.25
Thai - Arkansas	15.99*	0.66
Vietnam - Arkansas	34.64**	1.14
Thai - WMP	10.00	0.33
Vietnam - WMP	33.54**	0.64
Arkansas - WMP	12.51	1.00
CBOT Rice Future - USDA Rough Rice	36.81**	0.61
Arkansas Milled - USDA Rough Rice	33.95**	0.80

H(0): No cointegration, H(1): 1 cointegrating equation. The test specification assumes a constant term in VECM. Lag orders in each system are selected by the Akaike Info Criterion (Lutkepohl, 2005). *, and ** denote 5% and 1% significance level.

Johansen's Rank test also rejects the null hypothesis of no cointegration and fails to reject the hypothesis of one cointegrating equation between Thai and Arkansas rice prices as well as between Vietnamese and Arkansas rice prices. Strong evidence of price linkages among all major exporters seems to suggest that long-run price shocks could transmit from one region to another. More detailed analysis of the magnitude of these linkages will be presented and discussed in the next section.

As to questions relating to the validity of the USDA world market reference price among major exporters, the first weakness to the argument is seen when the cointegrating rank test finds only one linkage between the USDA World price and the Vietnamese market. The tests do not find any evidence of cointegration between the USDA World price and either the Thai or Arkansas market. While this result cannot be used as an indication of an invalid formula for the true world price, it does cast doubt over its performance.

Within U.S. domestic rice markets, for the spot and futures of rough rice, Johansen's test finds strong evidence to support a long-run linkage between the cash market and rough rice futures. The test also finds that Arkansas's milled price and the USDA cash rough rice price are cointegrated.

Results from Johansen's rank tests allow us to formally decompose the long-run parameter matrix π into α and β vector. This leaves to the next section detailed analysis of the estimated value of the α and β vectors, which will give us a clue to the role of price discovery in each market, and its long-term co-movement.

6 RESULTS

The subsection 6.1 comprises three pairs of estimated price relationships, namely Thai-Vietnamese markets, Vietnamese-US (Arkansas) markets, and Thai-US (Arkansas) markets. The second subsection analyzes an estimated price relationship between the adjusted world price and the Vietnamese rice price. The third subsection discusses a result from estimated price relationship between spot and futures of US rough rice. The final subsection includes an analysis from estimated price relationships between the spot price of rough rice and Arkansas milled rice price to demonstrate vertical market efficiency. Recall that all price relationships are estimated in the following form.

$$\Delta y_t = \alpha(\beta y_{t-1} + \mu) + \sum_{i=1}^{p-1} \Gamma_i \, \Delta y_{t-i} + \gamma + \epsilon_t \tag{14}$$

6.1 THREE LEADING RICE EXPORTING COUNTRIES

The results of the VECM estimation on the treated data of Thai-Vietnamese and US (Arkansas) -Vietnamese markets from 1998m1 to 2010m2 are shown in (15) and (16) respectively. To control for the effect of the government-imposed export restriction in Vietnamese rice market, the original price series were regressed on the dummy variable that indicates the export restriction during March 2008 – June 2008. The predicted residuals of the series were then used in the VECM estimation.

$$\begin{bmatrix}\Delta TH_t\\ \Delta VN_t\end{bmatrix} = \begin{bmatrix}-.0136\\ (.0368)\\ .4522^{**}\\ (.0670)\end{bmatrix}\left[TH_{t-1} - \underset{(.0548)}{1.300^{**}} VN_{t-1} — 5.662\right]$$

$$+\begin{bmatrix}.6038^{**} & -.2014\\ (.1511) & (.1650)\\ -.6027^{**} & .2221^{**}\\ (.1726) & (.0807)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-1}\\ \Delta VN_{t-1}\end{bmatrix} + \begin{bmatrix}-.259^{**} & .0992^{*}\\ (.1000) & (.0419)\\ -.0821 & .0921\\ (.1823) & (.0763)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-2}\\ \Delta VN_{t-2}\end{bmatrix}$$

$$+\begin{bmatrix}.1054 & .0092\\ (.0985) & (.0428)\\ .1309 & .1836^{**}\\ (.1795) & (.0781)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-3}\\ \Delta VN_{t-3}\end{bmatrix} + \begin{bmatrix}-.0150^{*}\\ (.0135)\\ .0151\\ (.0246)\end{bmatrix}[seasonal\ dummies]$$

$$+\begin{bmatrix}.0032\\ (.0041)\\ .0001\\ (.0076)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{TH}\\ \hat{\epsilon}_{VN}\end{bmatrix} \qquad (15)$$ [5]

$$\begin{bmatrix}\Delta \mathrm{AK_t}\\ \Delta VN_t\end{bmatrix} = \begin{bmatrix}-.0168\\ (.0136)\\ .2639\\ (.0452)\end{bmatrix}\left[AK_{t-1} - \underset{(.1012)}{1.1880^{**}} VN_{t-1} - 5.881\right]$$

$$+\begin{bmatrix}.7280^{**} & -.0732^{**}\\ (.0665) & (.0234)\\ -.2395 & .1807^{*}\\ (.2210) & (.0778)\end{bmatrix}\begin{bmatrix}\Delta AK_{t-1}\\ \Delta VN_{t-1}\end{bmatrix} + \begin{bmatrix}-.0051\\ (.0081)\\ .0264\\ (.0267)\end{bmatrix}[seasonal\ dummies]$$

$$+\begin{bmatrix}.0010\\ (.0026)\\ .0001\\ (.0088)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{AK}\\ \hat{\epsilon}_{VN}\end{bmatrix} \qquad (16)$$

Next, the estimated bivariate error correction model between Thai and Arkansas rice price is the following:

---

[5] * and,** denote 5% and 1% significance level. Standard errors are given in parentheses. The coefficient of the first variable in the cointegrating vector was normalized to one and thus does not have standard errors. The constant term in the cointegrating equation does not have standard errors since it is not directly estimated but is backed out from other estimates.

$$\begin{bmatrix}\Delta TH_t \\ \Delta AK_t\end{bmatrix} = \begin{bmatrix}-.0368 \\ (.0249) \\ .0384^{**} \\ (.0146)\end{bmatrix}\left[TH_{t-1} - \frac{1.2930^{**}}{(.1298)}AK_{t-1} + 1.9043\right]$$

$$+\begin{bmatrix}.4246^{**} & .3468^{*} \\ (.0929) & (.1526) \\ .1191^{*} & .5848^{**} \\ (.0544) & (.0893)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-1} \\ \Delta AK_{t-1}\end{bmatrix} + \begin{bmatrix}-.1967^{*} & -.1274 \\ (.0964) & (.1742) \\ -.0638 & -.0430 \\ (.0564) & (.1019)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-2} \\ \Delta AK_{t-2}\end{bmatrix}$$

$$+\begin{bmatrix}-.06889 & -.0328 \\ (.0919) & (.1512) \\ -.0226 & .2169^{*} \\ (.0537) & (.0885)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-3} \\ \Delta AK_{t-3}\end{bmatrix} + \begin{bmatrix}-.0234 \\ (.0146) \\ -.0016 \\ (.0085)\end{bmatrix}[seasonal\ dummy]$$

$$+\begin{bmatrix}.0018 \\ (.0146) \\ .0019 \\ (.0027)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{TH} \\ \hat{\epsilon}_{AK}\end{bmatrix} \qquad (17)$$

From (15), the cointegrating equation , or β vector, between Thai and Vietnamese rice price shows a coefficient of -1.30 on Vietnamese rice price with a 95% confidence interval between -1.41 and -1.19. The cointegrating equation above extracts the long-run co-movement of Thai and Vietnamese rice prices[6]. Since the absolute value of the lower range, 1.19, is larger than 1, we can reject the hypothesis of market efficiency between Thai and Vietnamese rice, and agree that the long-run percentage change in Thai rice price is slightly higher than that for Vietnam. The narrowed range of the estimate also gives us some comfort in terms of precision of magnitude.

---

[6] From the cointegrating equation $TH = 1.30VN - 5.66 + \hat{\varepsilon}$, we have
$lnP_{TH} = 1.30lnP_{VN} - 5.66 + \hat{\varepsilon}$

$$dlnP_{TH} / dlnP_{VN} = 1.30$$

$$\frac{dP_{TH}}{P_{TH}} \Big/ \frac{dP_{VN}}{P_{VN}} = 1.30$$

Figure 7 depicts the plot of the predicted cointegrating vector: a relatively stable graph, fluctuating in closed range, until the first half of 2008 when a large swing in the system occurred. The cointegration vector returns to the same but slightly elevated pattern of late 2008. Since the nature of shock is brief and temporary, one plausible explanation is a price-overshooting in the presence of supply-shock. Early 2008 was a time when many rice exporting countries decided to restrict their exports out of fear that escalating rice prices would create domestic turmoil. India restricted exports of non-basmati rice in February 2008 and Vietnam placed a ban on rice exports in March 2008.[7] This occurred while Thailand allowed business to go on as usual. Altogether, these decisions effectively pulled available tradable supply from the global market, causing a sharp deviation from the long-term co-movement trend. Although this non-market intervention was accounted for by an OLS regression on the dummy variable of the government intervention but the effect of the short-term price overshooting remained visible.

The Gonzalo and Granger (1995) framework allow us to extract more information from the alpha vector. The decomposition method was simplified for the case of bivariate systems by Baillie, Booth, Tse, and Zabotina (2002). Their simple formula, $\alpha_{\perp 1} = \frac{\alpha_2}{\alpha_2 - \alpha_1}, \alpha_{\perp 2} = \frac{-\alpha_1}{\alpha_2 - \alpha_1}$, computes the proportion of long-run fundamental value associated with each respective markets. Based on this formula, the proportion of long-run price discovery for the Thai market is 97% versus 3% for the Vietnamese market.[8] The p-value for the test of statistical significance for the price discovery measure is 0.00, and 0.71 for Thai and Vietnamese rice market

[7] See Appendix

[8] $\frac{.4522}{(.4522+.0136)} = 97\%, \frac{.0136}{(.4522+.0136)} = 3\%$

respectively. The result suggests that higher proportion of the price discovery associated with the Thai market pinpoints its dominance in information source compared to the Vietnamese market.

From (16), Arkansas and Vietnamese rice prices show a cointegrating relationship with the coefficient of -1 for Vietnam, and a 95% confidence level between -1.38 and -0.98. Since the absolute value of the range covers 1, we cannot reject the hypothesis that the long-run co-movement between Arkansas and Vietnamese rice prices is tightly integrated or equal to one. The estimated cointegration yields an interesting view that these two markets movements are roughly the same in the long run.[9]

The Gonzalo and Granger decomposition method shows that the proportion of long-run price discovery for the Arkansas market is 94%, versus 6% for the Vietnamese market.[10] The p-values associated with Arkansas and Vietnamese market's price discovery are 0.00 and 0.22 respectively. The higher proportion of the price discovery associated with the Arkansas market and its significance reveals its dominance in information source compared to the Vietnamese market.

---

[9] From the cointegrating equation $AK = 1.19VN - 5.88 + \hat{\varepsilon}$, we have

$$lnP_{AK} = 1.19lnP_{VN} - 5.88 + \hat{\varepsilon}$$

$$dlnP_{AK} / dlnP_{VN} = 1.19$$

$$\frac{dP_{AK}}{P_{AK}} \Big/ \frac{dP_{VN}}{P_{VN}} = 1.19$$

[10] $\frac{.2639}{(.2639+.0168)} = 94\%, \frac{.0168}{(.2639+.0168)} = 6\%$

From (17), the cointegrating equation between Thai and US (Arkansas) rice price shows a coefficient of -1.29 with 95% confidence interval between -1.55 and -1.04. The absolute value of -1.55 and -1.04 implies that in the long-run percentage change in Thai price is more volatile compared to that of Arkansas price.[11] As a result, market efficiency hypothesis is rejected between the Thai and US (Arkansas) markets. The relatively wider range of the interval suggests that while a long-run relationship may exist, there could be other driving forces outside this two market system for this long-run co-movement.

Figure 9 depicts the plot of the predicted cointegrating equation between Thai and Arkansas rice prices. The large variation in upswing and downswing between 2002 and 2004 is puzzling. This significant divergence from zero of the predicted cointegrating equation clearly suggests that there could be market-specific factors that drive deviation from the long-run co-movement. Over the longer term, however, the predicted cointegrating equation does return to the normal range, which suggests to us that shocks during 2002 and 2004 were temporary.

The Gonzalo & Granger decomposition method shows that the proportion of long-run price discovery for the Thai market is 51%, versus 49% for Arkansas market.[12] The p-values for the statistical test of significance for the price discovery

---

[11] From the cointegrating equation $TH = 1.29AK - 1.90 + \hat{\varepsilon}$, we have

$$lnP_{TH} = 1.29lnP_{AK} - 1.9 + \hat{\varepsilon}$$

$$dlnP_{TH} / dlnP_{AK} = 1.29$$

$$\frac{dP_{TH}}{P_{TH}} \Big/ \frac{dP_{AK}}{P_{AK}} = 1.29$$

[12] $\frac{0.0384}{(0.0368 + 0.0384)} = 51\%, \frac{0.0868}{(0.0368 + 0.0384)} = 49\%$

measure are 0.009 and 0.140 for Thai and Arkansas rice markets respectively. If we are to accept 0.15 significance level, this roughly similar proportion of permanent value associated with each market implies that both markets equally contribute to price discovery.

Table 8 Summary of Price Discovery among Leading Exporters

Time period (1998:1-2010:2)	Rice Market Price Discovery					
	TH	VN	AK	VN	TH	AK
Price Discovery	0.97	0.03	0.94	0.06	0.51	0.49
p-value	0.00	0.71	0.00	0.22	0.01	0.14

Source: Author's calculation

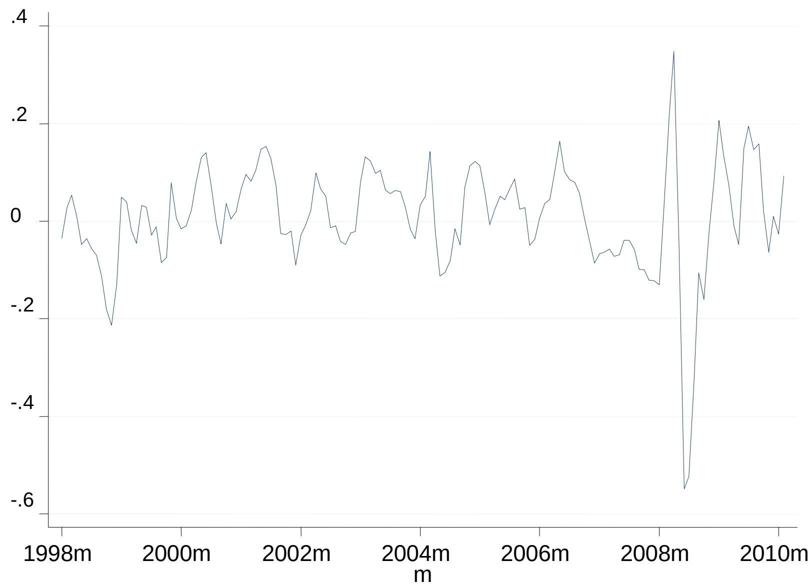


Figure 7 Predicted Cointegrated Equation between Thai and Vietnam rice markets

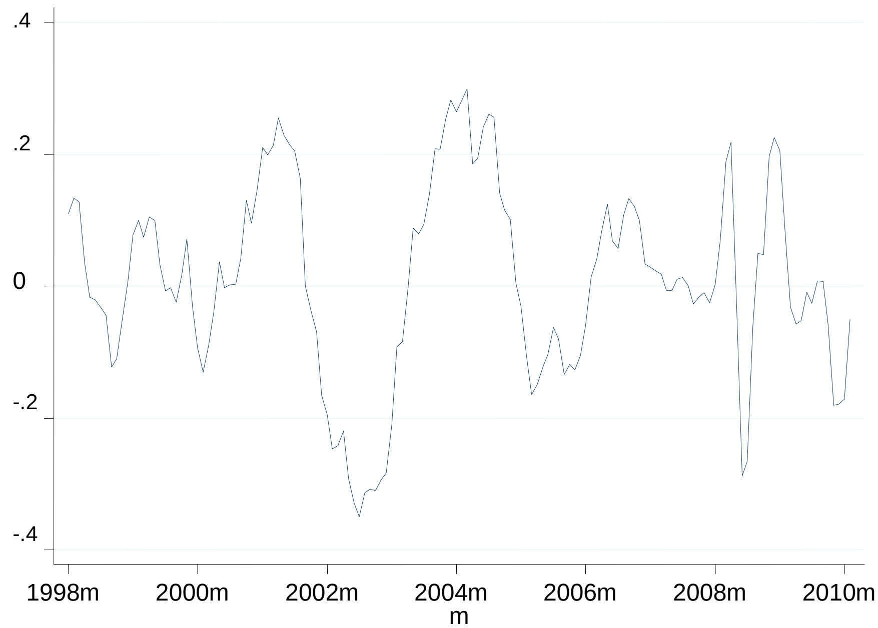


Figure 8 Predicted Cointegrated Equation between Arkansas and Vietnamese rice market

Figure 9 Predicted Cointegrated Equation between Thai and Arkansas rice market

To sum up, with respect to the first hypothesis which states "*There exists an efficient long-run price relationship among the Thai, Vietnamese, and US rice export market"*, the evidence above shows that there are long-run linkages among the three major rice exporters. This is a direct result from the fact that we are able to extract the long-run co-movement among the three pairs as shown by the price transmission elasticities. However, of the three cases, except for that between Vietnam and Arkansas, the market efficiency hypothesis is rejected for the two other cases, as the long-run co-movement represented by β_2 is significantly different from one. Building on these linkages, the finding shows the dominant sources of price discovery are Thailand and Arkansas, while the role of Vietnamese rice market appears to be informational laggard of the three.

The existence of long-run price linkages among the three rice-exporting countries is a sensible discovery as Thailand is the world's largest rice exporter while the United States dominates high-quality rice and high-income markets. Although the estimated price transmission elasticities do not comply with the market efficiency hypothesis, the analysis of cointegrating equation, with the aid of graphical plot, suggests that the sharp deviation from the cointegrating trend in 2008 could be the root cause of hypothesis's rejections. Accordingly, this view was reinforced by the result based on data prior to 2008 that saw non-rejection of the hypothesis of market efficiency on all three cases.[13]

[13] Results shown in Appendix

6.2 USDA ADJUSTED WORLD PRICE

In this subsection, the following analysis considers whether the characteristics of the USDA adjusted world price can be explained by the movement of one rice-exporting country. Recall that only Vietnamese rice price shows a cointegration with the USDA adjusted world price. Accordingly, the estimated bivariate error correction model between USDA adjusted world price and Vietnamese price is shown in (18).

$$\begin{bmatrix}\Delta WMP_t \\ \Delta VN_t\end{bmatrix} = \begin{bmatrix} .0128 \\ (.0333) \\ .2900^{**} \\ (.0514)\end{bmatrix}\left[WMP_{t-1} - \underset{(.0891)}{1.578^{**}} VN_{t-1} - 5.364\right]$$
$$+ \begin{bmatrix} .2333^{*} & -.1643^{**} \\ (.0931) & (.0511) \\ -.2852^{*} & .2062^{**} \\ (.1435) & (.0788)\end{bmatrix}\begin{bmatrix}\Delta WMP_{t-1} \\ \Delta VN_{t-1}\end{bmatrix} + \begin{bmatrix} .0703 & .1331^{**} \\ (.0945) & (.0519) \\ -.1071 & .0908 \\ (.1458) & (.0801)\end{bmatrix}\begin{bmatrix}\Delta WMP_{t-2} \\ \Delta VN_{t-2}\end{bmatrix}$$
$$+ \begin{bmatrix} .0368 & .0149 \\ (.0930) & (.0524) \\ .5335^{**} & .0978 \\ (.1435) & (.0808)\end{bmatrix}\begin{bmatrix}\Delta WMP_{t-3} \\ \Delta VN_{t-3}\end{bmatrix} + \begin{bmatrix} -.0119 \\ (.0157) \\ .0315 \\ (.0248)\end{bmatrix}[seasonal\ dummy]$$
$$+ \begin{bmatrix} .0014 \\ (.0050) \\ -.0001 \\ (.0078)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{WP} \\ \hat{\epsilon}_{VN}\end{bmatrix} \tag{18}$$

Although the result above portrays a cointegration between Vietnamese rice and USDA adjusted world price, the size of the estimated price transmission elasticity for the Vietnamese rice price significantly larger than one, suggesting that, in the long run, the USDA adjusted world price accelerates much faster than the Vietnamese rice price. Lacking clear justification of what drives the relationship, it becomes clear that the USDA adjusted world price might not represent a good proxy for the actual world price. Combined with the fact that we fail to establish long-run relationships with the other two leading rice exporting markets, the overall result does not satisfactorily provide an adequate evidence to support the second hypothesis, *"the USDA adjusted world price reflects a long-run price relationship with Thai,* *Vietnamese, and US rice"*. This conclusion reinforces Taylor, Bessler, Waller, and Rister's (1996)'s skepticism that other considerations besides market force are in place for formulation of the USDA adjusted world price.

6.3 SPOT-FUTURE OF U.S. ROUGH RICE

In this subsection, the following analysis considers whether there is a convergence of spot and futures price of rough rice in the United States domestic market. The estimated bivariate error correction model between spot price of rough rice and rough rice futures is shown below.

$$\begin{bmatrix}\Delta SP_t \\ \Delta FT_t\end{bmatrix} = \begin{bmatrix} -.1827^{**} \\ (.0312) \\ .0807 \\ (.0567) \end{bmatrix} \left[SP_{t-1} - \underset{(.0361)}{1.0194^{**}} FT_{t-1} + .0917 \right]$$

$$+ \begin{bmatrix} .1348^{*} & .1347 \\ (.0678) & (.1231) \\ .1421^{**} & .3920^{**} \\ (.0550) & (.0998) \end{bmatrix} \begin{bmatrix}\Delta SP_{t-1} \\ \Delta FT_{t-1}\end{bmatrix} + \begin{bmatrix} .0207^{*} \\ (.0092) \\ -.0011 \\ (.0167) \end{bmatrix} [seasonal\ dummy]$$

$$+ \begin{bmatrix} .0008 \\ (.0029) \\ .0018 \\ (.0053) \end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{SP} \\ \hat{\epsilon}_{FT}\end{bmatrix} \qquad (19)$$

From (19), the cointegrating equation, or β vector, between the cash price and futures of rough rice shows a coefficient of -1.02 on the futures rice price with a 95% confidence interval between -1.09 and -.95. This is a welcomed result, as the cointegrating equation above extracts the long-run co-movement of cash and futures rice prices[14] and the market efficiency hypothesis is not rejected since the interval also

---

[14] From the cointegrating equation $SP = 1.026FT - .0917 + \hat{\varepsilon}$, we have
$lnP_{SP} = 1.02lnP_{FT} - .09 + \hat{\varepsilon}$
$dlnP_{TH}/dlnP_{VN} = 1.02$

includes -1. The price discovery measure also reveals that 31% of permanent value is associated with the cash price, while 69% is associated with the future price.[15] This evidence seems to suggest that the rice future price is a dominant source of uncovering new information. To this end, the result provides supporting evidence to the <u>third hypothesis</u>, *"there exists a convergence between spot and future markets of rough rice in the US market"*. This finding is also consistent with the literature on price discovery of commodities (see Figuerola-Ferretti, and Gilbert, 2005)

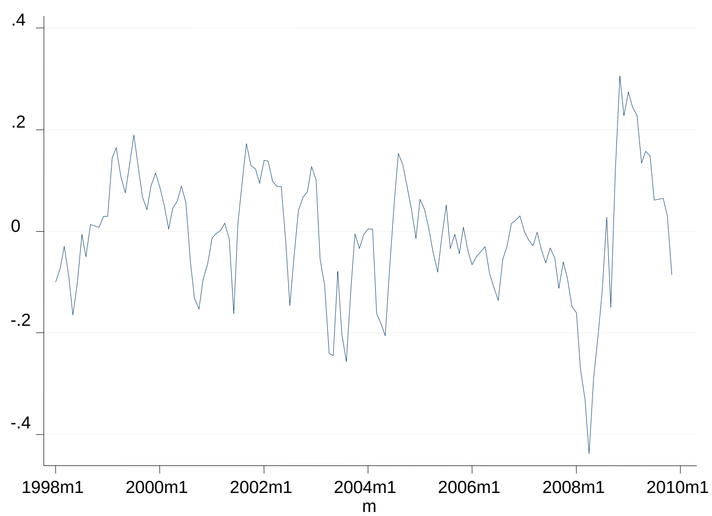


Figure 10 Predicted Cointegrated Equation between Spot and Future market of U.S. rough rice

Table 9 Price Discovery of Spot-Future market of U.S. rough rice

Time period (1998:1-2010:2)	Spot-Future Price Discovery	
	Spot	Future
Price Discovery	0.31	0.69
p-value	0.15	0.00

Source: Author's calculation

$$\frac{dP_{TH}}{P_{TH}} \Big/ \frac{dP_{VN}}{P_{VN}} = 1.02$$

15 $\frac{0.0807}{(0.1827+0.0807)} = 31\%, \frac{0.1827}{(0.1827+0.0807)} = 69\%$

6.4 VERTICAL MARKET

We now shift focus to the question of the degree of price discovery between two levels of market. The estimated bivariate error correction model between the F.O.B export price of U.S. milled rice and the spot price of rough rice is as follows:

$$\begin{bmatrix} \Delta AK_t \\ \Delta SP_t \end{bmatrix} = \begin{bmatrix} -.1768^{**} \\ (.0381) \\ .0818 \\ (.0510) \end{bmatrix} \left[AK_{t-1} - \frac{.8483^{**}}{(.0277)} SP_{t-1} - 1.1205 \right]$$
$$+ \begin{bmatrix} .6818^{**} & .0422 \\ (.0828) & (.0672) \\ .3530^{**} & .1542 \\ (.1109) & (.0900) \end{bmatrix} \begin{bmatrix} \Delta AK_{t-1} \\ \Delta SP_{t-1} \end{bmatrix} + \begin{bmatrix} .0252 & -.1605^{*} \\ (.0987) & (.0666) \\ -.0611 & .0373 \\ (.1322) & (.0892) \end{bmatrix} \begin{bmatrix} \Delta AK_{t-2} \\ \Delta SP_{t-2} \end{bmatrix}$$
$$+ \begin{bmatrix} .2957^{**} & -.0098 \\ (.0913) & (.0658) \\ .0417 & .0803 \\ (.1223) & (.0881) \end{bmatrix} \begin{bmatrix} \Delta AK_{t-3} \\ \Delta SP_{t-3} \end{bmatrix} + \begin{bmatrix} -.0037 \\ (.0079) \\ .0186 \\ (.0105) \end{bmatrix} [seasonal\ dummy]$$
$$+ \begin{bmatrix} .0006 \\ (.0025) \\ .0014 \\ (.0033) \end{bmatrix} + \begin{bmatrix} \hat{\epsilon}_{AK} \\ \hat{\epsilon}_{SP} \end{bmatrix} \quad (20)$$

The cointegrating relationship between the Arkansas milled rice and spot rough rice shows a coefficient of -.85 on spot rough rice price, with a 95% confidence interval between -90 and -.79. The significance of this finding is that it not only reveals a long-run co-movement between two marketing levels, but also provides evidence of a variable marked-up price structure, in contrast to a cost-plus structure. The price discovery measure reveals that only 32% of permanent value is associated with the F.O.B export price of milled-rice while 68% is associated with the spot price of rough rice. Together with results from the last section, it becomes clear that, for the U.S. domestic rice market, the first dominant price discovery process take place in the futures market, and this is followed by the cash market of the rough rice, and then, finally, is filtered through to the export market. The result provides supporting evidence to the fourth hypothesis that "*there exists a long-run price relationship between U.S. rough and milled rice.*"

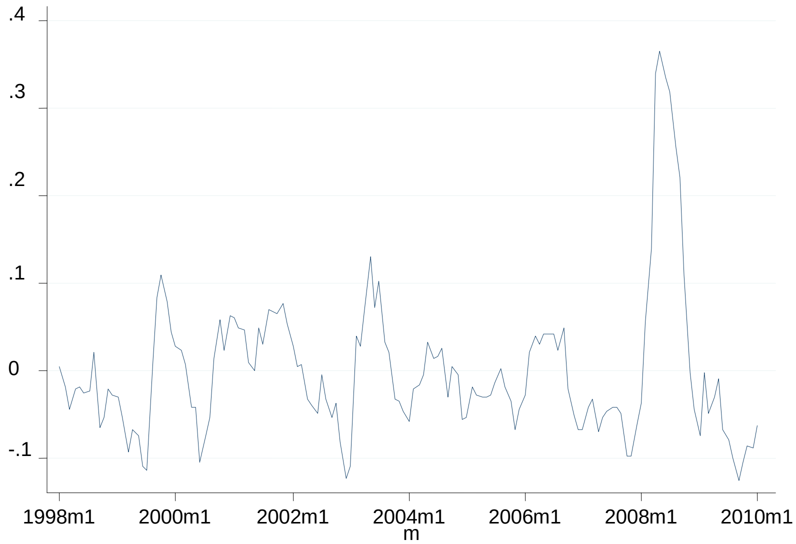


Figure 11 Predicted Cointegrated Equation between rough and milled rice

Table 10 Price Discovery of the Vertical U.S. Rice Markets

Time period (1998:1-2010:2)	Vertical Market Price Discovery	
	Milled	Rough
Price Discovery	0.32	0.68
p-value	0.14	0.00

Source: Author's calculation

7 POLICY IMPLICATIONS

The findings in the previous section establish that there exist long-term price co-movements among the three major rice-exporting countries examined in this study. Estimated price transmission elasticity also provides us with a preliminary tool to gauge the impact from price transmission when the price in one market changes. These estimates should be used with the possibility that deviation from long-run equilibrium can be sustained for a long time. The dominant role of price discovery of Thai and United States rice markets when compared to that of Vietnam underscores the importance of Thailand as the world's largest rice exporter and the United States' dominance in high-quality rice and high-income markets. Within the United States domestic markets, the success of the Chicago Board of Trade rough rice futures as a dominant source of price discovery also highlights the importance of the futures market as a hedging tool to farmers.

Recognizing that rice as a staple food is especially crucial for low income populations. Given the trajectory of population growth and climate change, together with scarcer land and waters, it is clear that rice supply will continue to be under pressure for the foreseeable future. The current thin market condition for rice does not allow room for a significant supply shock. To prevent malnutrition and improve food security means the more efficient role of rice market is required.

Since market inefficiency that arose from supply shock could be averted if the market infrastructure allows more rice to be traded. The first step would be to invest in viable storage technology in the rice growing emerging countries, allowing these countries to accumulate stocks that can be used as a buffer in the event of supply shock. This in effect would incur less degree of rice price volatility as well.

Secondly, governments must continue to support and nurture a viable centralized trading platform. In 2004, Thailand launched a trading of rice futures in the Agricultural Futures Exchange of Thailand (AFET) but the volume trade currently is dismal. Governments can engage more systematically through the exchange to accumulate or lower national rice stocks in order to boost and draw international traders to the market.

8 CONCLUSION

To summarize the main findings previously discussed. First through three bivariate cointegrated systems, this study found that from 1998m1 to 2010m2, long-run price co-movements exist among the three major rice-exporting countries, namely Thailand, Vietnam, and US (Arkansas). In addition to this anticipated finding, the hypothesis of market efficiency, as captured by the price transmission elasticity, $\beta_2 = 1$, are rejected in two of the three pairs, namely Thai-Vietnam and Thai-US(Arkansas) respectively. Based on the findings, the magnitudes of the estimated price transmission elasticities reveal higher acceleration of the Thai and United States rice price in the long-run co-movement than that of the Vietnamese rice.

Nonetheless, contrary to this result, the separate estimations with the data prior to 2008 show supporting evidence to the market efficiency hypothesis on all three price-pairs. This striking revelation highlights an effective mechanism of international rice market from 1998m1 up to 2008m1 and underscores the importance of further investigation on the 2008's period as a root cause of the market inefficiency.

Building on the 1998m1-2010m2 linkages, the Gonzalo & Granger common factor decomposition technique shows the dominant sources of price discovery are Thailand and Arkansas, while the role of Vietnamese rice market appears to be the informational laggard of the three. This finding is sensible as Thailand is the world's largest rice exporter while the United States dominates high-quality rice and high-income markets.

To evaluate the performance of the USDA adjusted world price relative to other major rice exporting countries, this study looks for evidence of long-run relationship of the adjusted world price with Thailand, Vietnam, and United States (Arkansas). In this regard, the USDA adjusted world price shows only a long-run linkage with Vietnamese market. The size of the estimated coefficient far from one also pinpoints that there could be other factors influencing the USDA world price. In combination with the fact that we fail to establish long-run price relationships between the USDA world rice price with the other two leading rice exporting markets, underscores a weakened argument to support validity and transparency of the USDA's methodology. This conclusion reinforces Taylor, Bessler, Waller, and Rister's (1996)'s skepticism that other considerations besides market force are in place for formulation of the USDA adjusted world price.

For the US domestic rice markets, the result shows efficient linkage between spot and future prices of rough rice, demonstrated by Chicago Board of Trade rough rice futures converging to USDA rough rice prices in a cash market. In the vertical markets, the result also reveals the significant price linkage between two marketing level of rough and milled rice in the United States. Building on these linkages, the Gonzalo & Granger's decomposition characterizes the dominant source of the price discovery process in the futures market followed by the cash market of the rough rice and then the milled rice export price.

Going forward, as the issue of food security becomes increasingly essential due to scarcer area of arable land due to competing demand for food as well as for energy, the understanding of the relationship among the major rice exporting countries remains key necessities especially to rice-importing countries to safeguard the failure of international rice supply as a transfer from rice-surplus countries to the rice-deficit countries has a meaningful welfare implication.

The advantage of our model's simplistic design also reveals its inherent shortcoming. The drawback of the model's reliance on price information is the limited interpretation of the estimated elasticities that precludes other significant factors underlying market force. A plausible explanation for the differences in size of percentage change that defines the co-movement includes fluctuating exchange rates, with exporters incurring costs in local currency but quoting prices in United States dollars, and transportation and transaction costs. Additional study can incorporate the role of traditional versus non-traditional investor in commodities market as well as investigate market efficiency in the context of bilateral, multilateral trade agreements.

APPENDIX

A. THREE LEADING RICE EXPORTING COUNTRIES VECM IN A SUB-PERIOD 1998m1-2008m1

In this section the result of the VECM estimation from 1998m1 to 2008m1, or the period prior to the 2008's price spike is provided for Thailand, Vietnam, and US (Arkansas). The hypothesis of market efficiency based on $\beta_2 = 1$ cannot be rejected on all three cases.

$$\begin{bmatrix}\Delta TH_t\\ \Delta VN_t\end{bmatrix} = \begin{bmatrix}-.1285^{*}\\ (.0570)\\ .2210^{**}\\ (.0548)\end{bmatrix}\left[TH_{t-1} - \underset{(.0470)}{1.054^{**}}VN_{t-1} + .2110\right] + \begin{bmatrix}.3514^{**} & -.0598\\ (.1026) & (.0885)\\ .3111^{**} & .2742^{**}\\ (.0985) & (.0849)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-1}\\ \Delta VN_{t-1}\end{bmatrix} + \begin{bmatrix}-.0101\\ (.0105)\\ .0354^{**}\\ (.0101)\end{bmatrix}[seasonal\ dummies] + \begin{bmatrix}.0021\\ (.0032)\\ .0012\\ (.0031)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{TH}\\ \hat{\epsilon}_{VN}\end{bmatrix} \quad \text{(A1)}^{16}$$

$$\begin{bmatrix}\Delta TH_t\\ \Delta AK_t\end{bmatrix} = \begin{bmatrix}-.0469^{*}\\ (.0231)\\ .0484^{**}\\ (.0172)\end{bmatrix}\left[TH_{t-1} - \underset{(.1587)}{1.010^{**}}AK_{t-1} + .2947\right] + \begin{bmatrix}.2955^{**} & .0831\\ (.0949) & (.1235)\\ .0134 & .5443^{**}\\ (.0705) & (.0918)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-1}\\ \Delta AK_{t-1}\end{bmatrix} + \begin{bmatrix}-.1159 & -.0297\\ (.0954) & (.1418)\\ -.0550 & .0381\\ (.0709) & (.1055)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-2}\\ \Delta AK_{t-2}\end{bmatrix} + \begin{bmatrix}-.0569 & -.0927\\ (.0928) & (.1279)\\ .0271 & .2042^{*}\\ (.0690) & (.0951)\end{bmatrix}\begin{bmatrix}\Delta TH_{t-3}\\ \Delta AK_{t-3}\end{bmatrix} + \begin{bmatrix}-.0126\\ (.0107)\\ .0027\\ (.0079)\end{bmatrix}[seasonal\ dummies] + \begin{bmatrix}.0016\\ (.0032)\\ .0015\\ (.0024)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{TH}\\ \hat{\epsilon}_{AK}\end{bmatrix} \quad \text{(A2)}$$

---

[16] * and,** denote 5% and 1% significance level. Standard errors are given in parentheses. The coefficient of the first variable in the cointegrating vector was normalized to one and thus does not have standard errors. The constant term in the cointegrating equation does not have standard errors since it is not directly estimated but is backed out from other estimates.

$$\begin{bmatrix}\Delta \mathrm{AK}_t\\ \Delta VN_t\end{bmatrix} = \begin{bmatrix}-.0428^{*}\\ (.0170)\\ .0695^{**}\\ (.0240)\end{bmatrix}\left[AK_{t-1} - \underset{(.1651)}{1.041^{**}} VN_{t-1} - .0977\right]$$

$$+\begin{bmatrix}.5347^{**} & .1052\\ (.0923) & (.0625)\\ .1215 & .3756^{**}\\ (.1298) & (.0880)\end{bmatrix}\begin{bmatrix}\Delta AK_{t-1}\\ \Delta VN_{t-1}\end{bmatrix} + \begin{bmatrix}.0487 & -.0473\\ (.1058) & (.0654)\\ .1527 & -.1594\\ (.1488) & (.0920)\end{bmatrix}\begin{bmatrix}\Delta AK_{t-2}\\ \Delta VN_{t-2}\end{bmatrix}$$

$$+\begin{bmatrix}.1981^{*} & -.0678\\ (.0935) & (.0935)\\ -.4180^{*} & .0940\\ (.1315) & (.0897)\end{bmatrix}\begin{bmatrix}\Delta AK_{t-3}\\ \Delta VN_{t-3}\end{bmatrix} + \begin{bmatrix}.0029\\ (.0074)\\ .0171\\ (.0104)\end{bmatrix}[seasonal\ dummies]$$

$$+\begin{bmatrix}.0016\\ (.0024)\\ .0009\\ (.0034)\end{bmatrix} + \begin{bmatrix}\hat{\epsilon}_{AK}\\ \hat{\epsilon}_{VN}\end{bmatrix} \qquad \text{(A3)}$$

B. SUMMARY OF THE COINTEGRATED EQUATIONS AND TEST OF MARKET EFFICIENCY HYPOTHESIS AMONG THREE LEADING RICE EXPORTING COUNTRIES

Cointegrated Price Relationships	1998m1-2008m1	1998m1-2010m2	
		Original	Treated Data[17]
TH-VN	$TH_{t-1} - \frac{1.054^{**}}{(.0470)} VN_{t-1} + .2110$	$TH_{t-1} - \frac{1.1676^{**}}{(.0381)} VN_{t-1} + .8145$	$TH_{t-1} - \frac{1.300^{**}}{(.0548)} VN_{t-1} — 5.662$
AK-VN	$AK_{t-1} - \frac{1.041^{**}}{(.1651)} VN_{t-1} - .0977$	$AK_{t-1} - \frac{.9703^{**}}{(.0982)} VN_{t-1} - .4852$	$AK_{t-1} - \frac{1.1880^{**}}{(.1012)} VN_{t-1} - 5.881$
TH-AK	$TH_{t-1} - \frac{1.010^{**}}{(.1587)} AK_{t-1} + .2947$	$TH_{t-1} - \frac{1.2930^{**}}{(.1298)} AK_{t-1} + 1.9043$	N/A

Source: compiled from the findings

Hypothesis of Market Efficiency	1998m1-2008m1	1998m1-2010m2	
		Original	Treated Data
TH-VN	Not rejected	Rejected	Rejected
AK-VN	Not rejected	Not rejected	Not rejected
TH-AK	Not rejected	Rejected	N/A

Source: compiled from the findings

[17] The price series were treated to account for non-market intervention in Vietnamese market by i) an OLS regression on dummy variable that indicates that rice exports ban is enforced, ii) then the computed residuals are used to enter the VECM estimation.

C. JARQUE-BERA TEST FOR NORMALITY DISTRIBUTED RESIDUAL TERMS

In this section, we perform misspecification tests to ensure that the underlying statistical assumptions are valid. First, Jarque-Bera test for normally distributed residuals are performed and the results are given below.

For each single-equation, the null hypothesis is that the error term has a univariate normal distribution. For the system with two equations jointly, the null hypothesis is that the disturbances come from bivariate normal distribution.

Equation	$\chi2$	df	Prob > $\chi2$
D_lnTH	29.247	2	0.00000
D_lnVN	4287.261	2	0.00000
ALL	4316.508	4	0.00000
D_lnTH	360.662	2	0.00000
D_lnAK	1.295	2	0.52324
ALL	361.958	4	0.00000
D_lnAK	7.819	2	0.02005
D_lnVN	6417.770	2	0.00613
ALL	6425.589	4	0.00000
D_lnWP	91.305	2	0.00000
D_lnVN	6799.894	2	0.00000
ALL	6891.199	4	0.00000
D_lnSP	9.621	2	0.00814
D_lnFT	12.3	2	0.00213
ALL	21.921	4	0.00021
D_lnAK	203.483	2	0.00000
D_lnSP	41.937	2	0.00000
ALL	245.42	4	0.00000

Source: Compiled from Stata’s output

D. LAGRANGE-MULTIPLIER TEST

Lagrange-multiplier tests are also performed for the possible autocorrelation of the residuals. The formula for the Lagrange-multiplier test statistic at lag j is

$$LM_s = (T - d - .5)\ln\left(\frac{|\hat{\Sigma}|}{|\tilde{\Sigma}_s|}\right)$$

where T is the number of observations in the estimation; $\hat{\Sigma}$ is the maximum likelihood estimate of the variance-covariance matrix of the error terms from the estimation; and the $\tilde{\Sigma}_s$ is the maximum likelihood estimate of Σ from the following augmented regression. First, for the case of 2 equations in the estimation, the error terms is defined to be a 2×1 vector of residuals. And for each lag s, an augmented regression in which the newly created error vector are lagged s times is formed. d is the number of coefficients estimated in the regression. The asymptotic distribution of LM_s in the case of 2-equation-system is χ^2 with 2^2 degree of freedom.

	Lag s	χ2	df	Prob > χ2
TH-VN	1	4.0299	4	0.40198
	2	3.4184	4	0.49039
	3	2.8971	4	0.57518
	4	2.8238	4	0.58773
TH-AK	1	2.4615	4	0.65155
	2	4.3421	4	0.36168
	3	4.0289	4	0.40211
	4	2.4307	4	0.65709
AK-VN	1	2.0583	4	0.72504
	2	2.8825	4	0.57767
	3	2.4107	4	0.66069
	4	3.0352	4	0.55194

Source: Compiled from Stata’s output

	Lag *s*	$\chi 2$	df	Prob > $\chi 2$
WMP-VN	1	7.0997	4	0.13071
	2	7.4833	4	0.11245
	3	4.3526	4	0.36038
	4	0.9407	4	0.91864
SP-FP	1	1.6864	4	0.79318
	2	3.1866	4	0.5271
	3	6.0689	4	0.19406
	4	5.6022	4	0.23089
AK-SP	1	1.3858	4	0.84665
	2	5.0438	4	0.28283
	3	2.0384	4	0.7287
	4	3.2234	4	0.52116

Source: Compiled from Stata's output

The Lagrange-multiplier test does not detect problems of autocorrelation in any systems. Since the likelihood function is derived under the assumption that the residuals are independently, identically, and normally distributed with zero mean and finite variance. If the normality assumption is rejected but the residual is independent and identically distributed, the estimates are consistent but not efficient. This implies that the estimates would converge to the true value in large sample but may not have minimum variance.

E. CHRONOLOGY OF THE 2007-08 RICE CRISIS

This table is an excerpt from Slayton (2009) appendix.

Date	Event
December 21st, 2007	**NFA** tender 500,000 tons and buys 422,701 tons at an average price of $410 CNF. Vietnam sells 410,701 tons.
December 31st	**Thailand**'s shipments exceed 9.5 million tons, while **India** records 6.3 million tons – including 5.25 million tons of non-Basmati. **Vietnam**'s exports top 4.5 million tons.
Janruary 14th, 2008	Cold spell begins in northern **Vietnam** which over several weeks destroys 148,000 ha of transplanted rice and 10,000 ha of seedlings.
Janruary 18th	**Vietnam**'s export target raised from 4.4 million tons to 4.5 million tons, including 700,000 tons January-March, 1.5 million tons each in second and third quarters, and 800,000 tons October-December. Exports are allowed to resume with MEP of $385 for 5% and $360 for 25% for January-February shipment and $400 for 5% for March. Informally, VFA asks that no sales of 25% be made.
Janruary 29th	**NFA** tenders for 550,000 tons and buys 463,750 tons at an average price of $475 CNF. Vietnam sells 300,000 tons. Facing large losses on unshipped contracts concluded in late 2007, major exporters in **Thailand** stop offering price quotes.
February 1st	**NFA**'s stocks are equivalent to 8 days requirements, just over half of the targeted level of 15 days. (NFA's stocks are to be at least 30 days during the "lean" season.)
February 5th	Vietnam informally banned new rice exports by revoking MEP and stop issuing any additional MEPs during February 2008.
March 6th	Export quotas announced in **Vietnam**: January-March 700-800,000 tons, April-June 1.3-1.5 million tons, July-September 1.3-1.4 million tons, and October-December 700-800,000 tons. This effectively bans further sales in March given existing sales to NFA and others.
March 14th	**Vietnam** Food Association issues letter to members banning exports through April and promises guidance for May shipment

at a later time.

March 17th — First arrivals in Bangladesh from initial contract by **India** under promised 500,000 tons. **Vietnam** bans further sales for March and April. Only sales for May allowed.

March 25th — **Vietnam** extends export ban through June and reduces export target by .5-1.0 million tons to 3.5-4.0 million tons.

March 26th — **Vietnam** decrees that no export contracts will be approved unless exporter is holding as stock 50% of the sale, prices must be in line w/ MEP, and shipment within 60 days; Export quotas revised: January-June 2.25 million tons (50% of total exports in ’06 and ’07); 3.5 million tons through September; MOU with the Philippines for 1.5 million tons announced.

March 28th — Commerce minister in **Thailand** told press that export prices would reach $1,000/ton by June and that farmers should not be in a hurry to sell.

March 29th — Export tax proposed in **Vietnam**.

April 1st — **India** bans non-Basmati exports, Basmati MEP rises to $1,200. Head of All India Rice Exporters’ Association forecasts a 5.25 million ton drop in exports over the next twelve months. In an effort to quell rising domestic prices, the minister of commerce indicates the government in **Thailand** will release up to 650,000 tons from its stockpiles into the domestic market at below-market prices. In the **Philippines**, the government rejects a proposal to reduce the import tariff on rice from its current rate of 50%.

April 4th — With hoarding underway by rice mills, traders, and public, prime minister in **Thailand** assures there is enough rice for domestic consumption, but indicates a commerce ministry plan to distribute subsidized rice is not necessary. Finance and commerce ministers repeat assurances that the country will not restrict exports. Following a “summit” with government officials and farm experts on how to contain the rice crisis, President Arroyo in the **Philippines** announces the government will spend over $1.0 billion to increase rice production. The agriculture secretary tells press that 2007 imports could reach as high as 2.7 million tons.

April 8th	In the **Philippines**, the government indicates it will tender for 500,000 tons in May. NFA tenders for 100 TMT of U.S. #2/4%; purchased 72,600 tons for June-Sep shipment at average price of over $1,058.
April 10th	**Thailand**'s commerce minister announces that effective from April 14 all retail rice prices will be reduced 10% for two months. Also, head of Internal Trade Department indicates that millers and exporters must report stock levels monthly and that army could be called in to guard rice warehouses. Also, exporters will be required to hold stock of 500 tons. Head of Thai rice exporters' association predicts 100% B will soon reach $1,000.
April 11th	In **Thailand**, the commerce minister asks the military to guard the government-held stocks.
April 14th	U.N. secretary general says global food crisis has reached "emergency proportions."
April 17th	**NFA** tenders for 500,000 tons and buys 364,000 tons at an average price of $1,075 CNF for 25% (323,375 tons) and $1,129 for 5% (40,625 tons each).. Vietnam sells 80,000 tons (not including 20,875 tons sourced out of Pakistan via Long An Food).
April 22nd	Prime minister of Thailand denies country will restrict exports.
April 23rd	Head of NFA indicates it is considering holding weekly import tenders.
April 25th	In **Vietnam** "rice fever" breaks out, prices have doubled in HCMC over the course of a couple of days. **India** announces it will build 5 million ton "strategic reserve" of food grains, including 2 million tons of rice – beyond its stocking norms. U.N. Secretary General Ban calls for concerted and immediate action to solve global food crisis. Head of FAO tells press global food crisis could result in "civil wars."
April 28th	Decree against speculators issued in Vietnam.
April 30th	Thailand's prime minister revives proposed rice exporter cartel, OREC.

May 5th	**NFA** tender for 675,000 tons fails as only one offer received and it without a sovereign guarantee; the Philippines talks of waiting until fall to buy. It indicates it will not, in any case, pay above $1,200/ton.
May 6th	Following objections from the Philippines and the ADB, **Thailand** scraps OREC proposal.
May 13th	Malaysia buys from Thailand 100,000 tons each of 5% at $950 and 15% at $940.
May 19th	Philippines discloses Japan may also provide 200,000 tons imported rice.
May 21st	Major exporters in Thailand resume offering price quotes.
May 23rd	**Thailand**'s visiting prime minister reportedly tells President Arroyo that Bangkok is prepared to sell its stocks to the Philippines at friendship prices.
June 2nd	At FAO summit on food crisis, **Japan**'s P.M. Fukuda commits "to release in the near future over 300,000 tons of imported rice" to the world market. Japan also discloses Sri Lanka has requested up to 200,000 tons of food aid. Prime minister takes commerce minister's proposal off cabinet agenda that would authorize **Thailand** to participate in NFA's request for G-to-G offers of 600,000 tons by June 13.
June 10th	Prime minister takes commerce minister's proposal off cabinet agenda that would authorize **Thailand** to participate in NFA's request for G-to-G offers of 600,000 tons by June 13.
June 13th	Philippines receives offers for G-to-G purchase of 600,000 tons.
June 18th	Vietnam's export ban lifted; MEP $800 for 5% established.

Source: Slayton (2009)

REFERENCES

Baillie, R. T., Booth, G. G., Tse, Y., & Zabotina, T. (2002). Price Discovery and Common Factor Models,. *Journal of Financial Markets* , 309-321.

Barrett, C. B. (2001). Measuring Integration and Efficiency in International Agricultural Markets. *Review of Agricultural Economics* , 19-32.

Bashir, K. A. (2002). *Analysis of Market Power and Trade Liberalization of International Rice Market Allowing for Rice Type:A Linear Complementarity Approach.* A Pennsylvania State University Graduate School.

Childs, N., & Burdett, A. (2000). *The U.S. Rice Export Market.* USDA Economic Research Service.

Childs, N., & Kiawu, J. (2009). *Factors Behind the Rise in Global Rice Prices in 2008.* Washington, D.C.: USDA Economic Research Service.

Conforti, P. (2004). Price Transmission in Selected Agricultural Markets. *FAO Commodity and Trade Policy Research Working Paper No.7* .

Cramer, G., Wailes, E., Chavez, E., & Hanzen, J. (1998). Arkansas Global Rice Modeling Project. In R. Norman, & T. Johnston, *B.R. Wells Rice Research Studies 1998* (pp. 380-388). Fayetteville: University of Arkansas.

Dawe, D. (2007). International Rice Market. *Manuscript unpublished* .

Dawson, P., & Dey, P. (2002). Testing for the Law of One Price: Rice Market Integration in Bangladesh. *Journal of International Development* , 473-484.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* , 427-431.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica* , 813-836.

Engle, R., & Granger, C. (1987). Cointegration and Error Coreection: Representation, estimation, and testing. *Econometrica* , 251-276.

Fackler, P., & Goodwin, B. (2001). Spatial price transmission. In B. Gardner, & G. Rausser, *Handbook of Agricultural Economics,Volume 1, Chapter 17.* Elsevier Science.

Figuerola-Ferretti, I., & Gilbert, C. (2005). Price Discovery in the Aluminium Market. *Journal of Futures Markets* , 967-988.

Figuerola-Ferretti, I., & Gonzalo, J. (2008). Modelling and Measuring Price Discovery in Commodities Markets. *Working Paper,* Madrid: Universidad Carlos III de Madrid.

Genetically Modified Rice. (2008, December 4). Retrieved from GMO Compass: http://www.gmo-compass.org/eng/grocery_shopping/crops/24.genetically_modified_rice.html

Gonzalo, J., & Granger, C. (1995). Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business & Economic Statistics* , 27-35.

Goodwin, B. K., & Piggott, N. E. (2001). Spatial Market Integration in the Presence of Threshold Effects. *American Journal of Agricultural Economics* , 302-317.

Granger, C. W. (1986). Developments in the Study of Cointegrated Economic Variables. *Oxford bulletin of Economics and Statistics* , 213-228.

Hands Off Our Rice. (2010, August). Retrieved August 24, 2010, from Greenpeace International: http://www.greenpeace.org/international/en/campaigns/agriculture/problem/genetic-engineering/hands-off-our-rice/

Harris, F., McInish, T., & Wood, R. (1997). Common Long-Memory Components of Intraday Stock Prices: A Measure of Price Discovery. *Wake Forest University University Working Paper* .

Harris, F., Mcinish, T., & Wood, R. (2002). Security Price Adjustment Across Exchangs: An Investiagation of Common Factor Components for Dow Stocks. *Journal of Financial Markets* , 277-308.

Hasbrouck, J. (2003). Intraday Price Formation of US Equity. *Journal of Finance* , 2375-2399.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* New York: Oxford University Press.

Kaosa-ard, M., & Juliano, B. O. (1990). Assessing Rice Quality Characteristics and Prices in Selected International Markets. *Rice Grain Marketing and Quality Issues* (pp. 23-35). Seoul, Korea: International Rice Research Institute.

Karp, L., & Perloff, J. (1989). Dynamic Oligopoly in the Rice Export Market. *Review of Economic Statistics* , 462-470.

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis.* Germany: Springer.

McNew, K., & Fackler, P. L. (1997). Testing Market Equilibrium: Is Cointegration Informative? *Journal of Agricultural and Resource Economics* , 191-207.

Ng, S., & Perron, P. (2000). Lag Length Selection and the Construction of Unit Root Tests with Good size and Power. *Working Paper.* Boston: Department of Economics, Boston College.

Ng, S., & Perron, P. (1995). Unit Root Tests in ARMA Models with Data-dependent Methods for the Selection of the Truncation Lag. *Journal of the American Statistical Association* , 268-281.

Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika* , 335-346.

Pintong, C. (1977). *A Price Analysis of the Thai Rice Marketing System.* Ph.D. Dissertation, Standford University.

Prasertsri, P. (2009). *Thailand Grain and Feed Annual.* Bangkok: USDA Foreign Agricultural Service.

Quan, T. Q. (2009). *Vietnam Grain and Feed Annual.* Hanoi: USDA Foreign Agricultural Service.

Rashid, S. (2004). *Spatial Integration of Maize Markets in Post-Liberalized Uganda.* Washington, D.C.: International Food Policy Research Institute.

Rice. (2010, June 24). Retrieved from GMO Compass: http://www.gmo-compass.org/eng/database/plants/064.rice.html

Rice:Background. (2009, November 24). Retrieved from USDA Economic Research Service.

Serapio, M. J. (2010, March 16). *After China, Philippines may approve GMO rice.* Retrieved August 20, 2010, from Reuters: http://www.reuters.com/article/idUSTRE62F0T720100316

Siamwalla, A., & Haykin, S. (1983). *The World Rice Market: Structure Conduct and Performance.* DC: International Food Policy Research Institute.

Slayton, T. (2009). Rice Crisis Forensics: How Asian Governments Carelessly Set the World Rice Market on Fire. *Working Paper Number 163.* DC: Center for Global Development.

Statistics 2008. (2010, Aug 23). Retrieved from Thai Rice Exporters Association : http://www.thairiceexporters.or.th/statistic_2008.htm

Taylor, E. L., Bessler, D. A., Waller, M. L., & Rister, E. M. (1996). Dynamic Relationships between US and Thai rice prices. *Agricultural Economics* , 123-133.

The Golden Rice Project. (2010, August 24). Retrieved from Golden Rice: http://www.goldenrice.org/Content4-Info/info.html

Thompson, S. R., Sul, D., & Bohl, M. T. (2002). Spartial Market Efficiency and Policy Regime Change: Seemingly Unrelated Error Correction Model Estimation. *American Journal of Agricultural Economics* , 1042-1053.

Timmer, C. P., & Dawe, D. (2007). Managing Food Price Instability in Asia: A Macro Food Security Perspective. *Asian Economic Journal* , 1-18.

USDA Economic Reserach Service. (2010, March). *Rice Yearbook: Dataset*. Retrieved August 23, 2010, from USDA Economics, Statistics, and Market Information System: http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1229

Westcott, P. C., & Price, M. J. (2001). *Analysis of the U.S. Commodity Loan Program with Marketing Loan.* Washington, D.C.: Economic Division, Economic Research Service, USDA.

Yan, B., & Zivot, E. (2007). A Structural Analysis of Price Discovery Measures. *Working Paper.* Seattle: University of Washington.

Yumkella, K., Unnevehr, L., & Garcia, P. (1994). Noncompetitive Pricing and Exchange rate Pass-through in selected U.S. and Thai Rice Markets. *Journal of Agricultural Applied Economics* , 406-416.

CHAPTER THREE

THE DETERMINANTS TO FORMAL CREDIT AND INFORMAL CREDIT PARTICIPATION: THE CASE OF THAI FARM HOUSEHOLDS

1 INTRODUCTION

This study examines the determinants of agricultural households' participation in formal and informal credit markets in Thailand. Formal credits are broadly defined as credits obtained from institutions that are registered to do lending business. Informal credits are credits obtained from all non-institution lenders, such as local lenders, landlords, traders, or more casually, a person outside of a household.

Increasing modernization, and export-oriented development that pull resources away from traditional agricultural sector and into manufacturing and services, entail an ever expanding income gap between urban rich and rural poor. And, despite growth in formal credit in recent years, rural farmers and entrepreneurs still have difficulties accessing funding, particularly those who are poorer and perceived to be at higher risk (see Paulason & Townsend, 2004). Insufficient funding, despite the need for farm households both to finance farming operations and smooth their consumption cycle, underscores the onset of a vicious cycle in which external shocks in income shortfalls or unexpected consumption needs can leave even less funding available for the productive cause, leading to the growing portion of principal (see Siamwalla, et al., 1993). With considerable impact on a large portion of Thai population, alleviating the debt situation of Thai farmers remains the central policy of all recent administrations.

While the literatures suggest that an informal credit market can emerge from credit rationing of the formal credit sector, what remains to be seen is a careful analysis of micro foundation of this informal credit market, itself. The lack of credible data on informal credit typically prevents a thorough empirical study on the informal market. And, indeed, there are only few empirical studies that investigate Thai farmers' credit situation (see Siamwalla, et al., 1993, Paulason and Townsend, 2004, Gine, 2005, Boonperm, Haughton, and Khandker, 2009, Kaboski and Townsend, 2009). Despite this setback, detailed analysis of the role of parallel credit market and its significance remains key to designing informed and sound policy. Accordingly, this study's main objective is to determine, based on household characteristics and their observed qualities, factors affecting probability of success in informal credit participation.

Using the 2006 Socio-Economic Household Survey of Thailand, this study follows Heckman's two-stage selection model approach to determine the informal loan participation of Thai agricultural households. Since we can observe the behavior of formal and informal credit participants only if they participate in new loans, one might view this pool of samples as self-selected. The dire consequence of self-selection bias is first illuminated in Heckman (1979). To demonstrate within this context, unobservable variables such as business acumen may affect both the probability of success in obtaining credit as well as the probability of success in participating in formal loans. Higher values of business acumen lead to higher success in obtaining credit, and thus successful borrowers are likely to have higher values of business acumen than non-borrowers. As a consequence of this, the probability of success in obtaining formal credit is overestimated from the influence of any selected sample. Hence, excluding the selection stage to engage in new loans from the full estimation induces a selection bias due to a non-random sample. See Achen (1986) and Sartori (2003) for more discussion of selection bias.

Results from the two-stage selection model provide supporting evidence that, facing credit rationing, households are more likely to 'stick' to the credit market they previously engaged in, whether the formal or informal credit market. This finding reinforces the vicious cycle which makes it more difficult for farmers to get out of debt. Secondly, the study finds supporting evidence that wealthier households are less likely to access credit, but, if they do, they are more likely to participate in formal credits, and the opposite is true of less wealthy households. Thirdly, the study finds supporting evidence for less credit participation during May and December coinciding with the planting and harvesting seasons. Specifically, the probability for accessing new credit starts off high in the beginning of the year then slowly declines while participation in informal credit, conditional on participating in new credit, rises toward the end of the year. Alarmingly, this indicates that, while overall access to new credit declines, people who need credit tend to participate in informal credit more as months pass. This piece of evidence pinpoints the limited role of BAAC in providing working capital only and not consumption-based loans underlining the importance of the lack of formal credit facility for farmers, and should be addressed quickly. Finally, the model finds that households with owned farmland are more likely to participate in the formal credit market, while households with rented farmland are more likely to participate in the informal credit market stressing the use of owned land as collateral to participate in formal credit channels.

There are two key contributions to this study. First, to the best of my knowledge, this is the first implementation of Heckman's approach in a context of formal/informal loan participation using Thai data. Second, processing the survey data into twelve monthly pieces allows us to uncover a pattern in seasonal debt participation of Thai farmers.

2 SURVEY OF LITERATURE RELATED TO THE INFORMAL CREDIT MARKET

The theoretical understanding of the informal credit market has been advanced by the study of imperfect information and imperfect enforcement in credit markets (see Hoff, Braverman, and Stiglitz, Introduction, 1993). This section briefly describes a theoretical foundation for understanding characteristics of the informal credit market.

In equilibrium with full information, lenders know types of borrowers beforehand, and charge high risk types with higher interest rates and low risk types with lower interest rates. There are two rates to clear the market. In a world with asymmetric information, since lenders cannot differentiate between high risk borrowers and low risk borrowers, lenders charge one rate for all. If the rate starts to rise, low-risk borrowers drop out first, and lenders' income discontinuously drops because high risk borrowers have lower probability of repayment. Results of important work by Stiglitz and Weiss (1981) serve to explain why, int equilibrium, some borrowers are still rejected by formal credit institutions, despite efforts to pay higher interest rates. Credit rationing arose from financial institutions setting interest rates lower than the market-clearing rate allows for the coexistence of the informal credit market to serve residue demand for loans.

This understanding of the coexistence of formal and informal credit markets has been emphasized by the empirical study of Bell, Srinivasan, & Udry (1997), who provide evidence of credit rationing in the formal sector and informal credit flourishing to serve this unsatisfied demand. Kochar (1997) provides theoretical and empirical evidence to support such coexistence but postulates that the extent of rationing is less than what is conventionally assumed. Guirkinger (2006) found empirical evidence that the informal sector in Piura, Peru serves both households excluded from the formal sector and households that prefer informal loans because of lower transaction costs.

Next, since informal credits usually lack a formal contract and are not legally binding, what motivates borrowers to repay their debts is simply a guarantee to get future loans. Aryeetey & Udry (1995) show that, by using a simple setup of repeated games, Subgame Perfect Nash Equilibrium reveals that borrowers have incentives to repay their loans in order to secure future borrowings. The implication of this work is that an informal credit market can continue to thrive without becoming a legal business entity as long as borrowers have easy access to informal creditors, and there are no competing creditors.

In this niche and segmented market, the work by Hoff and Stiglitz (1998) shows that increases in the number of moneylenders creates less incentive for borrowers to maintain a good credit reputation, since it is easier to defect to other money lenders. The increase in monitoring efforts of moneylenders can result in higher interest rates charged. Gine (2005) formulated a theoretical model and used Thai data to show that enforcement costs serve as a barrier to entry into informal credit markets.

Siamwalla, et al. (1993) also noted that interest rates of the informal sector were higher because transaction costs are high. These transaction costs reflect the information costs associated with current borrowers. Since only previous lenders have access to this information, if the same borrowers were to shift to different lenders, the marginal information costs to the new lender would be quite high.

3 BRIEF HISTORY OF BAAC AND VRF

In 1966, the Thai Government established the Bank of Agriculture and Agricultural Cooperatives (BAAC) as a state enterprise under the jurisdiction of the Ministry of Finance, to provide credit to the agricultural sector. The Thai Government mandates that commercial banks have to lend to the agricultural sector as a certain proportion of their overall portfolio, or can achieve that goal by making a deposit to the BAAC. In order to maintain a high repayment rate, the BAAC implements the following policies. The due date of a debt is usually right after the harvesting season. The amount of a loan overdue will be subject to a higher interest rate than the previous rate. New loans will not be provided until the old debt is sorted out -- either by paying the interest component alone, or both interest and principle. The BAAC also implements a pooled credits system. If a member of the pool, comprising eight to fifteen people, defaults, the group as a whole becomes ineligible for new loans. As a result, the operation at the BAAC has been very successful and become a central credit facility for Thai farmers. It is often observed that Thai farmers are more inclined toward servicing the debt of the BAAC, sometimes even at the cost of tapping into higher interest rate loans from local lenders. The success of the BAAC operation is emphasized by a rapid expansion in lending to farmers and farmers’ institutions with the amount of lending growing at a compound average growth rate of 8.5% per year over ten years, while the average loan per farmer grew at 6.2% per year.[18]

Table 1 BAAC's Operation

Number of Farm Households Serviced by BAAC (in millions)

	2004	2005	2006	2007	2008
Farmers	3.9	4.0	4.1	4.3	4.54
Members of Farmer Institutions	1.5	1.5	1.6	1.6	1.5
Total	5.4	5.5	5.7	5.9	6.1
%change	*0.3%*	*3.0%*	*2.5%*	*4.1%*	*2.7%*

BAAC's Lending Operation (in billions Baht)

	2004	2005	2006	2007	2008
Normal Operation					
Farmers	299.0	345.2	369.8	397.8	418.5
Farmer Institutions	15.0	17.8	21.3	22.1	24.7
Total	314.0	363.0	391.1	419.9	443.2
%change	*15.2%*	*15.6%*	*7.7%*	*7.4%*	*5.5%*
Average Loan to Farmers (in thousands Baht)	58.4	65.5	68.9	71.0	73.0
%change	*14.9%*	*12.3%*	*5.1%*	*3.2%*	*2.8%*
Government Mandate					
Government Secured Loans Project	37.1	32.2	12.0	5.6	4.25
Loans to VRF	9.0	6.4	4.9	2.6	1.0
Others	20.7	20.1	20.6	21.2	31.457
Total	66.8	58.7	37.5	29.3	36.7
%chng	*35.3%*	*-12.1%*	*-36.2%*	*-21.8%*	*25.2%*
BAAC's Deposit	354.3	431.4	496.6	514.7	585.9
Lending to Deposit Ratio	1.1	1.0	0.9	0.9	0.8

Source: compiled from BAAC's Annual reports

[18] Author's calculation from data compiled from BAAC annual reports

Despite its modest success, the BAAC is only allowed to give out loans based on ‘productive’ causes, specifically working capital toward farm operation. While this coverage specifically aims to exclude the portion that relates to household consumption, lacking other sources of credit, it is often observed that farmers find themselves tapping into these funds for consumption purposes, anyway. This leaves less available funding for productive agricultural purposes. As a result, farmers often find themselves entering the vicious cycle of never being able to pay off debt entirely as its principle portion grows larger every year (see Siamwalla, et al., 1993)

In 2001, the Thai Government introduced the Thailand Village Revolving Fund (VRF) as a government-funded microfinance project to lift farm households out of chronic poverty. Essentially, the government injected 1 million baht into each village via a locally-run village fund committee which then supervised the setting of interest rates, loan amounts, terms of loans, and other requirements for borrowing. To set up a committee requires a quorum of at least three quarters of adults in the village. The committee then consists of 15 members, half of whom are women (see Boonperm, Haughton, and Khandker, 2009). Boonperm, Haughton, & Khandker, (2009) report that average VRF borrower is 47% poorer than those who do not borrow based on income per capita, and twice as likely to be farmers and self-employed. The astounding success of the VRF in providing credit access to poor farmers as part of a set of populist policies helped bring the initiated-party a second term landslide victory in February of 2005 (see Thai Rak Thai's Victory: It was Thaksin who made up voters' minds, 2005, Mid-year Economic Review, 2004, Profile: Thaksin Shinawatra, 2009).

4 EMPIRICAL STRATEGY

This section describes a household loan participation model. First, households choose to access credit markets, then choose to participate in either informal or formal loans. Probability of success in borrowing and participating in either formal or informal credit channels is to be determined exclusively by the household characteristics.

This two-step decision—to gain access first, then choose to participate in either informal or formal loans—utilizes bivariate probit with a selection model, an equivalent of Heckman's selection model (Heckman, 1979), except that here both selection equation and outcome equation are probit. Since we can observe the behavior of formal and informal credit participants only if they participate in new loans, one might view this pool of samples as self-selected. For instance, an unobservable variable such as business acumen may affect both the probability of success in obtaining credit as well as probability of success in participating in a formal loan. Higher values of business acumen lead to higher success in obtaining credit, and thus successful borrowers are likely to have higher values of business acumen than non-borrowers. As a consequence of this, the probability of success in obtaining formal credit is overestimated from the influence of the selected sample. Hence, excluding the selection stage, to engage in new loans from the full estimation, induces a selection bias due to a non-random sample. See Achen (1986) and Sartori (2003) for more discussion of selection bias.

First consider the following binary model where z_i^* is the unobserved cost-benefit function of household *'s* participation in a credit market. And let z_i be the observed binary variable of household *i*, equal to 1 if household *i* accesses a new loan and 0 otherwise. z_i^* is determined by a vector of explanatory variables *w*.

$$z_i^* = w_i^{'}\gamma + u_i \quad (1)$$

$$z_i = \begin{cases} 1\ if\ z_i^* > 0 \\ 0_i\ if\ z_i^* \leq 0 \end{cases} \quad (2)$$

Further assume that,

$$u_i \sim N(0,1)$$

And let $\Phi(.)$ be a cumulative density function of standard normal distribution. Then we have,

$$\Pr(z_i = 0) = \Pr(z_i^* \leq 0)$$

$$= \Pr\left(w_i^{'}\gamma + u_i \leq 0\right)$$

$$= \Pr(u_i \leq -w_i^{'}\gamma)$$

$$= \Phi(-w_i^{'}\gamma) \quad (3)$$

And by symmetry,

$$\Pr(z_i = 1) = \Pr(z_i^* > 0)$$

$$= \Pr\left(w_i^{'}\gamma + u_i > 0\right)$$

$$= \Pr(u_i > -w_i^{'}\gamma)$$

$$= 1 - \Phi(-w_i^{'}\gamma)$$

$$= \Phi(w_i^{'}\gamma) \quad (4)$$

Next, consider at the same time that household *i* is accessing new credit, its characteristics also affect its probability of success to participate in either the informal or formal credit market. Let y_i be 1 if household *i* participates in the informal credit market, and 0 otherwise. And let the underlying index function y_i^* depend on a vector of explanatory variables x.

$$y_i^* = x_i^{'}\beta + \varepsilon_i \quad (5)$$

$$y_i = \begin{cases} 1 \; if \; y_i^* > 0, z_i^* > 0 \\ 0 \; if \; y_i^* \leq 0, \; z_i^* > 0 \end{cases} \quad (6)$$

Now, let $\phi_2(.)$ be a bivariate density function and assume further that ε, u are bivariate standard-normally distributed with mean zero and a correlation between u and ε equal to ρ.

Formally,

$$\varepsilon_i \sim N(0,1)$$

$$u_i \sim N(0,1)$$

$$corr(\varepsilon_i, u_i) = \rho$$

Then, it follows that

$$\Pr(y_i = 1, z_i = 1) = \Pr(y_i^* > 0\ ,\ z_i^* > 0; \rho)$$

$$= \Pr\left(x_i^{'}\beta + \varepsilon_i > 0\ ,\ w_i^{'}\gamma + u_i > 0; \rho\right)$$

$$= \Pr(\varepsilon_i > -x_i^{'}\beta\ ,\ u_i > -w_i^{'}\gamma; \rho) \qquad (7)$$

where $\Pr(y_i = 1, z_i = 1)$ is the probability of household *i* success in participating in informal credit. Recall the bivariate probability density function,

$$\phi_2(\varepsilon, u) = \frac{1}{2\pi\sigma_\varepsilon\sigma_u\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\left(\frac{\varepsilon^2 + u^2 - 2\rho\varepsilon u}{1-\rho^2}\right)\right]$$

Rearranging terms in (7) gives,

$$\Pr(y_i = 1, z_i = 1) = \int_{-\infty}^{w_i^{'}\gamma} \int_{-\infty}^{x_i^{'}\beta} \phi_2(\varepsilon, u; \rho)\, d\varepsilon du$$

$$= \Phi_2(x_i^{'}\beta\ ,\ w_i^{'}\gamma; \rho) \qquad (8)$$

and,

$$\Pr(y_i = 0, z_i = 1) = \Pr(z_i = 1) - \Pr(y_i = 1, z_i = 1)$$

$$= \Phi(w_i^{'}\gamma) - \Phi_2(x_i^{'}\beta, w_i^{'}\gamma; \rho). \quad (9)$$

where $\Pr(y_i = 0, z_i = 1)$ is the probability of household *i*'s success in participating in formal credit. Finally, the log-likelihood function for the bivariate probit model with selection is,

$$lnL = \sum_{i=1}^{N} \delta_i \{z_i y_i ln\, \Phi_2(x_i^{'}\beta, w_i^{'}\gamma; \rho)$$

$$+ z_i(1 - y_i) \ln[\, \Phi(w_i^{'}\gamma) - \Phi_2(x_i^{'}\beta, w_i^{'}\gamma; \rho)]$$

$$+ (1 - z_i) \ln \Phi(-w_i^{'}\gamma)\} \quad (10)$$

where δ_i is the weight for observation *i*.

Then, using the maximum likelihood procedure, we can obtain $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\rho}$.

$$\max_{\{\beta,\gamma,\rho\}} \, lnL = \max_{\{\beta,\gamma,\rho\}} \sum_{i=1}^{N} \delta_i \{z_i y_i ln\, \Phi_2(x_i^{'}\beta, w_i^{'}\gamma, \rho)$$

$$+ z_i(1 - y_i) \ln[\, \Phi(w_i^{'}\gamma) - \Phi_2(x_i^{'}\beta, w_i^{'}\gamma; \rho)]$$

$$+ (1 - z_i) \ln \Phi(-w_i^{'}\gamma)\} \quad (11)$$

It is important to note that when $\rho = 0$, the joint log-likelihood is the sum of the two log-likelihoods of two separate probits. And it would be more efficient to estimate each probit independently.

4.1 MARGINAL EFFECTS

To determine the probability of participation in informal credit conditional on accessing new credit, predicted marginal probability of the estimates can be calculated from the following formula.

$$\Pr(y_i = 1 | z_i = 1) = \frac{\Pr(y_i = 1, z_i = 1)}{\Pr(z_i = 1)}$$

$$= \frac{\Phi_2(x_i'\beta, w_i'\gamma; \rho)}{\Phi(w_i'\gamma)} \qquad (12)$$

Similarly, the probability of participating in formal credit conditional on accessing new credit is the following:

$$\Pr(y_i = 0 | z_i = 1) = \frac{\Pr(y_i = 0, z_i = 1)}{\Pr(z_i = 1)}$$

$$= \frac{\Phi(w_i'\gamma) - \Phi_2(x_i'\beta, w_i'\gamma; \rho)}{\Phi(w_i'\gamma)} \qquad (13)$$

For references in this section see Golder (2008) and Greene (2003).

5 DATA

The data source is the Socio-Economic Survey of Thailand (2006) obtained from the Thai National Statistics Office. The design of the survey employs a stratified two-stage sampling method with Changwat (76 provinces) stratum. For each province, two additional stratum are added and classified as municipal and non-municipal areas. Each municipal and non-municipal area or rural village is independently sampled with probability that is equal to the inverse of the number of households within that village. Total households within each chosen village are compiled and ranked by size and economic type. From each municipal area, 15 households are sampled. And, for a non-municipal area or a village, only 10 households are sampled. The number of total households is then equally divided into 12 groups. Each individual group has approximately the same number of households in both municipal and non-municipal areas. Each group is then randomly selected for a survey for each of the twelve months. The survey collects a wide variety of socio-economic data, including household characteristics, income, expenditure, debt, and an assets profile.

We have a total of 6195 household samples in the estimations. The weight structure associated with the survey-design was applied to the estimation.

5.1 VARIABLES

Table 2 lists the statistical properties of key variables used in the analysis. SES defines loans from the formal sector as loans that originated from: i) Commercial Banks, ii) the BAAC, iii) the Government Housing Bank and Saving Bank, iv) Other financial companies, v) Savings co-op organization's welfare, and vi) the VRF. Loans from informal source are defined as loans that originated from persons outside of households.

The dependent variable is a binary variable of household *i* having a new debt and a binary variable of household *i* having an informal debt. The key feature of this SES dataset is that the sample was collected throughout the year, allowing us to monitor the month/ harvest cycle effect using seasonal dummy variables.

Total income per capita measures net monthly household income, net of remittance, per capita. Age, age-squared, and gender are the variables associated with heads of households. Primary, secondary, and college education are the ratio variables associated with proportion of households with at least primary education, secondary education, and college education and above.

Family size counts the number of all members within the household. The dependency ratio measures the number of children under 15 and elderly, aged 65 and over, divided by the number of the household members between age 15 and 65.

$$Dependency\, ratio = \frac{(\#\, of\, people\, aged\, 0-\ 14) + (\#\, of\, people\, aged\, 65\, and\, over)}{\#\, of\, people\, aged\, 15-65}$$

Land, land 2, and land 3 are three agricultural land variables classified by ownership. The land variable is the area of farmlands owned by the head of household, including parents and relatives. Land2 is the area of farmlands rented from other persons. And Land3 is the area of public lands.

Region 2-5 are regional dummies, Central (2), North (3), North East (4), and South (5), capturing the regional variation in agricultural productivity, rainfall, and level of infrastructure development. Region 1 is Bangkok and its neighboring industrial provinces, with minimal presence of farmland.

Rice x Region 2 to 5 are interaction dummy variables associated with being rice farmers, operating in associated regions. Since the survey shows half of farm households are rice farmers, incorporating these variables allows us to investigate characteristics of rice farmers in different regions more carefully. Monthly dummies represent the surveyed month capturing variation in harvest cycles.

Other shock variables include medical illness, disability, and negative financial shocks. These represent dummy variables of external shocks to households. A medical shock is 1 if the household incurred medical expenses associated with in-patient hospitalization. Disability is 1 if the household has at least 1 disabled person. A financial shock is 1 if the household reports problems of being behind on house rent, water, or electricity bills, or school tuitions and fees.

Informal and formal outstanding debts represent dummy variables associated with a household having informal or formal debt outstanding up to the previous month. Based on the 2006 SES, Table 3 shows 11% of Thai farm households reported reliance on informal credit markets, with 4% reporting that they depend solely on informal credit. 78% of farm households reported access to formal credit, while 71% reported that they solely depend on formal credit. 7% of farm households depend on both sources, and 18% reported being debt-free. The Pearson Chi-squared and corrected F-test shows significant association between households with formal and informal debt access, (see Table 4).

Table 2 Statistical Summary of key variables

	Full Sample	No Existing Debts	Existing Debts		New Debts	
			Formal	Informal	Formal	Informal
Number of Observations	6,390	1,310	4,839	672	272	196
Value of Existing Debt (in Thousand Baht)	96	0	98	115	128	118
Value of Formal Debt (in Thousand Baht)	72	0	92	56	125	86
Value of Informal Debt (in Thousand Baht)	7	0	6	59	3	30
Monthly HH Income/capita (in Baht)	4,261	5,619	4,014	2,318	3,554	1,615
Monthly HH Expense/capita (in Baht)	3,738	4,585	3,565	3,170	4,017	2,390
Head Age	52	56	51	49	51	47
Family Size	3.64	3.32	3.71	3.78	3.86	3.92
Dependency Ratio	0.56	0.60	0.55	0.65	0.56	0.71
Owned Agricultural Land Size (in Rai)	16.7	13.5	17.6	15.5	17.8	12.8
Rental Agricultural Land Size (in Rai)	5.1	2.2	6.1	8.8	8.9	11.5
Public Agricultural Land Size (in Rai)	0.4	0.2	0.5	0.5	1.4	1.1
Number of Female Head Households--%	20	26	19	25	17	23
Number of households with head's education level.. (%)						
At least elementary	82	78	84	81	80	78
At least high school	10	9	11	11	14	13
At least college	7	13	6	8	5	9
Percentage of households with other members' education level..(%)						
At least elementary	62	63	62	58	61	63
At least high school	20	18	21	20	17	18
At least college	15	17	14	18	17	15
Percentage of households with education level.. (%)						
At least elementary	68	68	68	66	66	66
At least high school	17	15	18	16	17	16
At least college	13	15	12	15	14	14

Source: SES 2006

Table 3 Summary of key variables (continue)

	Full Sample	No Existing Debts	Existing Debts		New Debts	
			Formal	Informal	Formal	Informal
Percentage of households who have report having the following conditions. (%)						
Medical Treatment	16	15	16	17	25	18
Disability	0*	0*	0*	0*	0*	0*
Financial Problems	8	5	8	16	13	20
Percentage of households who live in the following area. (%)						
Bangkok and surrounding area (region 1)	0*	0*	0*	1	0*	1
Central (region 2)	14	16	13	19	8	20
North (region 3)	26	19	17	22	36	23
Northeastern (region 4)	46	38	48	41	38	45
South (region 5)	14	27	11	17	18	11
Number of households surveyed in each month. (%)						
Month 1	9	8	9	12	18	14
Month 2	9	9	9	10	20	8
Month 3	8	8	7	8	9	11
Month 4	9	11	8	6	16	6
Month 5	9	9	9	7	5	5
Month 6	7	6	8	6	5	9
Month 7	8	10	8	6	6	9
Month 8	8	8	8	5	3	7
Month 9	8	7	8	8	6	8
Month 10	8	6	8	11	4	8
Month 11	10	10	10	12	4	6
Month 12	7	8	7	9	4	8
Number of Rice farming households in each region (compared with other households-- %)						
Bangkok and surrounding area (region 1)	0*	0*	0*	0*	0*	0*
Central (region 2)	6	5	6	8	2	9
North (region 3)	14	10	14	12	14	15
Northeastern (region 4	38	33	39	35	32	40
South (region 5)	1	1	1	2	1	1

Source: SES 2006, 0* denotes less than 0.5%

Table 4 Percentage of farm households with debt

Formal	Informal 0	Informal 1	Total
0	18.03 (0.0084)	3.91 (0.0037)	21.94 (0.0081)
1	70.89 (0.0074)	7.18 (0.0021)	78.06 (0.0081)
Total	88.92 (0.0033)	11.08 (0.0033)	1
Pearson:			
Uncorrected chi2(1) =			82.2587
Design-based F(1, 23) =			34.979

Source: SES 2006
Note: Standard errors are given in parentheses

This model focuses on households choosing to participate in credit markets. Other households that do not want or need credit, and declared so in the survey, are *excluded* from the estimation. This leaves us with households that answer can borrow/ borrow some/ and unable to borrow. And since the model does not allow for households who borrowed both from formal and informal sources, 17 households are excluded out of 481 households that obtained new loans.

6 HYPOTHESES

This section summarizes main testable hypotheses and corresponding parameters of interest.

Hypothesis 1: *Households whose loan demand could not be met in the formal sector were credit rationed in that sector and had to rely on informal loans to finance agricultural production.*

The theoretical view on coexistence of parallel credit market as a result of credit rationing in the formal sector was shared by Stiglitz and Weiss (1981), and empirically tested by Bell, Srinivasan, & Udry (1997), Kochar (1997), and Guirkinger (2006). The presence of credit rationing in the formal sector, illustrated by households holding existing loans from the informal sector, would increase probability of households' borrowing again from informal sector. On the contrary, the presence of households holding an existing formal loan demonstrates the households' ability to borrow from the formal sector, and thus would increase probability of the households' borrowing again from the same sector. As a result of this, the coefficient for the existing informal loan should be positive and the coefficient for the existing formal loan should be negative for the probability of informal credit participation.

Hypothesis 2: *Less wealthy households are more likely to borrow, and are more likely to borrow from informal credit market than their wealthier peers.*

We are interested to see the net effect of income both as a factor determining the access to credit as well as a factor involving the formal/informal credit participation. Since poor households are more likely to face consumption constrains, and to rely more on credit markets, the probability of credit access should be higher for the households at lower level of wealth. Nonetheless, the less wealthy households are also considered riskier, due to the lack of collaterals, and are more likely to face credit rationing from the formal sector and to rely more on the informal credit market. As a result, the coefficient on income should be negative for the probability of credit access but positive for the probability of informal credit participation.

Hypothesis 3: *Farm households tend to borrow before the planting season, and less during the harvesting season.*

In our data, households that grow rice make up 59% of total agricultural households and the majority of these households are located in northeastern region where they grow only one crop of rice per year due to lack of proper irrigation systems. Rice productions require more working capital prior to the planting season to buy seeds, fertilizers, equipments, and machineries and less during the harvesting. Rice seeds are regularly planted following the dry season, mainly from May to June and harvested after the wet season, predominantly from October to December. As a result, activities involving new credit should be higher between January and May. Accordingly, the coefficients on monthly dummy variables should be positive between January and May and negative from June to December for the probability of credit access.

Hypothesis 4: *The greater the amount of land, the more working capital households need, thus the more likely they are to borrow.*

It is interesting to observe the effect of various types of land since the SES data provides us with three distinct types of land variables. First the size of land suggests a level of working capital the households might need for farm operations. In addition to this, the presence of land owned also suggests a possibility of land being used as collateral to acquire new loans from the formal credit sector. Accordingly, the coefficient on owned land should be positive on the probability of accessing new credit, and negative on the probability of informal credit participation. Additionally, the coefficients on both rental and public land should be positive on the probability of accessing new credit and positive on the probability of informal credit participation.

7 RESULTS

To show robustness of estimated parameters, results from four model specifications are offered in this section. The first specification provides a baseline model in which the probability of access to credit is determined by household characteristics, land size, external shocks, region, and monthly dummy variables. The probability of informal loan participation is determined by income, dummy variables for existing formal and informal loans, household characteristics, land size, and external shocks.

The second specification preserves all of the baseline covariates and adds dummy variables for rice farming in each region to the set of explanatory variables for the probability of accessing credit. Similar to the second specification, the third specification preserves all of the baseline covariates and adds dummy variables for rice farmers in each region to the set of explanatory variables for the probability of credit access. It also adds monthly dummy variables to explain the determination of the informal loan participation. Finally the fourth specification incorporates dummy variables for rice farming in each region in both sets of explanatory variables of credit access and informal loan participation.

7.1 PROBABILITY OF CREDIT ACCESS

Table 5 and 6 report results from the determination of credit access equation. Results from the credit access stage show that negative and highly significant coefficients on income across all four models suggest that wealthier households are less likely to access new loans. This is accentuated by the fact that, on average, increasing wealth substitutes the farmers' need to access funding. The positive and significant coefficient on family size indicates that households with higher numbers of family member are more likely to access new loans.

The model also finds coefficients on rental farmland and public land positive and highly significant, suggesting that households that rented lands or use public lands are more likely to take new loans. This result pinpoints the need for policymakers to monitor this particular risky group carefully, since rent payment can adversely affect the level of working capital and the longer-term viability of the farm's operations. Accordingly, the positive and significant coefficients on rental and public land coincide with the prediction of the fourth hypothesis, that the greater the amount of land, the more likelihood there is of borrowing due to more working capital required. Interestingly, the implication of the owned farmland remains unsolved.

Table 5 Parameters estimated from Access to Credit equation model I, II

Dependent Variables: 1{New loan}, 0{otherwise}	Model (I)		Model (II)	
	Coefficients	Linearized Standard Errors	Coefficients	Linearized Standard Errors
Income	-0.000006***	0.0000	-0.000006***	0.0000
Head of HH age	-0.0132	0.0125	-0.0149	0.0124
Head of HH age square	0.000091	0.0001	0.00010	0.0001
1{Head of HH is a female}	-0.0033	0.0609	-0.0037	0.0580
Number of households with head's education level.. (%)				
At least elementary	-0.0436	0.0908	0.0842	0.1299
At least high school			0.1167	0.1523
At least college	-0.1272	0.1537		
Family size	0.0458**	0.0204	0.0447**	0.0205
Dependency ratio	0.0633	0.0428	0.0621	0.0442
land size(owner)	-0.0014	0.0023	-0.0015	0.0023
land2 (rental)	0.0049**	0.0019	0.0051**	0.0020
land3 (public)	0.0144***	0.0048	0.0143***	0.0048
Medical	0.1353*	0.0696	0.1328*	0.0686
Disability	0.1283	0.5368	0.1399	0.5448
Financial problems	0.3937***	0.0908	0.3916***	0.0929
1{Region 2 - Central}	-0.0352	0.5184	0.0154	0.5360
1{Region 3 - North}	0.1480	0.5157	0.1655	0.5323
1{Region4 - Northeast}	-0.0039	0.5203	-0.1690	0.5292
1{Region 5- South}	0.1902	0.5210	0.1571	0.5309
1{Month 2}	0.0961	0.1407	0.0911	0.1328
1{Month 3}	-0.1672	0.1184	-0.1509	0.1158
1{Month 4}	-0.0460	0.0877	-0.0492	0.0853
1{Month 5}	-0.4497**	0.2135	-0.4510**	0.2050
1{Month 6}	-0.3308	0.2001	-0.3263	0.1981
1{Month 7}	-0.3532**	0.1623	-0.3686**	0.1531
1{Month 8}	-0.4940***	0.1486	-0.4983***	0.1357
1{Month 9}	-0.3026**	0.1277	-0.2974**	0.1253
1{Month 10}	-0.4345***	0.1505	-0.4389***	0.1444
1{Month 11}	-0.5426***	0.1046	-0.5472***	0.0972
1{Month 12}	-0.4466***	0.1408	-0.4596***	0.1440
1{rice farming*region 2}			-0.2637*	0.1364
1{rice farming*region 3}			-0.1391	0.0919
1{rice farming*region 4}			0.1353*	0.0699
1{rice farming*region 5}			-0.1551	0.1805
_cons	-1.0429*	0.5864	-1.0663*	0.5853

*,**,*** denote 10%, 5%, and 1% level of significance

Table 6 Parameters estimated from Access to Credit equation model III, IV

Dependent Variables: 1{New loan}, 0{otherwise}	Model (III)		Model (IV)	
	Coefficients	Linearized Standard Errors	Coefficients	Linearized Standard Errors
Income	-0.000006***	0.0000	-0.000006***	2.00E-06
Head of HH age	-0.0168	0.0128	-0.0157	0.0125
Head of HH age square	0.0001	0.0001	0.0001	0.0001
1{Head of HH is a female}	-0.0044	0.0608	0.0001	0.0597
Number of households with head's education level.. (%)				
At least elementary	0.0765	0.1268	0.0675	0.1270
At least high school	0.1113	0.1529	0.1042	0.1524
At least college				
Family size	0.0448**	0.0205	0.0439**	0.0210
Dependency ratio	0.0586	0.0433	0.0583	0.0437
land size(owner)	-0.0015	0.0023	-0.0014	0.0023
land2 (rental)	0.0051**	0.0020	0.0051**	0.0020
land3 (public)	0.0149***	0.0049	0.0148***	0.0048
Medical	0.1392**	0.0669	0.1393**	0.0692
Disability	0.1626	0.5460	0.1685	0.5373
Financial problems	0.3862***	0.0867	0.3852***	0.0913
1{Region 2 - Central}	0.0010	0.5396	-0.0251	0.5357
1{Region 3 - North}	0.1544	0.5394	0.1403	0.5320
1{Region4 - Northeast}	-0.1607	0.5352	-0.2023	0.5296
1{Region 5- South}	0.1464	0.5365	0.1334	0.5312
1{Month 2}	0.0001	0.1290	0.0922	0.1334
1{Month 3}	-0.1471	0.1327	-0.1423	0.1142
1{Month 4}	-0.1282	0.0968	-0.0474	0.0884
1{Month 5}	-0.5791**	0.2155	-0.4521**	0.2150
1{Month 6}	-0.3140	0.2198	-0.3277	0.1939
1{Month 7}	-0.3540**	0.1410	-0.3629**	0.1524
1{Month 8}	-0.5705***	0.1538	-0.5233***	0.1333
1{Month 9}	-0.3274**	0.1187	-0.3083**	0.1271
1{Month 10}	-0.4405***	0.1445	-0.4328**	0.1544
1{Month 11}	-0.6167***	0.1181	-0.5545***	0.1028
1{Month 12}	-0.4028**	0.1564	-0.4547***	0.1481
1{rice farming*region 2}	-0.2462*	0.1358	-0.1077	0.1819
1{rice farming*region 3}	-0.1416	0.0977	-0.1107	0.1102
1{rice farming*region 4}	0.1259	0.0736	0.2081*	0.0866
1{rice farming*region 5}	-0.1444	0.1916	-0.1275	0.2665
_cons	-0.9587	0.5860	-1.0266*	0.5862

*,**,*** denote 10%, 5%, and 1% level of significance

Negative and significant coefficients from May through December reveal that farm households are less likely to borrow during the planting and harvesting season. While the timing of credit access coincides with the prediction of the third hypothesis, it may also highlight the limited role of BAAC to support agricultural production loans and not to offer consumption-based loans. The estimation finds positive and significant coefficients on medical and financial shocks suggesting households faced with existing medical and financial problems are more likely to take new loans. Negative and significant coefficients on rice farmers of the central and northeastern region suggest less probability of these farmers to credit access.

7.2 PROBABILITY OF INFORMAL LOAN PARTICIPATION

Table 7 and 8 show results from probability of an informal-formal loan participation equation. Households choose to participate in either formal or informal loans. Thus an increase in probability of informal loan participation implies a decrease in probability of formal loan participation.

In this outcome stage, the estimation finds strong evidence to support the first hypothesis, "*Households whose loan demand could not be met in the formal sector were credit rationed in that sector and had to rely on informal loans to finance agricultural production.* The presence of credit rationing in the formal sector, illustrated by households holding existing loans from the informal sector, increases probability of households' borrowing again from the same sector as shown by positive and highly significant coefficient of existing informal loans. It turns out that the opposite is also true, as the coefficient for the existing informal loan is positive and the coefficient for the existing formal loan is negative for the probability of informal credit participation. The presence of households holding existing formal loan demonstrates the households' ability to borrow from the formal sector, and thus increases probability of the households' borrowing again from the same sector. The main application of this result is to predict the reoccurrences of loan sources based on previous, existing loan sources.

The estimation finds that higher income households are more likely to borrow from formal sources. This, combined with the discussion in the previous subsection, gives us a general e result of the wealth effect, that wealthier households are less likely to borrow, and more likely to borrow from formal sources. The opposite is also true, as wealthy households are more likely to borrow, and are more likely to borrow from the informal sources. The results coincide with prediction of the second hypothesis, *"Less wealthy households are more likely to borrow, and are more likely to borrow from informal credit market than their wealthier peers."*.

The estimation also finds that increases in family size positively affect the probability of informal credit participation. Together with the result from the previous subsection, this highlights the importance of welfare policy aiming at farm households with more than one child as the larger a family, the more likelihood to borrow, and to borrow from informal credit source.

Land ownership plays a defining role in determining the source of credit. The estimation finds the coefficient of owned land size negative and significant, albeit at a 10% level, suggesting that households with owned lands are more likely to participate in formal loans. The positive and highly significant coefficient of rented land size produces strong evidence that households with rental land are more likely to borrow from informal sources. Combined with the previous subsection, these results underscore collateral requirements in the formal credit market and the lack thereof in the informal credit sector.

Finally, the positive and significant coefficients on being a rice farmer in the central or northeastern region emphasize a likeliness of these farmers to borrow from the informal sources.

The estimated ρ reveals that the error terms in the two equations are closely linked. The unobservable factor influencing the probability of credit access is also determining the probability of informal loan participation. This result suggests that individual estimations would entail sample selection bias.

Table 7 Parameters estimated from Formal/Informal Loan Participation equation model I, II

Dependent Variables: 1{inf}, 0{fml}	Model (I)		Model (II)	
	Coefficients	Linearized Standard Errors	Coefficients	Linearized Standard Errors
Income	-0.00001***	0.0000	-0.00001***	2.94E-06
1{existing formal loan}	-0.7453***	0.2617	-0.7005***	0.2001
1{existing informal loan}	0.7476***	0.2056	0.7062***	0.1494
Number of households with head's education level.. (%)				
At least elementary	0.0371	0.1588	0.0294	0.1499
At least high school	0.0978	0.2283	0.0952	0.2222
Family size	0.0456**	0.0188	0.0465**	0.0177
Dependency ratio	0.0941	0.0843	0.0924	0.0779
land size(owner)	-0.0063*	0.0034	-0.0061*	0.0034
land2 (rental)	0.0055***	0.0009	0.0054***	0.0009
land3 (public)	0.0092	0.0071	0.0090	0.0067
Medical	0.0311	0.1106	0.0312	0.1083
Disability	-0.7736	0.5429	-0.7203	0.5278
Financial problems	0.3981*	0.2293	0.3781	0.2325
1{Month 2}				
1{Month 3}				
1{Month 4}				
1{Month 5}				
1{Month 6}				
1{Month 7}				
1{Month 8}				
1{Month 9}				
1{Month 10}				
1{Month 11}				
1{Month 12}				
1{rice farming*region 2}				
1{rice farming*region 3}				
1{rice farming*region 4}				
1{rice farming*region 5}				
_cons	-1.5563***	0.2416	-1.5915***	0.1918
ρ	0.9704	0.0328	0.9787	0.0198
Number of Observations	6211		6195	
Population size	3679000.2		3670866.5	

*,**,*** denote 10%, 5%, and 1% level of significance

Table 8 Parameters estimated from Formal/Informal Loan Participation equation model III, IV

Dependent Variables: 1{inf}, 0{fml}	Model (III)		Model (IV)	
	Coefficients	Linearized Standard Errors	Coefficients	Linearized Standard Errors
Income	-0.00001**	3.65E-06	-0.00001***	2.68E-06
1{existing formal loan}	-0.9211***	0.2643	-0.7664***	0.2096
1{existing informal loan}	0.8085***	0.1503	0.7426***	0.1468
Number of households with head's education level.. (%)				
At least elementary	0.0091	0.1345	0.0050	0.1340
At least high school	0.0742	0.2250	0.1171	0.2039
Family size	0.0684**	0.0275	0.0513**	0.0192
Dependency ratio	0.0793	0.0863	0.0515	0.0712
land size(owner)	-0.0066**	0.0032	-0.0059*	0.0034
land2 (rental)	0.0055***	0.0011	0.0047***	0.0008
land3 (public)	0.0107	0.0069	0.0114	0.0068
Medical	0.0483	0.1189	0.0366	0.1106
Disability	-0.7112	0.4859	-0.6991	0.4925
Financial problems	0.3771	0.2302	0.3504	0.2387
1{Month 2}	-0.4149**	0.1770		
1{Month 3}	-0.0080	0.1957		
1{Month 4}	-0.3276	0.2350		
1{Month 5}	-0.4482***	0.1175		
1{Month 6}	0.0110	0.3125		
1{Month 7}	0.0431	0.2071		
1{Month 8}	-0.2324	0.3054		
1{Month 9}	-0.1098	0.2116		
1{Month 10}	-0.0245	0.2196		
1{Month 11}	-0.1952	0.1919		
1{Month 12}	0.1323	0.1919		
1{rice farming*region 2}			0.4451**	0.2098
1{rice farming*region 3}			0.1303	0.1588
1{rice farming*region 4}			0.2375**	0.0973
1{rice farming*region 5}			0.1178	0.2736
_cons	-1.3201***	0.3062	-1.6448***	0.2196
ρ	0.9680	0.0330	0.9773	0.0209
Number of Observations	6195		6195	
Population size	3670866.5		3670866.5	

*,**,*** denote 10%, 5%, and 1% level of significance

7.3 GOODNESS OF FIT

As measures of goodness-of-fit this section reports percentages correctly predicted to show how each model performs. For each observation, predicted probability is calculated based on household covariates. If predicted probability exceeds 0.5, we predict the outcome to be certain or unity. If predicted probability is equal or less than 0.5, we predict the outcome to be zero. The percentage of times the predicted outcome matches the actual outcome is the percent correctly predicted. The overall percent correctly predicted is the weighted average of the percent correctly predicted of each outcome {0,1}, with weight being the fraction of each outcome.

Table 9 reports the comparison of percentages correctly predicted across four models with a high percentage of success in predicting no credit access outcome—higher than 99%, while success in predicting access to credit is much lower. Nonetheless, conditional on actual credit access, success rates in predicting informal-formal loan participation are higher than 60% in all cases. The third and fourth specifications give the best results in terms of overall percentage correctly predicted, while the third specification yields the highest success rate in predicting new informal loans. In light of this evidence, model three will be used to calculate marginal effects in the next section.

Table 9 Percent Correctly Predicted

Percent Correctly Predicted	Model I	Model II	Model III	Model IV
Success in predicting no new loan	99.9%	99.9%	99.9%	99.9%
Success in predicting new loan	1.2%	1.2%	1.2%	1.2%
Percent correctly predicted at credit access equation	92.4%	92.4%	92.4%	92.4%
Population size	3679000	3670866	3670866	3670866
Conditional on Actual new loan=1				
Success in predicting new formal loan	93.9%	95.3%	93.4%	94.3%
Success in predicting new informal loan	63.9%	65.9%	69.4%	68.2%
Percent correctly predicted at informal credit participation	80.8%	82.4%	82.9%	82.9%
Population size	281484	280858	280858	280858
Conditional on Predicted new loan=1				
Success in predicting new formal loan	38.1%	54.1%	46.1%	67.5%
Success in predicting new informal loan	100.0%	100.0%	100.0%	100.0%
Percent correctly predicted at informal credit participation	53.1%	65.2%	58.0%	75.4%
Population size	5413	5413	5934	5413

Source: Author’s calculation

7.4 MARGINAL EFFECTS

In this section, predicted probability based on model (III) is calculated using household characteristics, land size, external shocks, region, monthly dummy variables, and dummy variables for rice farmers in each region to determine the probability of credit access. The probability of informal loan participation is determined by income, dummy variables for existing formal and informal loans, household characteristics, land size, external shocks, and monthly dummy variables to explain the determination of the informal loan participation.

The average of calculated probabilities is then grouped by income quartile, month, and land ownership, and displayed in Table 9. Based on the results, there are a few interesting observations.

First, wealthier farmers are less likely to access new credit. However, when they do borrow, they are more likely to participate in formal credit. Conditional on accessing new credit, the poorest farmers, indicated by being in the first income quartile, have 9 percentage points higher[19] in probability to participate in informal credit compared to the richest farmers (fourth income quartile).

Secondly, the probability to access new credit starts off higher in the beginning of the year, then slowly declines while participation in informal credit, conditional on participating in new credit, rises toward the end of the year. This indicates that while overall access to new credit declines, people who need credit tend to participate in informal credit more as months pass. The probability difference between May and

[19] 50% versus 40.9%

December is, markedly, 27.7 percentage points.[20] This striking evidence raises questions about the diminished role of formal credit institutions in the latter part of the year.

Third, farmers who own land show the lowest probability of accessing new credit, while farmers who both own and rent land, or rent land only, show higher probability of accessing new credit. Among the four groups, farmers who do not own but use public land show the highest probability of accessing new credit. Conditional probability of participation in informal credit is highest among farmers who rent and do not own land, followed closely by farmers who rely on public land only. In both cases, the probability is higher than 50%, indicating they are more likely to borrow from informal credit sources. Farmers who both rent and own some lands appear to have the lowest conditional probability to participate in informal credit.

[20] 68.1% in December versus 40.4% in May.

Table 10 Predicted probability based on households' characteristics

Marginal Effect	P(new debt=1)	P(inf=1)	P(inf=1|new debt=1)
1st income quartile	8.4%	7.0%	50.0%
2nd income quartile	7.8%	5.9%	47.5%
3rd income quartile	7.2%	5.2%	47.1%
4th income quartile	5.8%	3.7%	40.9%
Jan	13.4%	8.6%	42.9%
Feb	12.9%	3.6%	22.6%
Mar	9.9%	6.9%	45.8%
Apr	10.2%	3.8%	29.2%
May	4.2%	3.0%	40.4%
Jun	7.4%	7.0%	52.1%
Jul	6.9%	7.5%	56.7%
Aug	4.3%	4.1%	51.0%
Sep	7.0%	5.7%	48.2%
Oct	5.8%	6.8%	57.7%
Nov	3.9%	5.4%	59.7%
Dec	6.0%	9.0%	68.1%
Owned land only	7.0%	5.5%	47.6%
Owned and Rented land	8.8%	5.5%	43.2%
Rented land	9.3%	9.0%	56.8%
Public land only	12.8%	11.1%	53.6%

Source: Author’s calculation

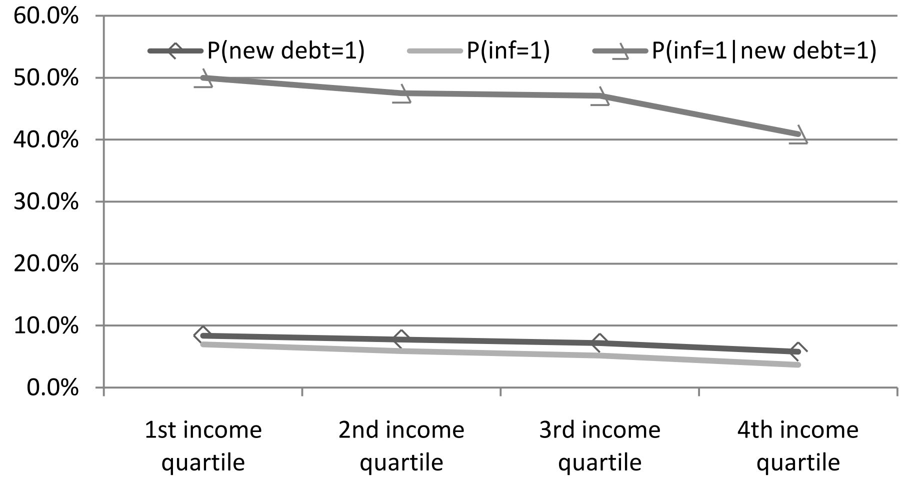


Figure 12 Average predicted probability based on income quartiles

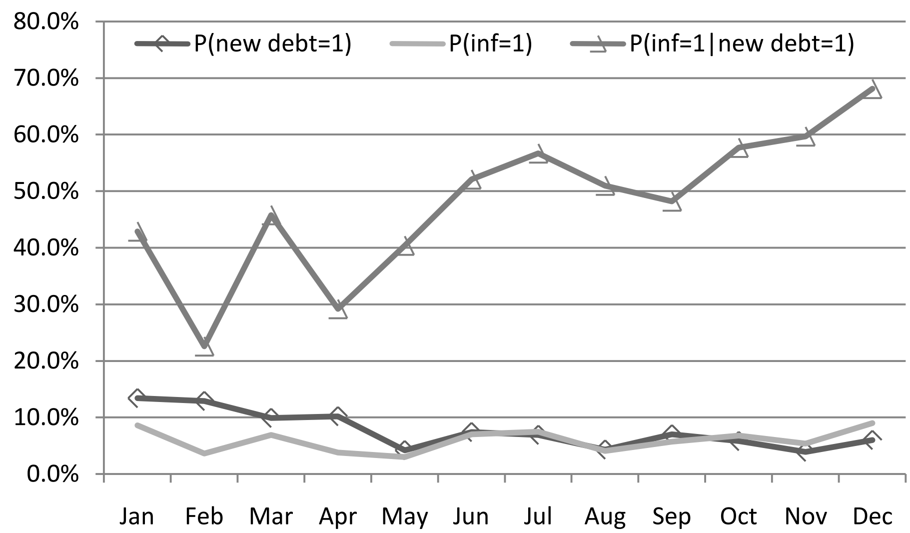


Figure 13 Average predicted probability based on month

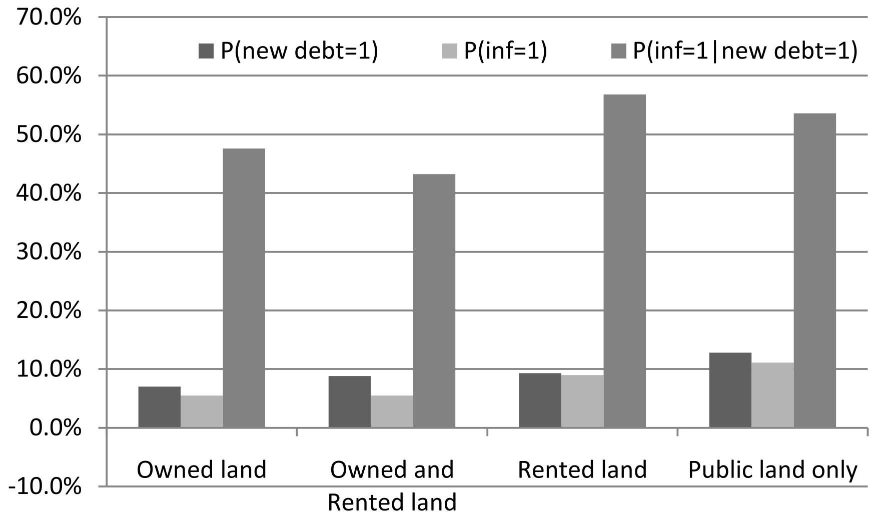


Figure 14 Predicted Probability based on land ownership

8 POLICY IMPLICATIONS

From the previous section, it is clear that the degree of participation in the informal credit market for Thai farmers is still significant. In particular, farmers' increased participation in informal credit conditional on accessing new credit rising toward the end of the year is the most worrisome discovery. This piece of evidence pinpoints the limited role of BAAC in providing working capital only and not consumption-based loans hence underlining the importance of the lack of formal credit facility for farmers, and should be addressed quickly. A policy response, perhaps together with BAAC's loan allocation and other social safety net programs, to boost the role of formal credit channels through other and lower farm operation and household expenses should be considered.

Secondly, level of wealth plays a significant role in determining both the probability of accessing to credit, and probability of participating in either formal or informal credit markets. Table 9 shows that conditional on accessing new credit, the poorest farmers have equal probability of participation in either formal or informal credit, while the richest farmers have 59% probability of participation in formal credit. This demonstrates room for improvement in participation rates of formal credit across the board. A form of micro-finance lending, and other network-lending programs should be considered to boost the allocation of credits through formal channels more efficiently. A new innovative credit delivery system is also encouraged

Third, land ownership is crucial to accessing new credit. What is interesting is that farmers who own and rent land at the same time have a higher probability of accessing new credit but lower conditional probability for participating in the informal credit market. This may indicate the use of owned land as collateral to participate in formal credit channels. Without land ownership, the probability indicates that participation in the informal credit market is more than 50%. The policy discussion should focus on working with the BAAC to establish a framework that will allow farmers to maintain land ownership.

9 CONCLUSION

Using the 2006 Socio-Economic Household Survey of Thailand, this study utilizes Heckman's style selection model to determine loan participation of Thai agricultural households. The result from this model provides supporting evidence that, facing credit rationing, households are more likely to 'stick' to the credit market they previously engaged in, whether the formal or informal credit market. This finding reinforces the vicious cycle which makes it more difficult for farmers to get out of debt. Secondly, the study finds supporting evidence that wealthier households are less likely to access credit, but, if they do, they are more likely to participate in formal credits, and the opposite is true of less wealthy households. Thirdly, the study finds supporting evidence for less credit participation during May and December coinciding with the planting and harvesting season. Specifically, the probability for accessing new credit starts off high in the beginning of the year then slowly declines while participation in informal credit, conditional on participating in new credit, rises toward the end of the year. Alarmingly, this indicates that, while overall access to new credit declines, people who need credit tend to participate in informal credit more as months pass. This piece of evidence pinpoints the limited role of BAAC in providing working capital only and not consumption-based loans underlining the importance of the lack of formal credit facility for farmers, and should be addressed quickly. Finally, the model finds that households with owned farmland are more likely to participate in the formal credit market, while households with rented farmland are more likely to participate in the informal credit market stressing the use of owned land as collateral to participate in formal credit channels.

Future work can incorporate characteristics of informal lenders in order to model the behavior of suppliers of informal credit, although such data is extremely difficult to find. The inclusion of reliable and comprehensive interest rate data will also benefit discussion of cost-benefit analysis of formal and informal credit, such as the calculation of risk premium based on households' characteristics.

REFERENCES

Achen, C. (1986). *The Statistical Analysis of Quasi-Experiments.* Berkeley: University of California Press.

Akoten, J. E., Sawada, Y., & Otsuka, K. (2006). The Determinants of Credit Access and Its Impacts on Micro and Small Enterprises: The Case of Garment Producers in Kenya. *Economic Development and Cultural Change* , 927-944.

Aryeetey, E., & Udry, C. (1995). The Characteristics of Informal Financial Markets in Africa. *Bi-annual Research Conference of the African Economic Reseach Consortium.* Nairobi, Kenya.

Bell, C., Srinivasan, T. N., & Udry, C. (1997). Rationing, Spillover, and Interlinking. *Oxford Economic Papers* , 557-585.

Boonperm, J., Haughton, J., & Khandker, S. R. (2009). *Does the Village Fund Matter in Thailand?* The World Bank Development Research Group.

Gine, X. (2005). *Access to Capital in Rural Thailand: An Estimated Model of Formal vs. Informal Credit.* Wasington, D.C.: The World Bank.

Golder, M. (2008). *Methods IV: Advanced Quantitative Analysis* . Retrieved from http://homepages.nyu.edu/~mrg217/homepage.htm

Greene, W. H. (2003). *Econometric Analysis.* Upper Saddle River, NJ: Prentice Hall.

Guirkinger, C. (2006). *Understanding the Coexistence of Formal and Informal Credit Markets in Piura, Peru.* Belgium: Center for Research in Economic Development, University of Namur.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica* , 153-161.

Hoff, K., & Stiglitz, J. E. (1998). Moneylenders and bankers: Price-increasing subsidies in a monopolistically competitive market. *Journal of Development Economics* , 485-518.

Hoff, K., Braverman, A., & Stiglitz, J. E. (1993). Introduction. In K. Hoff, A. Braverman, & J. E. Stiglitz, *The Economics of Rural Organization: Theory, Practice, and Policy* (pp. 1-30). New York: Oxford University Press, Inc.

Kaboski, J., & Townsend, R. (2009). *The Impacts of Credit on Village Economies.* Boston: Massachusetts Institute of Technology, Department of Economics Working Paper Series.

Kochar, A. (1997). An empirical investigation of rationing constraints in rural credit markets in india. *Journal of Development Economics* , 339-371.

Mid-year Economic Review 2004. (2004, July). Retrieved November 2009, from Bangkok Post: http://www.bangkokpost.com/midyear2004/economy02.html

Paulason, A. L., & Townsend, R. (2004). Entrepreneurship and financial constraints in Thailand. *Journal of Corporate Finance* , 229-262.

Profile: Thaksin Shinawatra. (2009, November 11). Retrieved November 2009, from BBC News: http://news.bbc.co.uk/2/hi/asia-pacific/1108114.stm

Sartori, A. (2003). An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions. *Political Analysis* , 111-138.

Siamwalla, A., Pinthong, C., Poapongsakorn, N., Satisanguan, P., Nettayarak, P., Mingmaneenakin, W., et al. (1993). The Thai Rural Credit System: Public Subsidies, Private Information, and Segmented Markets. In K. Hoff, A. Braverman, & J. E. Stiglitz, *The Economics of Rural Organization: Theory, Practice, and Policy* (pp. 154-185). Oxford University Press.

Stiglitz, J. E., & Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *The American Economic Review* , 393-410.

Thai Rak Thai's Victory: It was Thaksin who made up voters' minds. (2005, February 21). Retrieved November 2009, from Nation Multimedia: http://www.nationmultimedia.com/Election2005/news/news.php?news=02/21020501.htm

CHAPTER FOUR

ON THE CHALLENGE OF TESTING WEAK-FORM MARKET EFFICIENCY USING HIGH FREQUENCY DATA

1 INTRODUCTION

The increased availability of high frequency intraday trading data has made it possible for researchers to explore the microstructure of financial markets. Due to the sheer quantity of data and limited ability to manipulate and control for inherent errors, and outliers, the data-cleaning procedure of the high-frequency or intraday dataset becomes one of the most difficult tasks microstructure researchers face. Lacking the appropriate treatment to screen the data, subsequent analyses based on it are at high risk of misrepresentation. To demonstrate the magnitude to which such misrepresentation can grow, this paper shows that running a simple technical trading strategy on improperly treated data can generate excess profits which are large, significant, and robust. While this could be misinterpreted as another refutation of the *Weak Form Efficient Market Hypothesis* (Fama, 1970), this paper shows that more stringent data cleaning measures also eliminate these profits.

To begin, Joel Hasbrouck, a finance professor at New York University and a key pioneer in the use of the Trade And Quotes dataset (TAQ), in a 2003 article in the Journal of Finance (Hasbrouck, 2003), presented a "$0.5 filter" treatment to clean TAQ data for a relatively new class of security called the Exchange Traded Fund, or the ETF of the S&P 500 (SPDR). We show that the method implemented by Hasbrouck (2003) does not remove enough mistakes and outliers to be reliable. This is critical as we find that a simple technical trading strategy yields spectacularly abnormal returns when the data used is filtered with Hasbrouck's $0.5 filter. However, when a more conservative filter is applied, these abnormal profits go away. We further show precisely how the profitability of this simple strategy is dependent on the filtering criteria. If one believes that the S&P 500 ETF is weak form efficient, that the current price fully reflects all available information, then a simple technical strategy based on past price patterns should not yield abnormal returns. Thus, the fact that data cleaned by Hasbrouck's filter yields abnormal returns is evidence that this filter is not appropriate for testing weak form market efficiency. This filter is simply too lax or too wide. We also provide graphic evidence that clearly identifiable outliers and mistakes are still visible on the charts of prices that were filtered with the Hasbrouck (2003) filter. However, these outliers are removed when we use a more appropriately stringent filter. These findings cast a shadow on all papers that have already used the Hasbrouck (2003) filter. They also represent a cautionary tale for researchers who are interested in testing the weak form efficiency of high frequency prices.

Following Hasbrouck (2003), this paper uses the S&P 500 ETF as its main vehicle of study. From its inception in 1993, investment in Exchange-Traded Funds, or ETFs, has been associated with low cost and highly diversified investment. ETFs were created to provide retail investors opportunities to purchase a single security representing diversified baskets of stocks, similar to an index itself. Accordingly, the SPDR trust contains the same proportionate weighting of all common stocks that comprise Standard and Poor's 500 Composite Price Index. In addition to being the first of its kind and having the longest price record of all ETFs, the SPDR is also the most traded ETF to date. The SPDR's success has led to the establishment of many similar index products.

The Efficient Market Hypothesis (EMH) (Fama, 1970) conjectures a price sequence for a security following a sub-martingale with respect to the available information set, with a main implication that patterns in market prices cannot be reliably used to make greater expected profits than a buy-and-hold strategy. Three sufficient but not necessary conditions to achieve frictionless market efficiency are i) zero transaction costs in trading securities, ii) all available information is accessible without cost to all market participants, and iii) all agree on the implications of current information for current price and distributions of future prices of each security. Accordingly, empirical work that demonstrates profitability of trading system that can outperform the buy-and-hold strategy is a direct evidence against market efficiency.

Even before Fama's (1970) paper gained momentum, there were attempts to test this weak form market hypothesis. To mention an important few, Alexander (1961, 1964) implemented a trading system on Standard & Poor's Industrials, the Dow Jones Industrial Average using daily data during 1928-61. Alexander's filter rule gives a buy signal if the price has risen x% from the most recent bottom, and gives a sell signal when the price has dropped x% from the most recent peak. This is seen today as a momentum strategy. Alexander's result was positive out-performance, compared to the buy-and-hold (B&H) strategy, but profitability was significantly reduced when a transaction cost of 2% for each round-trip was included. Fama and Blume (1966) repeated Alexander's (1964) results and found only three small filter rules[21], 0.5, 1,

[21] This filter rules trading algorithm is not to be confused with the filtering technique this study uses to clean data.

and 1.5%, which yielded higher gross average returns of 11.4%-20.9% per year per security. However after transaction costs were included, the returns became similar to buy-and-hold returns. While early empirical results seem to support weak form market efficiency, more recent attempts that employ advanced econometrics methods and intraday dataset yield mixed results. Farrell and Olszewski (1993) used a nonlinear trading strategy based on the Autoregressive Moving Average (1,1) model on daily data which includes a 0.024% transaction cost per round-trip on S&P 500 futures. The result was slightly more profitable than the B&H strategy, but statistically insignificant. Levich and Thomas (1993) applied filter rules and a moving average strategy on daily International Monetary Market (IMM) currency futures during 1976-1990 with 0.025% and 0.04% transaction costs per one-way. They found that their filter rules and moving average rule generated substantial positive mean net returns for all currencies but the Canadian dollar. And, for both trading systems, the null hypothesis that there is no information in the original time series was rejected in 25 of 30 cases. A more exhaustive review of the literature involving technical trading can be found in Park and Irwin (2004).

Back to the study related to ETF, Hasbrouck (2003) studied price discovery among floor-traded index futures contracts, exchange-traded funds (ETFs), electronically traded, small denomination futures contracts (E-minis), and sector ETFs. Using a Vector Error Correction Model (VECM), he found that prices among ETFs, E-minis, and futures contracts in the S&P 500 Index converge rapidly within a five to ten minute window after a shock for each security. The evidence of these small windows led to the conclusion that a high degree of substitutability exists among these ‘similar’ products and that prices react quickly to offset any arbitrage opportunity. This finding also casually supports market efficiency.

In summary, the main contributions of this paper are three fold. First, the analysis shows that the profitability of a simple technical trading strategy depends critically on filtering the data. Secondly, the analysis shows that the wider 50 cent filter of Hasbrouck (2003) is not appropriate to use because it leaves in too many outliers and mistakes. This point is not apparent in early research since the availability of high frequency data is relatively limited. In fact, using this data filter, the spectacularly positive out-performance for a simple technical strategy is everywhere. However, the positive region is not robust as we treat the data with lower filters of 0.3% and 0.2%. Third, all previously profitable strategies disappear with the 0.1% filter, the smallest filter we use, and the most appropriate, in my view. This shows that the 0.1% filter is not too wide. When risk is factored in, the buy-and-hold strategy remains superior to any of our technical trading strategies, and this result is robust across all three sub-periods. Having much of the profitability hinging on a wide filter level casts a shadow over the validity of results from studies that employed a wider band, such as 50 cents.

The rest of this paper is organized as follows. Discussion of data is given in the next section. Section three describes transaction costs, trade algorithms, measures of performance, and the filtering method. Section four presents results. Section five concludes the study.

2 DATA

The main analysis in this paper uses price series of the S&P 500 ETF from the Trades and Quotes (TAQ) database, which contains high frequency intraday transaction data (trades and quotes) for all securities listed on the New York Stock Exchange, American Stock Exchange and NASDAQ. From the TAQ dataset, second-by-second price series of an ETF of the S&P 500 were created for the regular floor trading session, from 9:30 am to 4:00 pm, for the period between January 29, 1993 and December 31, 2006. To reflect an actual "real-time" trading data set, or second by second data, this study aggregates the trade transactions of the same second and calculates the weighted average price of that second as well as filling-in any missing seconds with the previous active price traded.

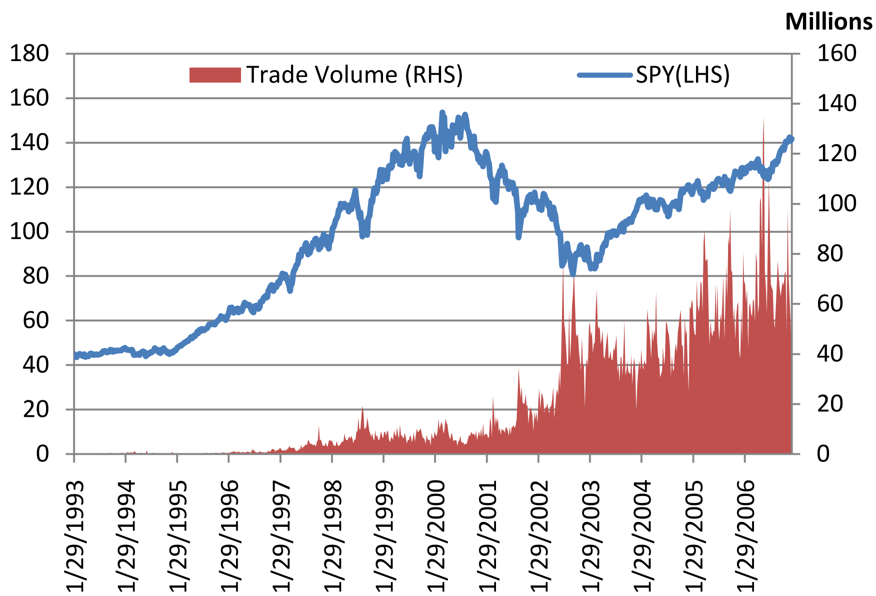


Figure 1 S&P 500 ETF price and average trade volume

This study focuses on the issue of eliminating mistakes and outliers in the TAQ dataset since TAQ is notorious for containing a high number of erroneous transactions. Many filtering criteria will be examined in this study. First, a filtering method proposed in Hasbrouck (2003) is used. It is aimed at removing exchange-identified erroneous trades as well as any outlier trades that differ by more than 50 cents from a centered moving average over the nearest 10 second prices. The Hasbrouck (2003) technique of filtering TAQ with a fixed 50 cent band during April and May of 2003 is equivalent to a 0.35% filter. As this turns out to have considerable impact on the results, this analysis will elaborate further on the filtering technique in section (4.1). A different result after adjusting the data with a new filtering treatment is offered in section (4.7).

Benchmark monthly returns, which consist of holding period returns, value-weighted portfolio returns, equal-weighted portfolio returns, and S&P 500 composite returns were obtained from the Center for Research in Security Prices (CRSP). The next section discusses the details of the trading strategy, including buying rules and selling rules.

3 METHODOLOGY

This section describes all components of the trading system underlying the analysis, starting with transaction costs in section (3.1) and (3.2), as vital elements in determining profitability of transactions. Elaborations of mechanics in trading strategies are given in section (3.3) and (3.4). Section (3.5) examines calculation of returns from the strategy. Section (3.6) discusses risk-adjusted returns, or the Sharpe ratio. Section (3.7) examines filtering techniques to clean TAQ data.

3.1 ANALYSIS OF TRANSACTION COSTS

Transaction costs play a crucial role in determining the real cost of trade. This analysis breaks down transaction cost into two major components: 1) the Bid/Ask spread and 2) the brokerage commission fee.

Trade prices, as opposed to quotes, are used in this study. Trade prices could result from a buyer initiated trade or from a seller initiated trade. This, however, is not identified in the TAQ database. Therefore, there is a need to account for the bid/ask spread. When trade prices are used to buy and sell, there are 4 possibilities. First,

i. *Buy at a Bid price, and Sell at a Bid price: in such a case, the return is inflated by 1 spread.*

ii. *Buy at a Bid price, and Sell at an Ask price, then the return is inflated by 2 spreads.*

iii. *Buy at an Ask price, and Sell at an Ask price; the return is inflated by 1 spread.*

iv. *Buy at an Ask price, and Sell at a Bid price. This is the correct case. Therefore, the return reflects actual value.*

Based on these possibilities, on average, one spread should be deducted from the round-trip profit of each transaction. In addition, a fixed minimal commission of 2 cents for each round-trip is used for transaction costs. For a reality check, Interactive Brokers charges $0.01 for each transaction for each share. TD Ameritrade charges a flat fee of $10 per each trade. For trade of 1000 ETFs, commission per share = $10/1,000 = $0.01.

A major event that directly impacts trading spread took place on January 29, 2001. The American Stock Exchange approved a decision to decimalize Exchange Traded Funds, for which the main implication is that quotes were to be reported in decimals instead of fractions, such as 1/32. This decision has led to more flexibility in trading and a more favorable trading environment, as the next section will show in more detail.

3.2 ANALYSIS OF BID/ASK SPREAD

The analysis in this section is based on average daily spreads of the S&P 500 ETF derived from the CSRP. Table 1 shows the evolution of the Bid/Ask spread's basic statistics over the time period studied. The distribution of the Bid/Ask spread is highly and positively skewed; thus, the analysis that follows renders mode as a representation of the spread distribution. An interpretation of means would be misleading, as they suffer from extreme price posting by market makers in later years. Analysis using medians is parallel to that for mode.

Figure 2 shows that spreads were stable at around 3 cents in an absolute term from 1993 to 1995. But, from 1995 to 2000, the spread quickly rose from 3 cents to 25 cents, an 833% increase while prices increased from 54.8 to 142.2, or a 259% increase, a far smaller change, over the same period. The asymmetric increase in spread and price unveiled unique characteristics of this market microstructure, as quotes are trivially based on 1/32 price system.

Table 1 Summary of Bid/Ask spread of S&P500 ETFs from 1993-2006

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Bid/Ask Spread (in $US)														
Median	0.03	0.03	0.03	0.06	0.16	0.23	0.25	0.25	0.1	0.05	0.03	0.02	0.02	0.02
Mode	0.03	0.03	0.03	0.06	0.13	0.19	0.19	0.25	0.08	0.03	0.03	0.01	0.02	0.01
Interquartile Range	0.00	0.00	0.00	0.03	0.06	0.20	0.23	0.23	0.09	0.03	0.02	0.02	0.02	0.02
Bid/Ask Spread (in 1/32 $US)														
Median	1.00	1.00	1.00	2.00	5.00	7.50	8.00	8.00	3.20	1.60	0.96	0.64	0.64	0.64
Mode	1.00	1.00	1.00	2.00	4.00	6.00	6.00	8.00	2.56	0.96	0.96	0.32	0.64	0.32
Bid/Ask Spread (% of Price)														
Median	0.07%	0.07%	0.11%	0.09%	0.18%	0.22%	0.19%	0.18%	0.08%	0.05%	0.03%	0.02%	0.02%	0.02%
Mode	0.07%	0.07%	0.11%	0.09%	0.14%	0.17%	0.14%	0.18%	0.07%	0.03%	0.03%	0.01%	0.02%	0.01%
Closing price* ($US)	45.4	46.1	54.8	67.5	87.6	108.8	133.1	142.2	118.8	99	97.2	113.8	120.9	132

*Closing price is an average 12-month closing price of S&P500 ETF.

At the height of the dot-com bubble in 2000-01, prices peaked and then began to drop sharply. As average prices dropped 14% in 2001 from 2000, the spread dropped from 25 to 8 cents, a 68% decrease over the same period. Although an asymmetric decrease could play a role this time, the decline in the spread is also likely to be partially driven by decimalization that took effect in January 2001. This became more evident when prices started rising again in 2004 but the spread went the opposite way, declining even further to only one cent. Based on this finding, the decision to decimalize quotes has proven to be a productive measure for reducing transaction costs and smoothing market activities.

Figure 3 shows spread as a percentage of price. The spread has varied from 0.07% of price in 1993 to a height of 0.18% in 2000 and down to 0.01% between 2004 and 2006. In our trading simulation, we allow the bid/ask spread to vary every year.

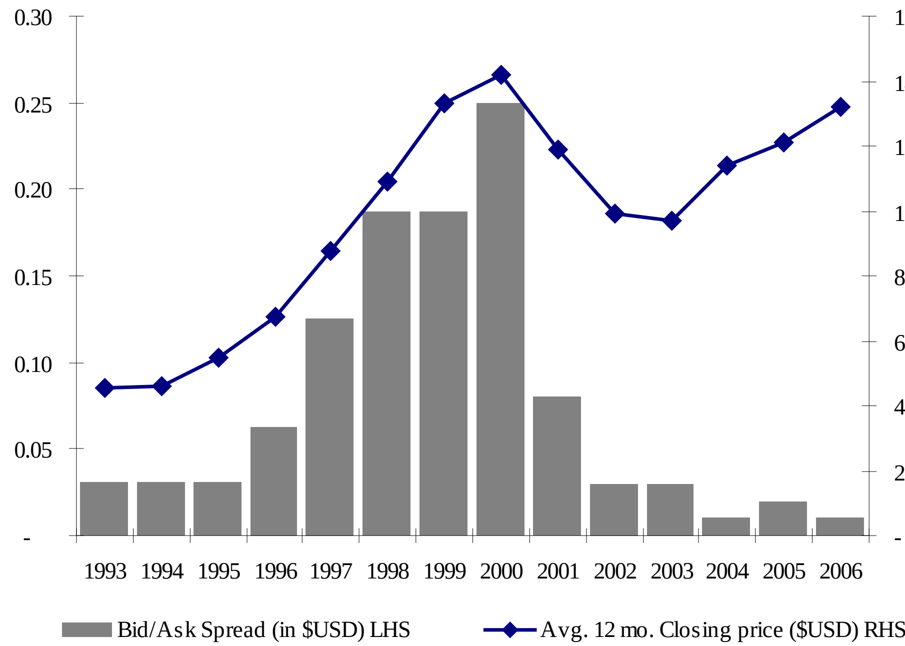


Figure 2 Modes of Bid/Ask Spread of SPDR from 1993 to 2006 in ($USD)

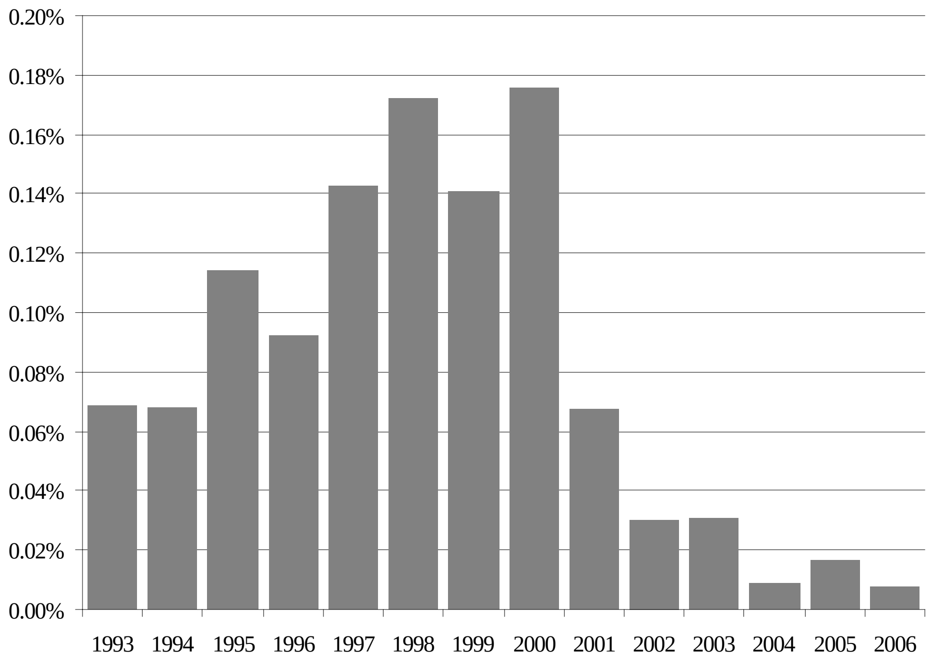


Figure 3 Modes of Bid/Ask Spread of SPDR from 1993 to 2006 in (% of SPDR Price)

3.3 SIMPLE ALGORITHM

This section describes the construction of a simple trading rule: selling when prices rise and buying back when prices drop. The investor can hold cash and one security. Since the maximum number of securities is restricted to one unit, the investor can place a buy order again only after he closes his previous position. And he will close this position only when a predetermined profit margin has been reached. Shorting is not allowed.

Consider the price that an investor pays for the ETF is P^b, with a selling rule $S = 0.04\%$, a sell order will go through if there exists a price P such that,

$$P \geq (1+0.04\%)P^b.$$

Let us call this a selling price P^s. Then the investor realizes the gain of P^s minus P^b and any costs associated with the trade. Now, the investor is looking for another buying opportunity to long or add a position. He will place another buy order when the current price is lower by a certain amount from the most recent price peak, defined as local maximum price between the last sell order and the execution of a new buy order.

That is, for a price at time t, P_t,

$$P^{max} = Max(P_{t-j}\ , P_{t-j+1}\ , \ldots, P_t)$$

Consider for a buying rule $B = 0.05\%$, a buy order will go through if there occurs a price P such that,

$$P \leq (1-0.05\%)P^{max}.$$

This price becomes the buying price or P^b.

The price peak is reset after a sell order is executed. That is, following a new P^s,

$$P^{max} = P^s .$$

Then another buying cycle begins with,

$$P^{max} = Max(P_{t-j} , P_{t-j+1} , \ldots, P_t).$$

This investor is forced to close his open position by 4:00 pm on the last trading day of the month to make the comparison conform to a buy-and-hold strategy. It should be noted at this point, following a successful buy order, that if the price never rises above the triggering point, the investor will hold the position regardless of how much unrealized losses have accumulated. He will never sell. Usually investors have a different time frame for cash needs; in this case, the model assumes that the investor has the ability to withstand a long period of market downturn. The first buy order is initiated arbitrarily at 10.30am, an hour after the opening of active trading hour, on 29 January, 1993.

To make this trading as close to reality as possible, this program also imposes special requirements that every buy and sell order needs to be executed on the price that has supporting market depth for the transaction to go through. A graphic illustration is given in Figure 4.

3.4 ALGORITHM GRID

The set of rules for buying and selling are from 0.05% to 2% for buying and from 0.04% to 2% for selling. The full set of the buying and selling rules are given here:

B = {0.05%,0.1%,0.15%,0.2%,0.25%,0.3%,0.5%,1%,1.5%,2%}

and

S = {0.04%,0.06%,0.08%,0.1%,0.12%,0.14%,0.16%,0.18%,
0.2%,0.3%,0.4%,0.5%,0.6%,0.7%,0.8%,0.9%,1%,2%}.

Altogether this gives us about 180 single strategies that account for each combination of buying and selling rules. This paper restricts the upper-bound to 2% to focus on the intraday trading window. The range beyond 2% is left for future exploration. To simplify future notations, the following pattern is applied. Strategy *(B,S) = (0.1%,0.5%)* refers to the strategy that imposes a buying rule of 0.1% drop from the previous nearest peak and a selling rule of 0.5% profit margin. With enough combinations, we will be able to draw a 'region' of performance, as well as determine the direction of higher performance.

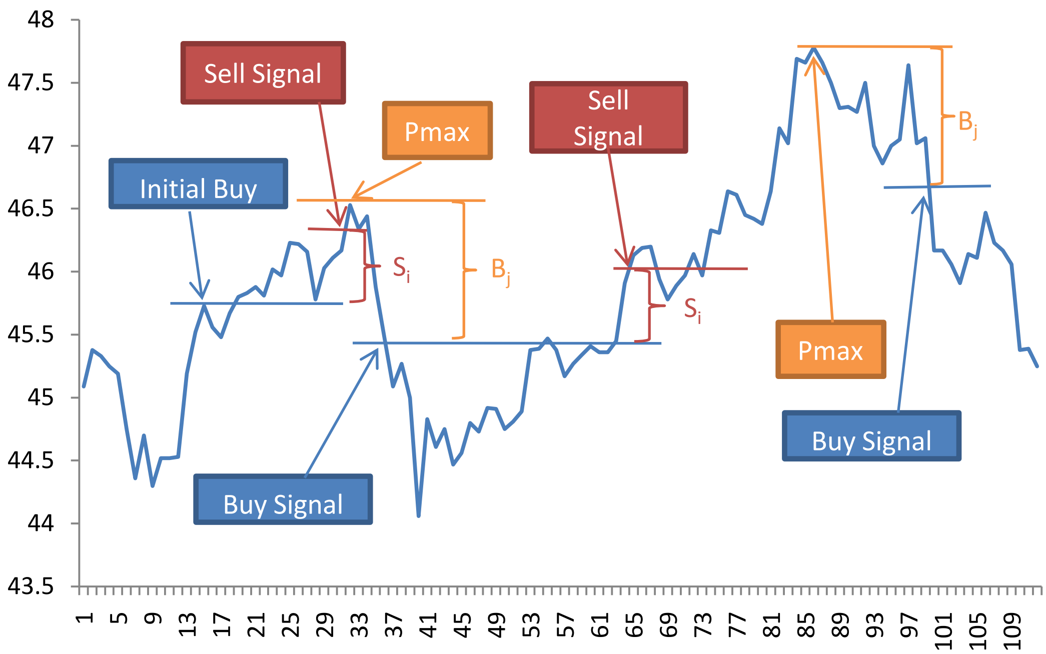


Figure 4 Illustration of Trading Strategy (Bi, Sj). The gap between Buying and Selling Price reflects Profit Margin or Selling Rule. The gap between the maximum price and the buying price reflects Buying Rule.

3.5 CALCULATION OF RETURNS

A return is the change in the total value of an investment in a common stock over some period of time per dollar of initial investment. The benchmark monthly return series, or monthly returns of the buy-and-hold strategy (including dividends) is the difference of the current month's closing price and the previous one plus a dividend then divided by the previous month's closing price. This series was obtained from the CRSP data. The construction of this series is given below.

For month t, let:

t-1 = time of last month's closing price

$R_{(t)}(B\&H)$ = return on purchase at t-1, sale at t

$P_{(t)}$ = last sale price or closing bid/ask average at current month

$d_{(t)}$ = cash adjustment for month t, usually a dividend

$P_{(t-1)}$ = last sale price or closing bid/ask average at time of the previous month

then,

$$R_t(B\&H) = \frac{P_t + d_t}{P_{t-1}} - 1$$

Next, monthly net returns of any particular combination of buying rule and selling rule (Bi, Sj), are created by summing all differences in buying and selling price from each round-trip transaction minus a spread and round trip commission fees, then divided by the buying price of that transaction, to make it more comparable to the benchmark return.

For month t, let:

$R_{(t)}(B_i,S_j)$ = return of the strategy *(Bi,Sj)* of month t

$n(t)$ = a round trip transaction n of month t

$P^{b}_{n(t)}$ = 'buying' price of transaction n of month t

$P^{s}_{n(t)}$ = 'selling' price of transaction n of month t

$s_{(t)}$ = spread between bid price and ask price of month t

c = round trip commission fees

then,

$$R_t(B_i, S_j) = \sum_n \frac{P^s_{n(t)} - P^b_{n(t)} - s_t - c}{P^b_{n(t)}}$$

Substituting Bi, and Sj with a different combination of buying rules and selling rules gives a unique algorithm that generates returns based on that particular combination.

3.6 SHARPE RATIO

This section analyzes the risk and return trade-off of the strategy *(Bi,Sj)* using an Ex Post Sharpe ratio that closely follows Sharpe (1994). This performance measure uses historic data to determine the best investing strategy. Validating this connection requires that historic results can be reconciled with future performance.

From this construction, the Sharpe ratio measures the excess return from the strategy to rate of return from a risk-free asset, per unit of risk, or standard deviation.

Let,

t = Month t

Rf_t = Rate of return from risk-free asset

$R_t\,(Bi,Sj)$ = Rate of returns from strategy *(Bi,Sj)* in the month t

$R(Bi,Sj)_t - Rf_t$ = excess return in month *t*,

$$E[R_t(B_i,S_j) - Rf_t] = \frac{1}{T}\sum_{t=1}^{T}(R_t(B_i,S_j) - Rf_t)$$

σ_D = *standard deviation of the excess return*

$$\sigma_D = \sqrt{\frac{\sum_{t=1}^{T}\{[R_t(B_i,S_j) - Rf_t] - E[R_t(B_i,S_j) - Rf_t]\}^2}{T-1}}$$

then, an ex post, historic Sharpe Ratio is

$$S_h = \frac{E[R_t(B_i,S_j) - Rf_t]}{\sigma_D}.$$

An increase in returns from algorithm *(Bi,Sj)* or decrease in standard deviation of excess return, will increase the Sharpe ratio. A higher Sharpe ratio suggests higher reward-to-risk strategy.

3.7 FILTERING TECHNIQUE

TAQ data is notorious for containing mistakes and outliers, as shown in Figures 14, 18, and 21. To the naked eye, these outliers represent trading transactions that have large deviations from the main trend. If these outliers are not treated appropriately, a computerized trading algorithm could potentially use these outliers and create an illusion of 'unrealistic' returns.

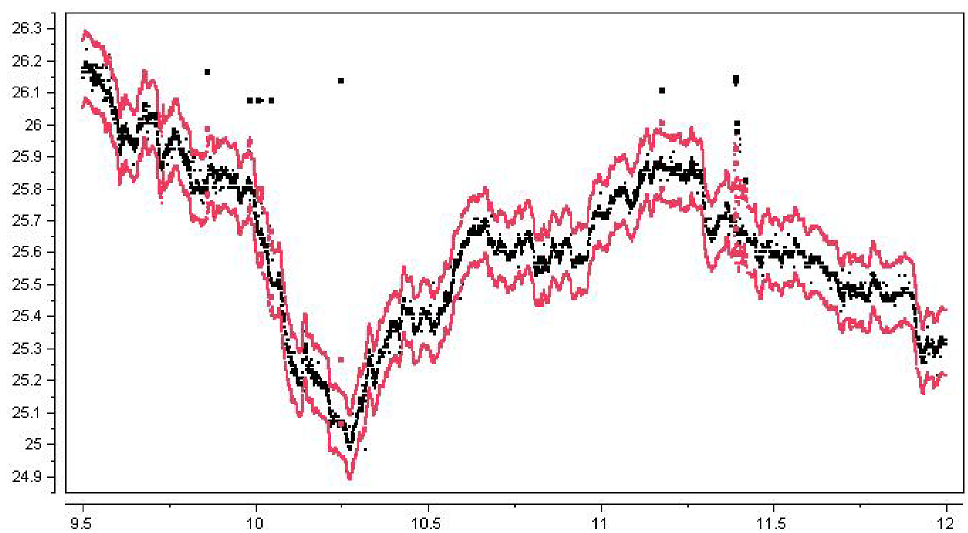


Figure 5 Illustration of Filter Band around moving average

This study examines filters based on narrower bands than that of Hasbrouck (2003). Here filters are studied that screen any price outlier that has 0.3%, 0.2%, or 0.1% price deviation from the central moving average of the nearest 10 prices, or 11 prices with the center price included. A graphic example of how the band works is depicted in Figure 5. The price series is plotted second by second. And the outliers can be clearly seen at prices much higher than the price curve (isolated points). The filter draws a band around the price curve (the 11 second moving average, to be precise). Then, any price that is outside of the band is labeled as erroneous and removed.

To aid readers, and demonstrate how different levels of filter band can affect microstructure research, a graphic presentation of an actual filter-removed-outlier is given in Figures 14 to 17, Figure 18 to 21, and 22 to 25. Three sets of daily prices on December 17, 2002, December 26, 2002, and May 11, 2005 show price movement and outliers removed from three different filters, starting with the most generous 0.3%, then 0.2%, and, finally, 0.1%. For the following figures, an outlier that is identified by the filter will be marked by an "X" symbol. The SPDR price series is represented by a circle symbol.

Based on visual inspection of the scatter plots, it becomes obvious that the 0.3% filter does not adequately remove outliers. However, a filter that is less than 0.1% is not investigated since the 0.1% filter removed the large majority of outliers and a narrower filter might remove too many valid observations. The 0.1% filter seems to be a more appropriate filter to 'clean' the TAQ data than that adopted by Hasbrouck (2003). The downside of our narrower filter is that during a high price volatility period, it could misidentify a 'real' price jump as an outlier. Based on the evidence presented in this section, it is clear that Hasbrouck's (2003) technique of filtering TAQ with a fixed 50 cent band during April and May of 2003, an equivalent of 0.35% Filter, might be too wide. The next section will confirm this hypothesis.

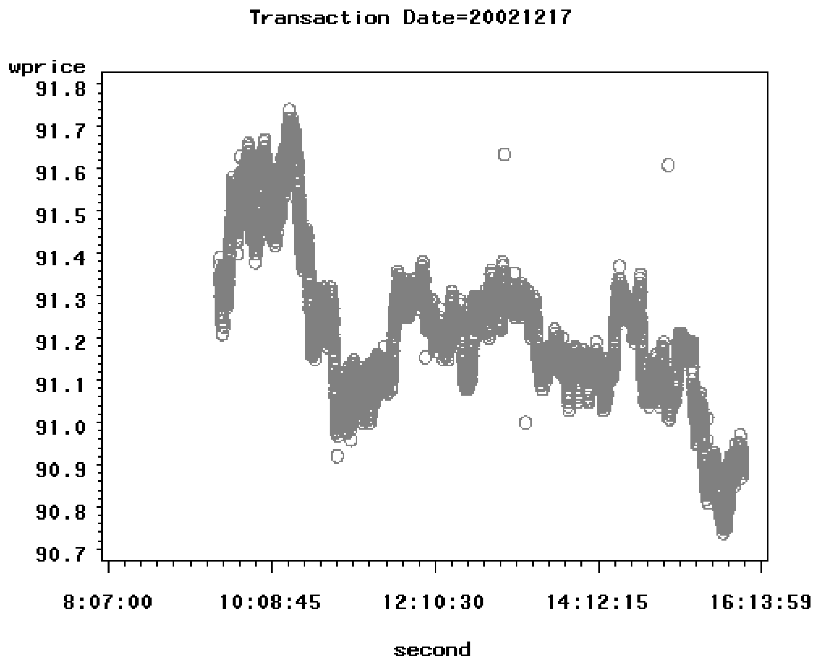


Figure 6 SPDR Price without Filter

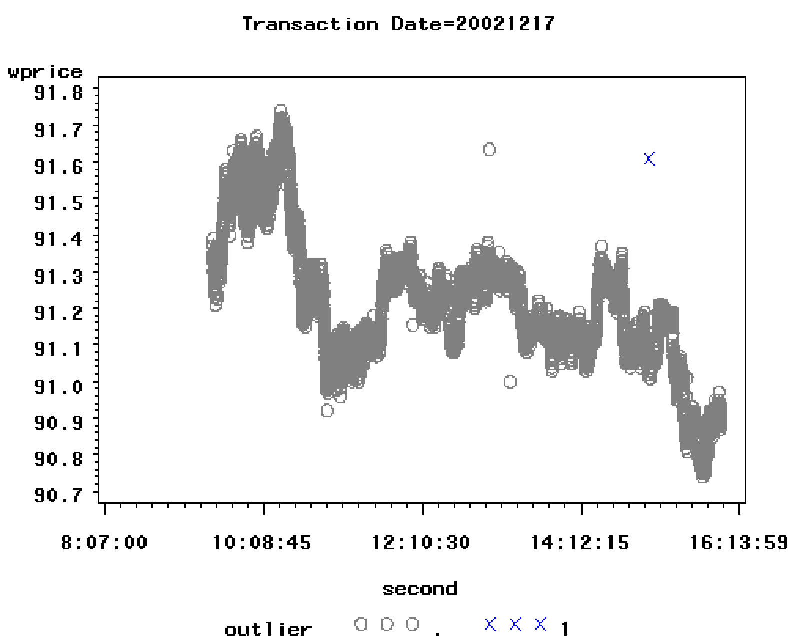


Figure 7 SPDR Price with 0.3% Filter

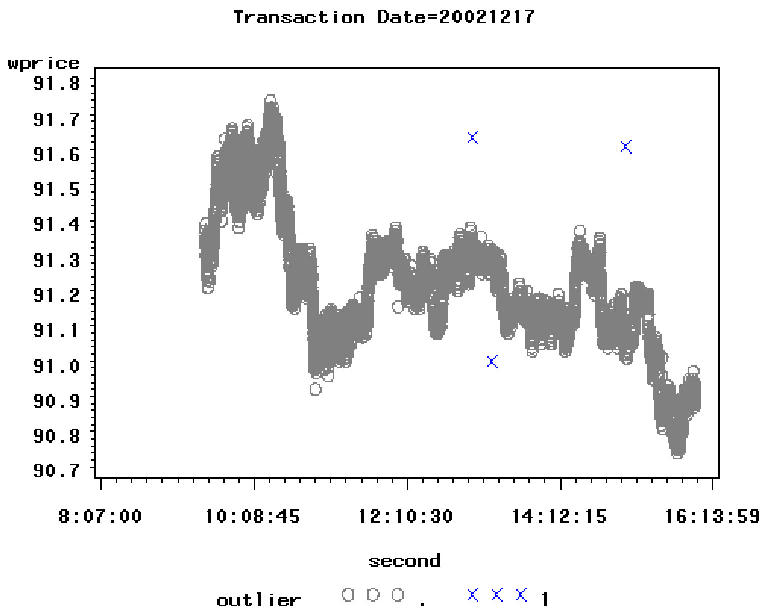


Figure 8 SPDR Price with 0.2% Filter

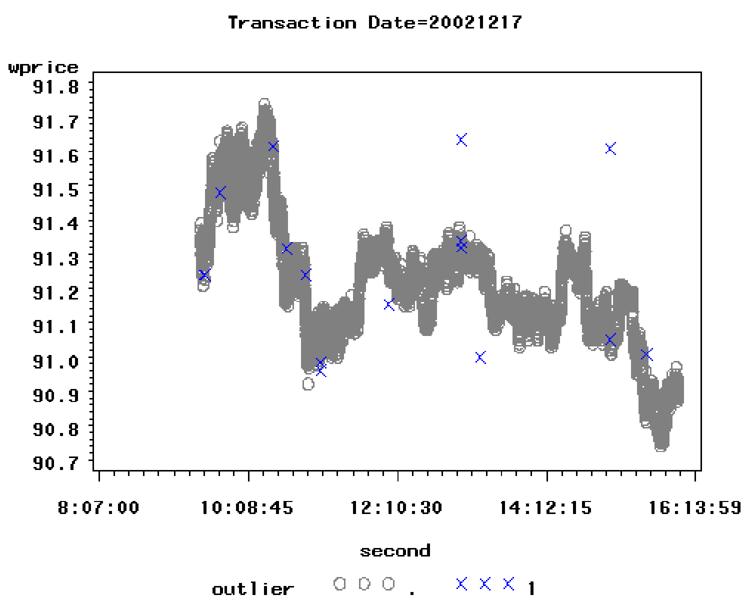


Figure 9 SPDR Price with 0.1% Filter

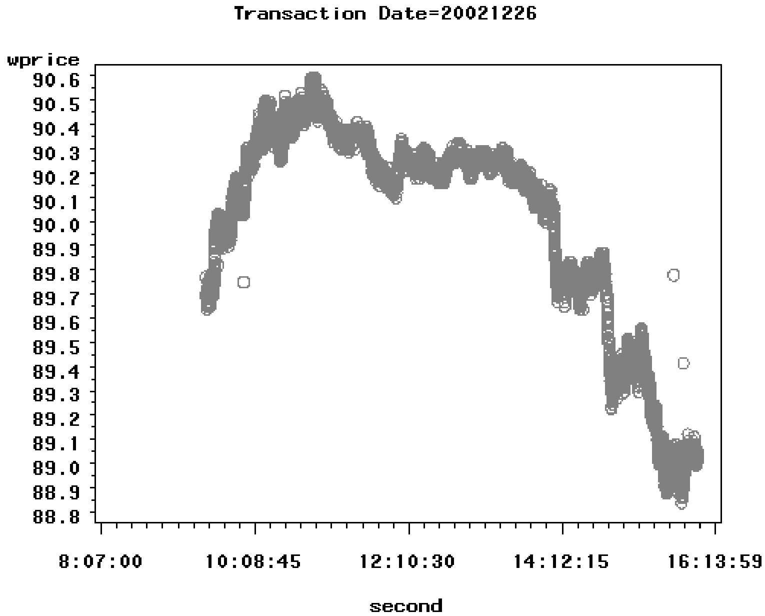


Figure 10 SPDR Price without Filter

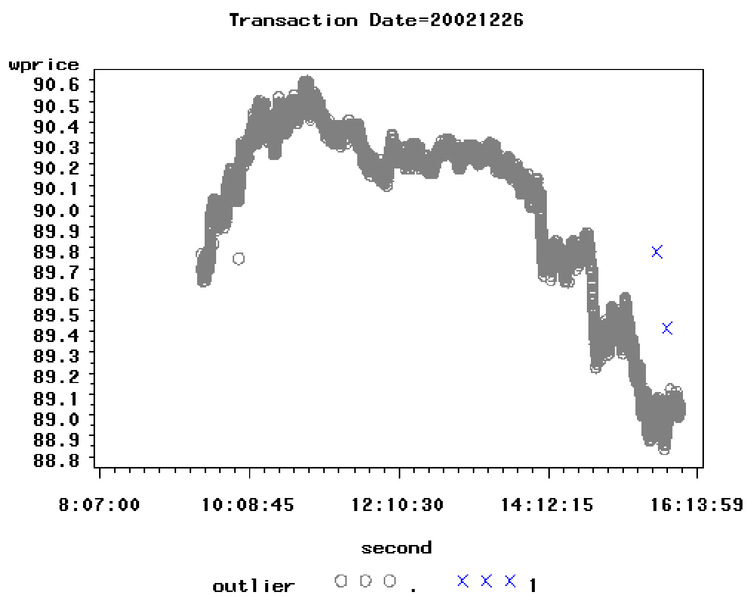


Figure 11 SPDR Price with 0.3% Filter

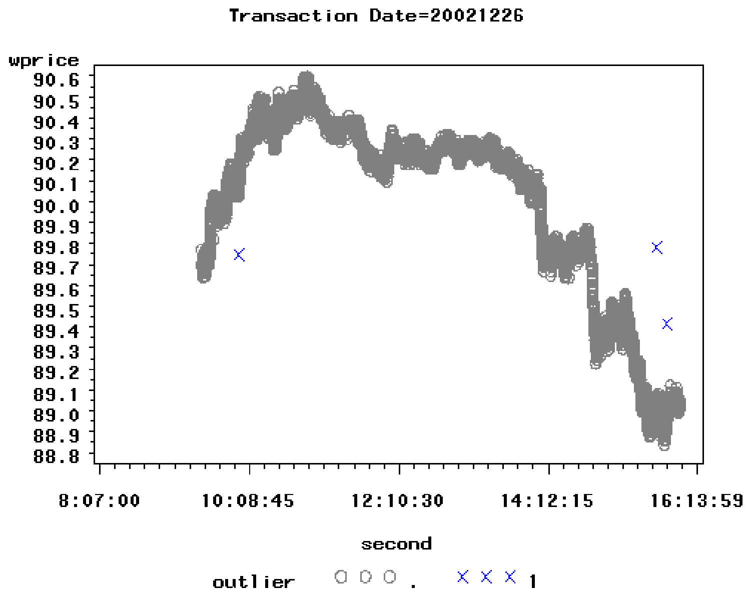


Figure 12 SPDR Price without 0.2% Filter

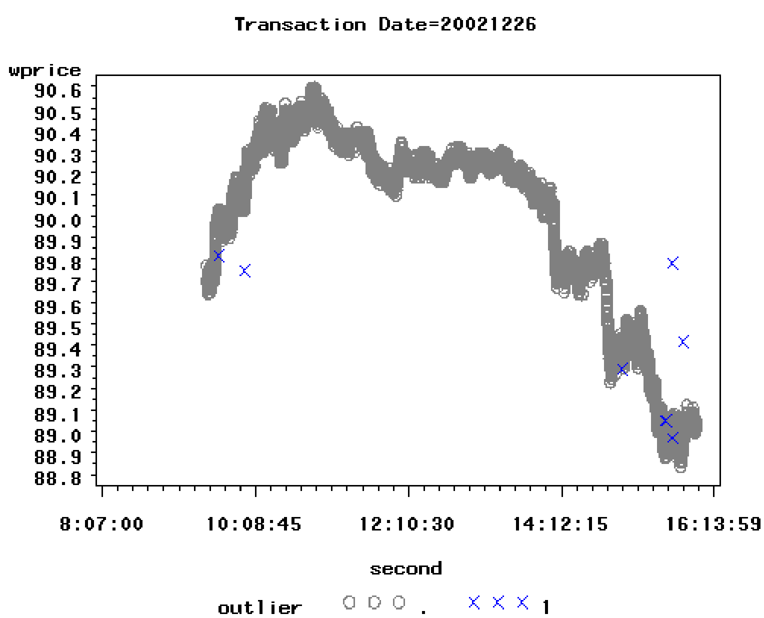


Figure 13 SPDR Price with 0.1% Filter

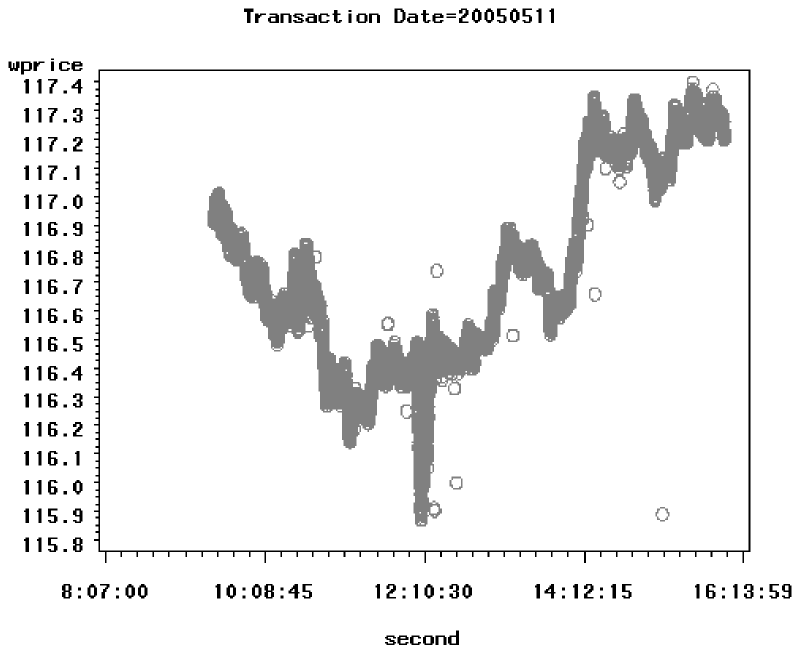


Figure 14 SPDR Price without Filter

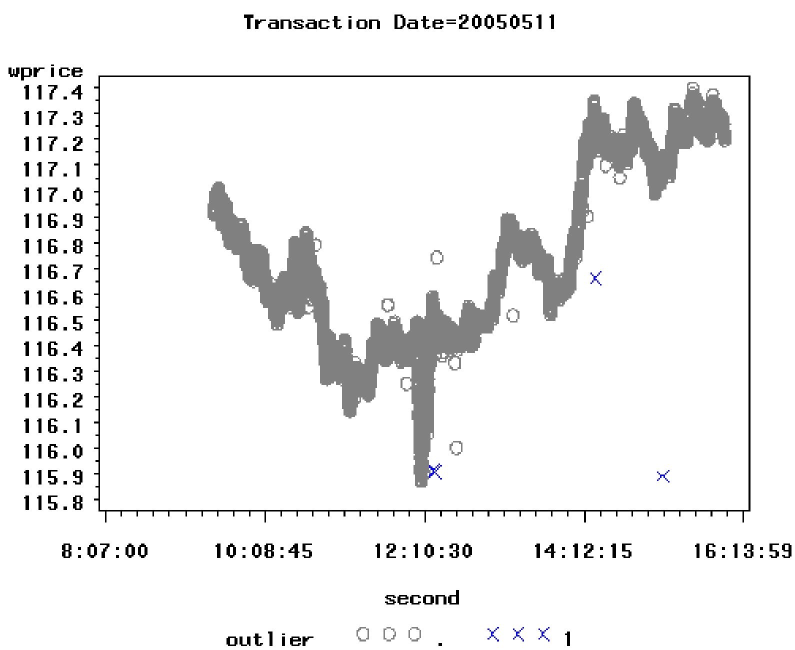


Figure 15 SPDR Price with 0.3% Filter

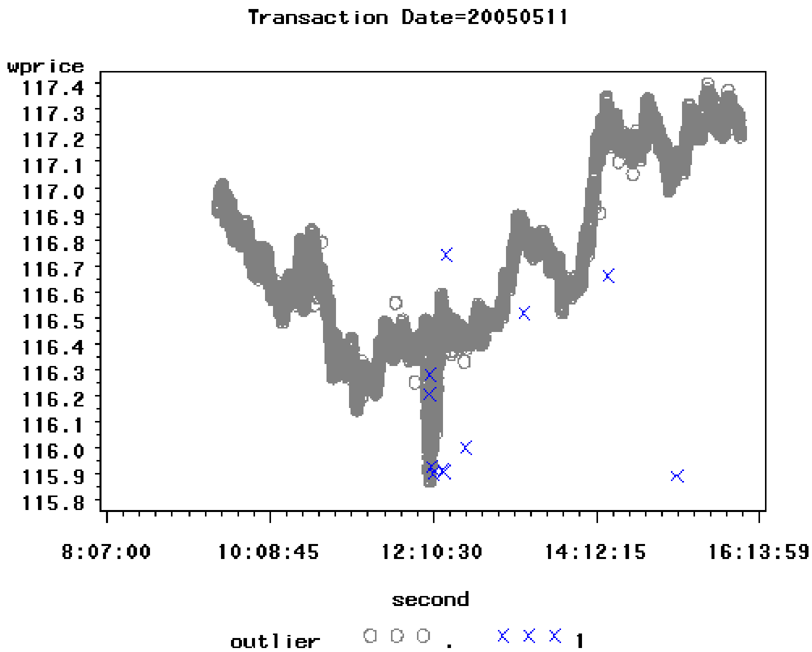


Figure 16 SPDR Price with 0.2% Filter

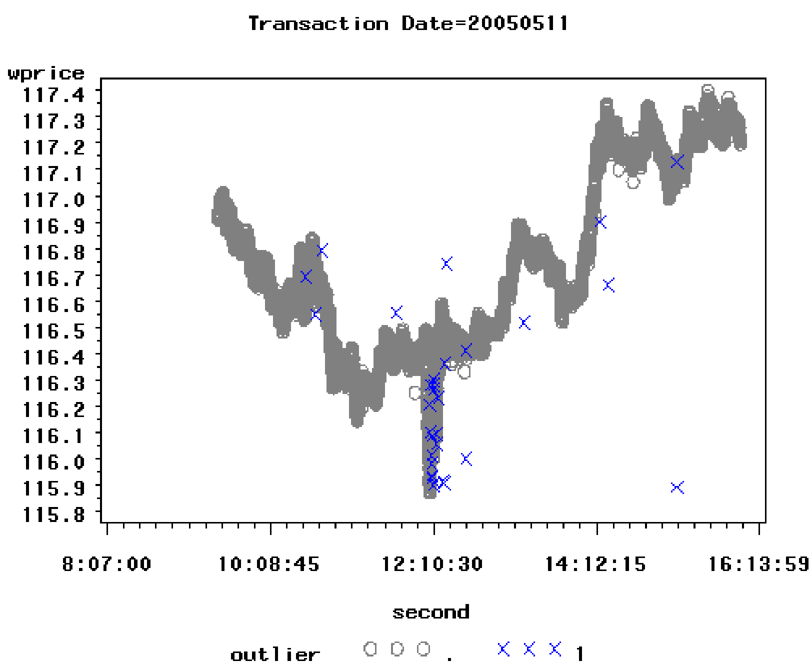


Figure 17 SPDR Price with 0.1% Filter

4 RESULTS

The main findings in this section are three-fold. First, the analysis shows that the profitability of a simple technical trading strategy depends critically on filtering of the data. Second, the analysis show that the wider 50 cent filter of Hasbrouck (2003) is not appropriate to use because it leaves in too many outliers and mistakes. In fact, using this data filter, this analysis found spectacularly positive out-performance for the simple technical strategy. However, the positive region is not robust when the data is treated with lower filters of 0.3% and 0.2%. Third, all previously profitable strategies disappear with the 0.1% filter, the smallest filter we use, and the most appropriate, in our view. This shows that the 0.1% filter is not too wide. When risk is factored in, the buy-and-hold strategy remains superior to any of the technical trading strategies, and this result is robust across all three sub-periods. Having much of the profitability hinging on a wide filter level casts a shadow on the validity of results from studies that employed a wider band, such as 50 cents.

The organization of this part is as follows. Section (4.1) visits returns and out-performance based on the 50 cent (0.35%) filter in the three periods we study. Section (4.1.1) to (4.1.5) analyzes a characteristic of a profitable strategy based on the 50 cent filter. Section (4.2) re-visits returns and shows the disappearance of the out-performance region as the lower filter criteria of 0.3%, 0.2%, and 0.1% are used.

4.1 RETURNS FROM THE 50 CENT (0.35%) FILTER

The dynamic of transaction costs from section (2.2) suggests that further analysis should break the time studied into three periods, that is, 1993-1996, 1997-2001, and 2002-2006. Transaction costs have varied substantially among these sub-periods, both in terms of absolute value, as well as relative to price. Breaking the analysis into three sub-periods will illustrate the importance of market structure for high-frequency trading strategies.

Table 2 shows average annualized monthly net returns based on the 50 cent (0.35%) filter, and equivalents of 0.9%, 0.4%, 0.4% of average prices in the first, second, and third periods. To interpret results from Table 2, the left heading is the range of buying rules, from a 0.05% drop from the most recent peak to 2% while the top heading is the range for selling rules, from a 0.04% profit margin to 2%.

There are interesting implications to be drawn from these tables. First, during 1993-1996, the more profitable strategies are those with high selling rules or those that set high profit margins. With the same profit margins, strategies with smaller retraction, or lower buying rules, tend to do better. Hence, the top performers' regions are those to the top right of the table. The annualized net returns from each combination of buying and selling rules in this first period range from -21.28% to 12.86%.

In the second period (1997-2001), while higher profit margins or selling rules continued to generate higher profitability, larger retraction, or higher buying rules tended to generate better returns. Thus, the top performer's region shifts from the top right to the bottom right of the table. Annualized net returns from each combination of buying rules and selling rules in this second period range from -356.62% to 13.72%. The heavily negative returns region in the top left of the table reflects losses occurring from increases in transaction costs. As the price of SPDR increases, the spread between bid and offer price also increases. As was shown in section 2.1. This increase in transaction costs considerably penalizes strategies with low buying and selling rules.

In the final period (2002-2006), the situation reverses again, and strategies that set lower profit margins turn out to be the most profitable this time. The region of top performers shifts from the bottom-right corner to the medium-left of the table. Annualized net returns from each combination of buying rules and selling rules in this first period ranged from 4.31% to 94.62%. The marked increase in the profitability during this period was primarily driven by the decrease in the bid-ask spread.

Now that the existence of a positive returns region in the sample is established, we shift the focus to comparing these returns from our trading rules *(Bi,Sj)* to the benchmark returns from the buy-and-hold strategy; as shown in Tables 3 and 4, it becomes immediately obvious that all of the strategies have underperformed the buy-and-hold strategy in the first period, 1993-1996. The grey color highlights the 'over-performance' region. And while three strategies have survived the performance test in the second period, 1997-2001, the sizes of out-performance were relatively small, only 0.18%, 1.73%, and 2.17% based on annualized monthly returns.

However, results during 2002-06 are different. The positive region, or the 'out-performance', has surfaced across the ranges of the buying and selling rules. The size of out-performance ranges from 0.03% to 87.49% over the buy-and-hold strategy, based on annualized net returns. A close inspection reveals that the out-performance region includes strategies that set low profit margins, or selling rules, often lower than 0.3%, and have retraction points, or buying rules, between 0.1% and 0.5% from the previous price peak. However, before rushing to the conclusion that there is a true 'out-performance' of the buy-and-hold strategy, and that the weak-form market efficiency can be rejected, it is proposed that these 'abnormal' returns are a product of not filtering the data appropriately for mistakes and outliers. It is asserted that the 50 cent (0.35%) filter of Hasbrouck (2003) is too "wide."

Table 2 Average Annualized Benchmark Returns on the periods studied

Periods	(SPY) Buy&Hold Return incl. Dividends	(SPY) Buy&Hold Return excl. Dividends	Market Value-weighted* Returns incl. Dividends	Market Value-weighted* Returns excl. Dividends	Market Equal-weighted** Returns incl. Dividends	Market Equal-weighted** Returns excl. Dividends	S&P Composite Returns
Jan 93 - Dec 96	15.55%	13.80%	15.46%	13.05%	14.67%	13.08%	13.84%
Jan 97 - Dec 01	11.55%	10.30%	10.91%	9.53%	13.35%	11.87%	10.39%
Jan 02 - Dec 06	6.80%	5.06%	8.72%	6.90%	18.46%	16.70%	5.00%
Jan 02 - Jun 04	2.88%	1.33%	4.85%	3.15%	21.75%	20.04%	1.00%
Jul 04 - Dec 06	10.71%	8.79%	12.58%	10.65%	15.17%	13.36%	8.99%

*Market Value -Weighted Returns contains average monthly returns, including all disstributions, on a value-weighted market portfolio, excluding American Depository receipts(ADRs)

**Market Equal-Weighted Returns contains average monthly returns, on an equally-weighted market portfolio, including American Depository(ADRs)

Table 3 Average Annualized Monthly Returns in the Three Periods with 50 cents Filter

January 1993 - December 1996

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	-21.3%	-15.0%	-6.8%	-3.5%	-1.6%	0.9%	2.8%	3.3%	3.9%	8.3%	9.7%	10.5%	10.9%	11.3%	11.7%	12.0%	12.0%	12.4%
0.10%	-18.0%	-12.2%	-5.7%	-3.0%	-1.0%	1.0%	2.4%	3.5%	4.3%	8.0%	10.4%	10.6%	11.7%	11.1%	12.2%	12.8%	12.1%	12.5%
0.15%	-15.4%	-10.6%	-4.8%	-1.8%	-0.7%	1.1%	2.5%	3.2%	3.8%	8.0%	9.9%	11.4%	11.0%	11.0%	11.5%	12.1%	12.4%	12.8%
0.20%	-12.1%	-7.5%	-3.3%	-0.9%	-1.5%	1.1%	1.9%	3.0%	3.9%	7.9%	9.0%	9.9%	10.7%	10.6%	10.9%	12.4%	12.1%	12.7%
0.25%	-10.1%	-7.9%	-2.9%	-1.2%	-1.2%	0.7%	5.2%	2.1%	3.2%	6.9%	9.7%	9.8%	11.2%	11.0%	10.9%	12.0%	11.7%	12.9%
0.30%	-4.8%	-4.5%	-2.0%	0.0%	1.9%	3.8%	0.9%	4.2%	3.4%	6.5%	9.3%	9.5%	10.7%	11.2%	11.4%	12.0%	12.6%	12.5%
0.50%	-3.4%	-2.0%	0.6%	2.4%	5.0%	5.0%	4.8%	5.8%	5.3%	6.1%	6.6%	8.3%	10.2%	11.3%	11.1%	10.3%	11.8%	10.7%
1.00%	0.6%	1.7%	2.8%	3.4%	4.1%	4.1%	5.0%	4.8%	5.7%	6.1%	6.6%	5.8%	6.5%	6.7%	8.0%	9.0%	8.2%	8.0%
1.50%	1.0%	1.5%	2.1%	2.8%	4.1%	4.1%	4.6%	4.0%	3.5%	4.6%	6.3%	5.8%	7.4%	8.0%	7.3%	7.9%	8.8%	9.0%
2.00%	0.6%	0.9%	1.2%	1.5%	1.8%	1.8%	2.2%	1.9%	2.2%	3.2%	4.5%	3.5%	5.2%	6.5%	6.1%	7.2%	7.6%	7.5%

January 1997 - December 2001

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	-356.6%	-184.1%	-101.9%	-59.6%	-41.9%	-29.9%	-22.9%	-18.3%	-15.5%	-4.7%	-0.3%	1.0%	2.5%	3.2%	4.3%	4.8%	5.1%	6.9%
0.10%	-347.2%	-200.9%	-122.0%	-71.7%	-42.7%	-27.4%	-19.0%	-14.9%	-12.4%	-1.7%	1.3%	2.8%	3.5%	4.7%	5.0%	5.5%	5.7%	7.5%
0.15%	-256.3%	-162.7%	-97.3%	-55.0%	-35.3%	-22.5%	-15.2%	-11.8%	-9.8%	0.9%	2.9%	4.4%	4.7%	5.1%	5.2%	6.7%	5.7%	7.5%
0.20%	-201.1%	-129.7%	-77.8%	-43.8%	-27.5%	-16.0%	-9.6%	-5.6%	-3.5%	3.5%	5.0%	5.6%	5.4%	7.0%	6.1%	6.9%	6.5%	7.2%
0.25%	-149.5%	-98.3%	-63.5%	-36.0%	-22.2%	-13.4%	-6.5%	-2.6%	-0.8%	6.9%	5.7%	6.9%	6.7%	6.8%	7.6%	7.3%	8.0%	7.5%
0.30%	-121.0%	-84.5%	-53.2%	-30.0%	-17.2%	-9.0%	-2.8%	-0.4%	1.5%	9.2%	6.5%	7.5%	7.9%	7.5%	7.2%	8.4%	7.8%	8.0%
0.50%	-49.5%	-29.9%	-19.1%	-12.4%	0.9%	0.9%	4.1%	7.1%	9.2%	13.7%	8.2%	7.5%	7.1%	9.5%	9.7%	9.0%	8.9%	8.9%
1.00%	-10.0%	-5.1%	-2.7%	-0.9%	-0.9%	-0.9%	1.5%	-0.2%	-1.0%	3.2%	5.6%	9.3%	10.1%	10.3%	11.7%	11.4%	9.7%	11.1%
1.50%	-6.4%	-4.2%	-2.5%	0.8%	0.2%	0.2%	1.9%	2.7%	4.4%	10.4%	9.3%	9.1%	5.6%	8.2%	10.4%	13.3%	10.4%	10.0%
2.00%	-1.8%	-0.5%	0.3%	-0.1%	-0.6%	-0.6%	0.4%	1.8%	2.0%	5.0%	5.6%	5.1%	2.9%	4.9%	5.7%	8.3%	6.2%	8.8%

January 2002 - December 2006

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	23.2%	16.0%	11.8%	10.1%	7.4%	7.2%	6.0%	5.2%	4.8%	6.8%	6.2%	5.6%	6.0%	5.1%	4.7%	4.4%	4.6%	4.3%
0.10%	80.8%	59.1%	43.1%	28.4%	21.2%	16.2%	14.4%	12.5%	12.1%	11.3%	9.8%	8.4%	7.0%	6.7%	6.2%	5.0%	5.4%	4.8%
0.15%	94.3%	73.4%	57.3%	47.9%	38.0%	29.7%	21.7%	18.7%	16.7%	17.3%	13.2%	9.3%	8.4%	8.5%	6.5%	6.0%	6.9%	5.5%
0.20%	94.6%	83.7%	68.6%	58.7%	46.8%	41.2%	33.4%	27.9%	22.5%	22.2%	15.4%	11.4%	10.2%	9.5%	7.2%	6.8%	6.8%	5.5%
0.25%	91.0%	72.3%	65.2%	55.2%	49.5%	45.1%	38.4%	33.0%	32.0%	28.2%	19.9%	13.9%	12.2%	10.3%	9.0%	7.8%	7.1%	5.6%
0.30%	77.9%	71.9%	63.9%	53.2%	48.5%	44.6%	39.5%	36.2%	33.2%	34.7%	22.8%	17.0%	15.5%	11.9%	10.1%	9.1%	7.8%	5.9%
0.50%	88.2%	83.4%	67.7%	62.7%	56.9%	56.2%	51.7%	50.5%	49.3%	41.4%	34.2%	27.8%	22.5%	19.7%	16.7%	12.5%	12.3%	6.7%
1.00%	21.9%	23.9%	17.6%	17.3%	14.5%	14.4%	12.0%	13.4%	13.3%	15.3%	14.9%	13.8%	14.7%	13.9%	14.0%	13.6%	13.3%	7.2%
1.50%	9.0%	11.1%	13.4%	15.4%	14.8%	13.0%	12.8%	9.5%	11.1%	7.9%	6.6%	6.3%	4.7%	4.2%	6.3%	6.1%	6.8%	4.9%
2.00%	5.4%	6.9%	8.3%	8.7%	9.7%	8.7%	10.3%	10.2%	9.8%	5.2%	7.5%	7.3%	7.9%	5.4%	4.5%	4.7%	4.6%	3.8%

Note : Left heading represents Buying Signal. And Top heading represents Selling Signal. For example, in the final period, R(B,S) = R(0.25 %, 0.30 %) = 28.24 %. Darker region represents more profitable strategies.

[22] For this and following tables, darker shade of gray indicates higher profitable strategies.

Table 4 Average 'Out-performance' of R(B,S) to R(B&H) in Annualized Returns

January 1993 - December 1996

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	-36.8%	-30.6%	-22.4%	-19.0%	-17.1%	-14.7%	-12.8%	-12.2%	-11.7%	-7.3%	-5.9%	-5.0%	-4.6%	-4.2%	-3.8%	-3.5%	-3.6%	-3.2%
0.10%	-33.5%	-27.8%	-21.2%	-18.6%	-16.5%	-14.5%	-13.1%	-12.0%	-11.3%	-7.5%	-5.2%	-4.9%	-3.9%	-4.4%	-3.3%	-2.8%	-3.5%	-3.0%
0.15%	-30.9%	-26.1%	-18.9%	-17.4%	-16.3%	-14.4%	-13.1%	-12.3%	-11.8%	-7.6%	-5.6%	-4.2%	-4.5%	-4.5%	-4.0%	-3.5%	-3.1%	-2.7%
0.20%	-27.6%	-23.1%	-18.5%	-16.5%	-17.1%	-14.5%	-13.6%	-12.6%	-11.7%	-7.7%	-6.5%	-5.6%	-4.9%	-4.9%	-4.6%	-3.2%	-3.4%	-2.9%
0.25%	-25.6%	-23.5%	-17.6%	-16.7%	-16.8%	-14.9%	-10.3%	-13.4%	-12.4%	-8.7%	-5.8%	-5.8%	-4.3%	-4.5%	-4.7%	-3.6%	-3.9%	-2.7%
0.30%	-20.3%	-20.0%	-14.9%	-15.5%	-13.7%	-11.8%	-14.7%	-11.4%	-12.2%	-9.0%	-6.3%	-6.0%	-4.9%	-4.3%	-4.2%	-3.6%	-2.9%	-3.1%
0.50%	-18.9%	-17.6%	-12.8%	-13.2%	-10.6%	-10.6%	-10.7%	-9.7%	-10.2%	-9.4%	-8.9%	-7.2%	-5.3%	-4.3%	-4.5%	-5.2%	-3.8%	-4.8%
1.00%	15.0%	-13.8%	-13.5%	-12.2%	-11.5%	-11.5%	-10.6%	-10.8%	-9.9%	-9.5%	-8.9%	-9.7%	-9.1%	-8.9%	-7.5%	-6.6%	-7.3%	-7.6%
1.50%	-14.5%	-14.1%	-14.4%	-12.7%	-11.5%	-11.5%	-11.0%	-11.6%	-12.0%	-11.0%	-9.3%	-9.8%	-8.2%	-7.6%	-8.2%	-7.7%	-6.7%	-6.6%
2.00%	-14.9%	-14.6%		-14.0%	-13.8%	-13.8%	-13.6%	-13.7%	-13.4%	-12.4%	-11.1%	-12.1%	-10.4%	-9.1%	-9.5%	-8.4%	-7.9%	-8.1%

January 1997 - December 2001

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	-368.2%	-195.7%	-113.5%	-71.1%	-53.4%	-41.4%	-34.5%	-29.9%	-27.0%	-16.3%	-11.8%	-10.5%	-9.1%	-8.3%	-7.2%	-6.8%	-6.5%	-4.6%
0.10%	-358.7%	-212.5%	-133.6%	-83.3%	-54.2%	-39.0%	-30.6%	-26.4%	-24.0%	-13.3%	-110.3%	-8.8%	-8.1%	-6.9%	-6.5%	-6.1%	-5.8%	-4.0%
0.15%	-267.8%	-174.3%	-108.9%	-66.5%	-46.9%	-34.1%	-26.8%	-23.3%	-21.4%	-10.7%	-8.6%	-7.1%	-6.8%	-6.4%	6.4%	-4.9%	-5.9%	-4.1%
0.20%	-212.6%	-141.2%	-89.4%	-55.3%	-39.1%	-27.6%	-21.2%	-17.1%	-15.0%	-8.0%	-6.5%	-6.0%	-6.1%	-4.6%	-5.4%	-4.7%	-5.1%	-4.3%
0.25%	-161.1%	-109.9%	-75.1%	-47.5%	-33.7%	-24.9%	-18.1%	-14.2%	-12.4%	-4.7%	-5.9%	-4.6%	-4.9%	-4.8%	-4.0%	-4.2%	-3.6%	-4.0%
0.30%	-132.6%	-96.0%	-64.7%	-41.5%	-28.8%	-20.6%	14.4%	-12.0%	-10.1%	-2.4%	-5.1%	-4.1%	-3.6%	-4.1%	-4.3%	-3.2%	-3.7%	-3.5%
0.50%	-61.0%	-41.5%	-30.7%	-24.0%	-10.7%	-10.7%	-7.5%	-4.4%	-2.3%	2.2%	-3.3%	-4.1%	-4.5%	-2.1%	-1.8%	-2.6%	-2.7%	-2.7%
1.00%	-21.5%	-16.7%	-14.3%	-12.5%	-12.5%	-12.5%	-10.1%	-11.7%	-12.6%	-8.4%	-5.9%	-2.3%	-1.4%	-1.2%	0.2%	-0.2%	-1.9%	-0.5%
1.50%	-17.9%	-15.7%	-14.1%	-10.8%	-11.3%	-11.3%	-9.7%	-8.8%	-7.2%	-1.2%	-2.3%	-2.5%	-5.9%	-3.3%	-1.2%	1.7%	-1.2%	-1.6%
2.00%	-13.4%	-12.1%	-11.3%	-11.7%	-12.2%	-12.2%	-11.2%	-9.8%	-9.6%	-6.5%	-6.0%	-6.5%	-8.7%	-6.6%	-5.9%	-3.3%	-5.4%	-2.7%

January 2002 - December 2006

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	16.4%	9.2%	5.0%	3.3%	0.6%	0.4%	-0.8%	-1.6%	-2.0%	0.0%	-0.6%	-1.2%	-0.8%	-1.7%	-2.1%	-2.4%	-2.2%	-2.5%
0.10%	74.0%	52.3%	36.3%	21.6%	14.4%	9.4%	7.6%	5.7%	5.3%	4.5%	3.0%	1.6%	0.2%	-0.1%	-0.6%	-1.8%	-1.4%	2.0%
0.15%	87.5%	66.6%	50.5%	41.1%	31.2%	22.9%	14.9%	11.9%	9.9%	10.5%	6.4%	2.5%	1.6%	1.7%	-0.3%	-0.8%	0.1%	-1.3%
0.20%	87.8%	76.9%	61.8%	51.9%	40.0%	34.4%	26.6%	21.1%	15.7%	15.4%	8.6%	4.6%	3.4%	2.7%	0.4%	0.0%	0.0%	-1.3%
0.25%	84.2%	65.5%	58.4%	48.4%	42.7%	38.3%	31.6%	26.2%	25.2%	21.4%	13.1%	7.1%	5.4%	3.5%	2.3%	1.0%	0.3%	-1.2%
0.30%	71.1%	65.1%	57.1%	46.4%	41.7%	37.8%	32.7%	29.4%	26.4%	27.9%	16.0%	10.2%	8.7%	5.1%	3.3%	2.3%	1.0%	-0.9%
0.50%	81.4%	76.6%	60.9%	55.9%	50.1%	49.4%	44.9%	43.7%	42.5%	34.6%	27.4%	21.0%	15.7%	12.9%	9.9%	5.7%	5.5%	-0.1%
1.00%	15.1%	17.1%	10.8%	10.5%	7.7%	7.6%	5.2%	6.6%	6.5%	8.5%	8.1%	7.0%	7.9%	7.1%	7.2%	6.8%	6.5%	0.4%
1.50%	2.2%	4.3%	6.6%	8.6%	8.0%	6.2%	6.0%	2.7%	4.3%	1.1%	-0.2%	-0.6%	-2.1%	-2.6%	-0.5%	-0.7%	0.0%	-1.9%
2.00%	-1.4%	0.1%	1.5%	1.9%	3.0%	1.9%	3.5%	3.4%	3.0%	-1.6%	0.8%	0.5%	1.1%	-1.5%	-2.3%	-2.1%	-2.2%	-3.0%

Note : Left heading represents Buying Signal. And Top heading represents Selling Signal. For example, in the final period, R(B,S) = R(0.25 %, 0.30 %) = outperforms R(B&H) by 21.44%. Darker region represents more profitable strategies.

4.1.1 Robustness Check

Given the number of successful strategies in the final period, it is appropriate to continue to explore the robustness of the profitable strategies. This section will break down the final period into an additional two sub-periods.

Table 5 compares returns from both sub-periods. All of the strategies in the set have generated positive returns and are robust across the two periods. The next table presents excess returns, that is returns from strategy *(B,S)* minus returns from strategy *(B&H)*. This gives the out-performance measure. A broad observation suggests that the excess returns, or out-performance of strategy *(B,S)*, tends to persist over the second period as well, although the out-performance region has somewhat contracted from the first sub-period, and the size of out-performance has reduced across the board.

From the region of positive returns, one (typical) single strategy can be drawn out and analyzed in detail. Figure 5 depicts the out-performance of strategy (0.25%, 0.12%) from January 2002 to December 2006. The break point for the two sub-periods is July 1, 2004. In the first sub-period, there were substantial excess-returns, many of which were larger than five percent on a monthly basis. Another interesting observation is that these higher than usual excess returns concentrated on only market upturns. In the second sub-period, although magnitudes were generally smaller, these excess returns continue to persist throughout the period. There is no evidence of substantial excess returns, but consistent two to five percent excess returns are generally observed in this second period.

Table 5 Average Annualized Returns from strategy (B,S) in two sub-periods

January 2002 - June 2004

Period 3.1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	27.6%	18.7%	10.7%	8.9%	5.0%	4.7%	3.1%	2.1%	1.3%	3.2%	2.8%	1.6%	2.5%	0.7%	0.2%	0.3%	0.1%	-0.2%
0.10%	107.5%	81.3%	59.1%	36.9%	24.9%	18.4%	15.3%	13.3%	12.6%	9.6%	6.1%	4.8%	3.7%	2.2%	1.9%	0.6%	2.2%	0.5%
0.15%	132.9%	99.1%	78.3%	63.9%	50.8%	39.9%	27.1%	22.5%	20.0%	15.3%	9.6%	5.0%	4.8%	4.5%	2.2%	1.9%	2.9%	1.2%
0.20%	137.4%	117.8%	93.4%	75.8%	61.2%	52.2%	41.0%	33.2%	26.6%	20.4%	11.3%	7.0%	5.7%	4.9%	3.1%	1.9%	1.5%	1.3%
0.25%	133.5%	99.0%	88.4%	73.3%	62.6%	54.4%	46.3%	38.6%	37.2%	28.6%	17.9%	11.0%	8.0%	6.1%	6.1%	3.0%	2.0%	1.5%
0.30%	117.2%	106.5%	92.0%	70.7%	60.7%	54.4%	49.1%	44.3%	39.3%	36.0%	22.7%	14.5%	12.0%	9.1%	6.1%	5.4%	3.0%	2.3%
0.50%	124.6%	110.0%	87.0%	79.2%	72.7%	72.7%	64.9%	65.1%	64.3%	48.7%	39.0%	30.8%	25.2%	21.9%	16.0%	9.5%	8.8%	2.6%
1.00%	32.3%	35.6%	25.8%	25.1%	17.1%	17.1%	12.4%	14.3%	14.3%	19.0%	19.0%	17.7%	18.0%	18.2%	17.0%	16.4%	15.6%	2.8%
1.50%	14.7%	18.6%	21.8%	24.6%	17.9%	17.9%	16.9%	9.7%	11.7%	7.3%	4.1%	6.6%	2.9%	4.1%	7.3%	8.1%	8.6%	3.3%
2.00%	9.6%	11.9%	14.2%	14.4%	14.4%	14.4%	16.6%	16.8%	15.6%	6.0%	10.2%	9.7%	10.2%	4.4%	3.9%	6.0%	5.5%	2.5%

July 2004 - December 2006

Period 3.2	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	18.8%	13.4%	12.8%	11.2%	9.8%	9.7%	8.8%	8.3%	8.4%	10.4%	9.7%	9.5%	9.5%	9.6%	9.2%	8.6%	9.0%	8.9%
0.10%	54.2%	36.8%	27.2%	19.9%	17.5%	13.9%	13.5%	11.7%	11.6%	12.9%	13.5%	12.0%	10.2%	11.2%	10.5%	9.3%	9.6%	9.1%
0.15%	55.7%	47.6%	36.4%	31.9%	25.2%	19.4%	16.3%	14.9%	13.4%	19.4%	16.9%	13.5%	12.0%	12.5%	10.7%	10.1%	10.8%	9.7%
0.20%	51.8%	49.7%	43.8%	41.5%	32.3%	30.3%	25.8%	22.6%	18.4%	24.1%	19.4%	15.8%	14.7%	14.2%	11.4%	11.7%	12.1%	9.8%
0.25%	48.5%	45.7%	41.9%	37.2%	36.3%	35.8%	30.6%	27.4%	26.9%	27.9%	21.9%	16.8%	16.4%	14.5%	12.1%	12.5%	12.2%	9.6%
0.30%	38.7%	37.4%	35.7%	35.8%	36.2%	34.7%	29.9%	28.2%	27.0%	33.4%	22.9%	19.5%	19.0%	14.7%	14.1%	12.8%	12.6%	9.5%
0.50%	51.8%	56.9%	48.4%	46.2%	39.9%	39.7%	38.6%	35.8%	34.2%	34.2%	29.4%	24.9%	19.9%	17.4%	17.4%	15.6%	15.8%	10.8%
1.00%	11.4%	12.2%	9.4%	9.5%	10.2%	11.8%	11.6%	12.4%	12.4%	11.6%	10.8%	9.9%	11.5%	9.7%	11.0%	10.9%	10.9%	11.6%
1.50%	3.3%	3.6%	5.1%	6.1%	7.4%	8.2%	8.7%	9.4%	10.6%	8.5%	9.2%	5.9%	6.5%	4.3%	5.3%	4.1%	4.9%	6.5%
2.00%	1.2%	1.9%	2.3%	3.1%	2.6%	3.1%	3.9%	3.7%	3.9%	4.5%	4.9%	4.9%	5.7%	6.3%	5.1%	3.5%	3.8%	5.1%

Note: Left heading represents Buying Signal. And Top heading represents Selling Signal. For example, in the final period, R(B,S) = R(0.25%,0.30%) = 27.90%. Darker region represents more profitable strategies.

Table 6 Average Annualized Excess Returns of Strategy (B,S) to strategy (B&H) during two sub-periods

January 2002 - June 2004

Period 3.1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	24.7%	15.8%	7.8%	6.1%	2.1%	1.8%	0.2%	-0.8%	-1.6%	0.3%	-0.1%	-1.3%	-0.4%	-2.2%	-2.7%	-2.6%	-2.8%	-3.1%
0.10%	104.6%	78.4%	56.2%	34.0%	22.0%	15.5%	12.4%	10.4%	9.7%	6.7%	3.3%	2.0%	0.9%	-0.7%	-1.0%	-2.3%	-1.6%	-2.4%
0.15%	130.0%	96.2%	75.4%	61.0%	47.9%	37.0%	24.2%	19.6%	17.1%	12.4%	6.8%	2.1%	1.9%	1.6%	-0.7%	-1.0%	0.0%	-1.7%
0.20%	134.6%	114.9%	90.5%	73.0%	58.4%	49.3%	38.1%	30.3%	23.7%	17.5%	8.4%	4.1%	2.9%	2.0%	0.2%	-1.0%	-1.4%	-1.6%
0.25%	130.6%	96.1%	85.5%	70.4%	59.7%	51.5%	43.4%	35.7%	34.3%	25.7%	12.0%	8.1%	5.1%	3.2%	3.2%	0.1%	-0.9%	-1.4%
0.30%	114.3%	103.6%	89.1%	67.8%	57.8%	51.5%	46.2%	41.4%	36.5%	33.1%	19.8%	11.6%	9.1%	6.2%	3.2%	2.6%	0.1%	-0.6%
0.50%	121.8%	107.2%	84.2%	76.3%	69.8%	69.8%	62.0%	62.3%	61.4%	45.8%	36.1%	27.9%	22.3%	19.0%	13.1%	6.6%	5.9%	-0.3%
1.00%	29.4%	32.7%	22.9%	22.2%	14.2%	14.2%	9.5%	11.4%	11.4%	16.1%	16.1%	14.8%	15.1%	15.3%	14.1%	13.5%	12.8%	-0.1%
1.50%	11.8%	15.7%	18.9%	21.7%	15.0%	15.0%	14.0%	6.8%	8.8%	4.4%	1.2%	3.7%	0.0%	1.2%	4.4%	5.2%	5.7%	0.4%
2.00%	6.7%	9.0%	11.3%	11.5%	11.5%	11.5%	13.7%	13.9%	12.7%	3.1%	7.3%	6.8%	7.3%	1.5%	1.0%	3.1%	2.6%	-0.4%

July 2004 - December 2006

Period 1	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	8.1%	2.7%	2.1%	0.5%	-1.0%	-1.0%	-1.9%	-2.4%	-2.3%	-0.3%	-1.0%	-1.2%	-1.2%	-1.2%	-1.5%	2.1%	-1.7%	-1.9%
0.10%	43.5%	26.1%	16.5%	28.4%	6.8%	3.2%	2.8%	1.0%	0.9%	2.2%	2.8%	1.3%	-0.5%	0.5%	-0.2%	-1.7%	-1.1%	-1.6%
0.15%	45.0%	36.9%	25.6%	47.9%	14.5%	8.7%	5.5%	4.1%	2.7%	8.6%	6.1%	2.8%	1.3%	1.8%	0.0%	-0.6%	0.1%	-1.0%
0.20%	41.1%	38.9%	33.1%	58.7%	21.6%	19.6%	15.0%	11.9%	7.6%	13.3%	8.7%	5.1%	4.0%	3.4%	0.6%	1.0%	1.4%	-0.9%
0.25%	37.8%	35.0%	31.2%	55.2%	25.6%	25.1%	19.9%	16.7%	16.2%	17.2%	11.2%	6.0%	5.7%	3.8%	1.4%	1.8%	1.5%	-1.1%
0.30%	27.9%	26.7%	25.0%	53.2%	25.5%	24.0%	19.2%	17.4%	16.3%	22.7%	12.2%	8.8%	8.3%	4.0%	3.4%	2.1%	1.9%	-1.2%
0.50%	41.1%	46.1%	37.7%	35.5%	29.1%	29.0%	27.9%	25.1%	23.5%	23.5%	18.7%	14.1%	9.2%	6.7%	6.7%	4.8%	5.1%	0.1%
1.00%	0.7%	1.5%	-1.3%	-1.2%	-0.5%	1.1%	0.9%	1.7%	1.6%	0.9%	0.1%	-0.8%	0.8%	-1.1%	0.3%	0.2%	0.2%	0.9%
1.50%	-7.4%	-7.1%	-5.6%	-4.6%	-3.4%	-2.6%	-2.1%	-1.3%	-0.2%	-2.2%	-1.5%	-4.8%	-4.2%	-6.4%	-5.4%	-6.6%	-5.8%	-4.2%
2.00%	-9.5%	-8.8%	-8.4%	-7.7%	-8.1%	-6.8%	-6.8%	-7.0%	-6.8%	-6.2%	-5.8%	-5.8%	-5.0%	-4.4%	-5.6%	-7.2%	-6.9%	-5.6%

Note: Left heading represents Buying Signal. And Top heading represents Selling Signal. For example, in the final period, R(B,S) = R(0.25%,0.30%) outperforms R(B&H) by 17.19%

**Darker region represents more profitable strategies.*

Table 7 Ex Post Sharpe Ratios of Benchmark Strategies during January 2002 and December 2006

	Buy&Hold Returns incl. Dividends	Buy&Hold Returns excl. Dividends	Market Value-Weighted* Returns incl. Dividends	Market Value-Weighted* Returns excl. Dividends	Market Equal-Weighted** Returns incl. Dividends	Market Equal-Weighted** Returns excl. Dividends	S&P Composite Returns
Jan 02 - Dec 06	0.10	0.06	0.15	0.10	0.29	0.26	0.06

*Market Value-Weighted Returns contains average monthly returns, including all distributions, on a value-weighted market portfolio, excluding American Depository receipts (ADRs).

**Market Equal-Weighted Return contains average monthly returns, on an equally-weighted market portfolio, including American Depository (ADRs).

Table 8 Ex Post Sharpe Ratios of Strategy (Bi,Sj) during January 2002 and December 2006

	0.04%	0.06%	0.08%	0.10%	0.12%	0.14%	0.16%	0.18%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%	2.00%
0.05%	35%	26%	19%	17%	11%	11%	8%	6%	6%	10%	9%	7%	8%	6%	5%	5%	5%	4%
0.10%	64%	58%	50%	40%	33%	27%	24%	21%	20%	18%	16%	13%	10%	10%	9%	6%	7%	6%
0.15%	69%	66%	57%	55%	48%	42%	33%	30%	27%	28%	22%	15%	13%	13%	9%	8%	10%	7%
0.20%	70%	65%	60%	59%	55%	53%	49%	43%	36%	36%	25%	19%	16%	16%	11%	10%	10%	7%
0.25%	67%	73%	65%	59%	57%	54%	50%	46%	45%	40%	31%	22%	19%	16%	14%	12%	10%	7%
0.30%	67%	72%	66%	57%	57%	54%	50%	47%	46%	48%	36%	27%	26%	19%	16%	15%	12%	8%
0.50%	98%	94%	76%	69%	63%	61%	59%	57%	55%	49%	43%	36%	31%	28%	24%	20%	20%	10%
1.00%	52%	49%	32%	34%	27%	26%	20%	22%	22%	23%	22%	20%	21%	20%	21%	20%	20%	10%
1.50%	33%	40%	46%	49%	36%	34%	32%	23%	27%	16%	11%	10%	5%	4%	9%	8%	10%	6%
2.00%	19%	26%	32%	34%	36%	23%	28%	27%	24%	8%	13%	13%	14%	7%	5%	6%	5%	3%

Table 7 shows benchmark ex post Sharpe ratios during January 2002 and December 2006. Table 8 shows ex post Sharpe ratios from strategy *(Bi,Sj)* during the same period. The ex post Sharpe ratio that is based on the buy-and-hold returns during January 2004 and December 2006 is 0.10, implying 10% excess return to risk-free asset per one unit of risk. The gray region in Table 8 represents Sharpe's ratios from strategy *(Bi,Sj),* which outperform this benchmark criterion.

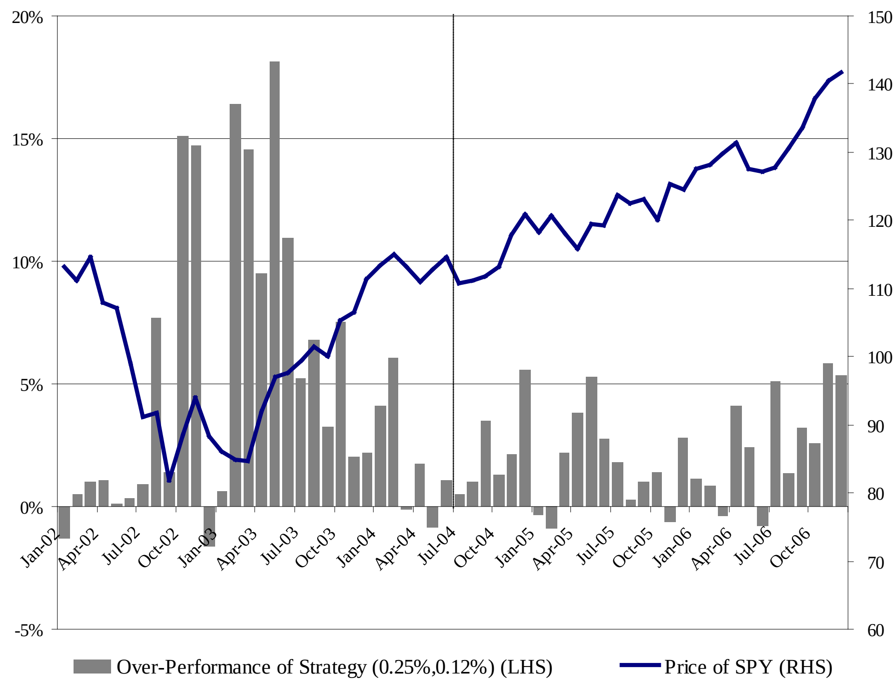


Figure 18 Monthly Over-Performance of Strategy (0.25%,0.12%) and SPDR Price during January 2002 and December 2006

The interpretation of this gray region represents strategies with better reward to risk ratios compared to the buy-and-hold strategy. In retrospect, this graphic representation is consistent with Table 4, suggesting that the out-performance of strategy *(Bi,Sj)* holds true even when risk is taken into account.

4.1.2 Analysis of Selected Profitable Strategy from the 50 Cent Filter

The analysis of the following section relies on monthly returns from strategy *(B,S) = (0.25%,0.12%)* during 2002-2006.

Table 9 Descriptive Statistics of the Monthly Returns of the Strategy and the Benchmark

	R(B,S)	*R(B&H)*
Mean	0.0412	0.0057
Standard Error	0.0089	0.0046
Median	0.0302	0.0104
Standard Deviation	0.0691	0.0355
Sample Variance	0.0048	0.0013
Kurtosis	1.2306	1.3860
Skewness	0.8853	(0.5923)
Range	0.3272	0.1893
Minimum	(0.0907)	(0.1047)
Maximum	0.2365	0.0846
Count	60	60

Source: compiled from Stata's output

Sixty month returns are compiled to create the distribution histogram in Figure 6. Broad observation suggests that the distinct feature of the *R(B,S)* lies in its right tail, which spans very far right, covering a more positive region of the returns than *R(B&H).* This unique feature of *R(B,S)* allows its computed mean to be significantly larger than zero, and much higher than *R(B&H).* Moreover, the left tail of the *R(B,S)* does not extend further than that of *R(B&H).*

Figure 19 shows a large majority of dots on the left side of the 45 degree line, which indicates out-performance of the buy-and-hold strategy. Figure 20-21 shows a large majority of dots above the x-axis, which indicates out-performance of the buy-and-hold strategy. The next section addresses analysis of this out-performance using the Capital Asset Pricing Model approach.

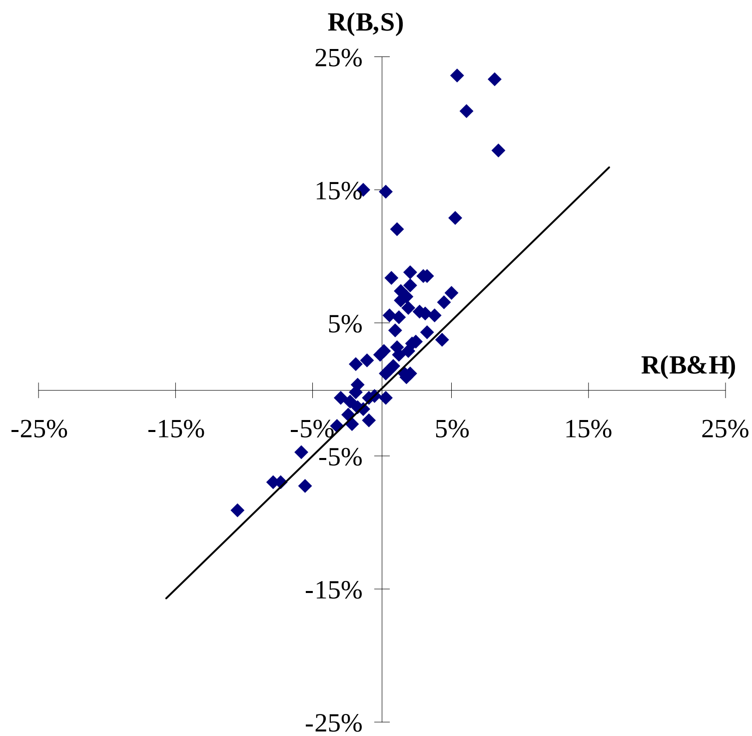
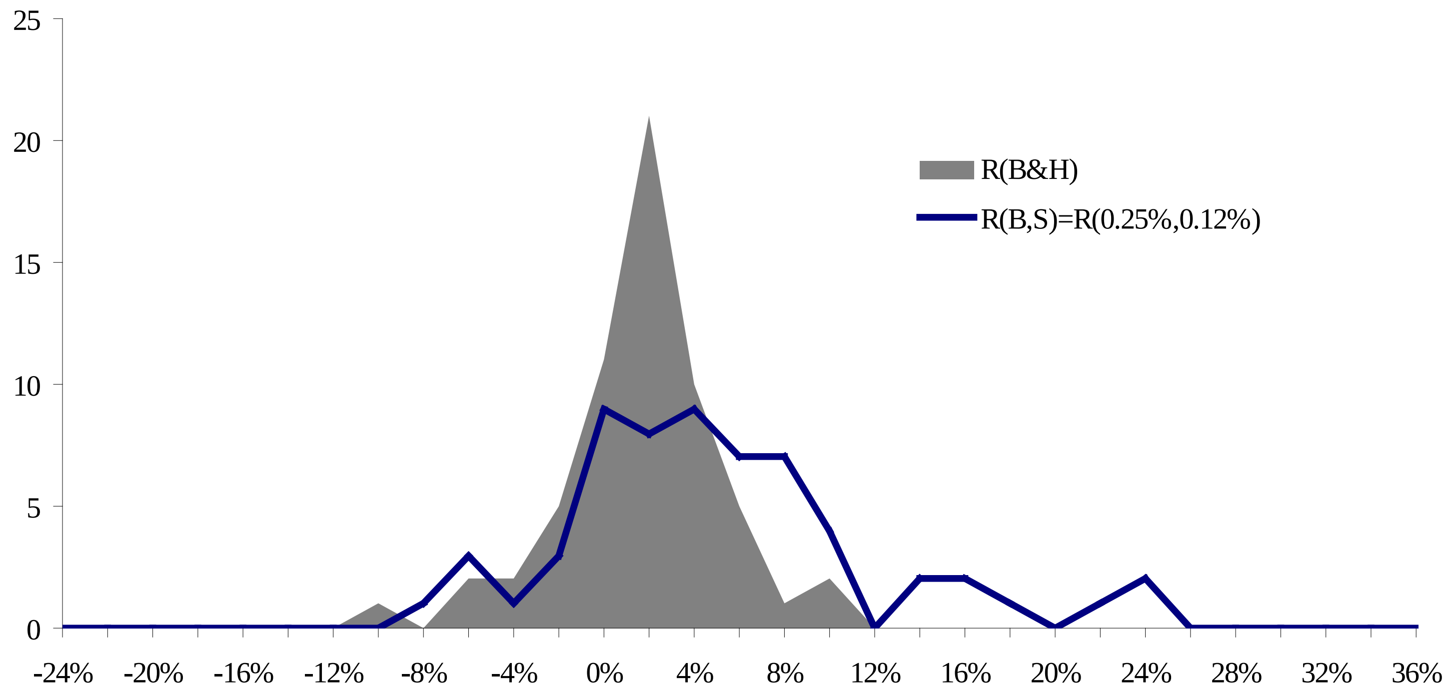


Figure 19 Scatter Plot of R(B,S)=R(0.25%,0.12%) and R(B&H) with 45 degree line

Figure 20 Distributions of the annualized returns between the strategy and the benchmark



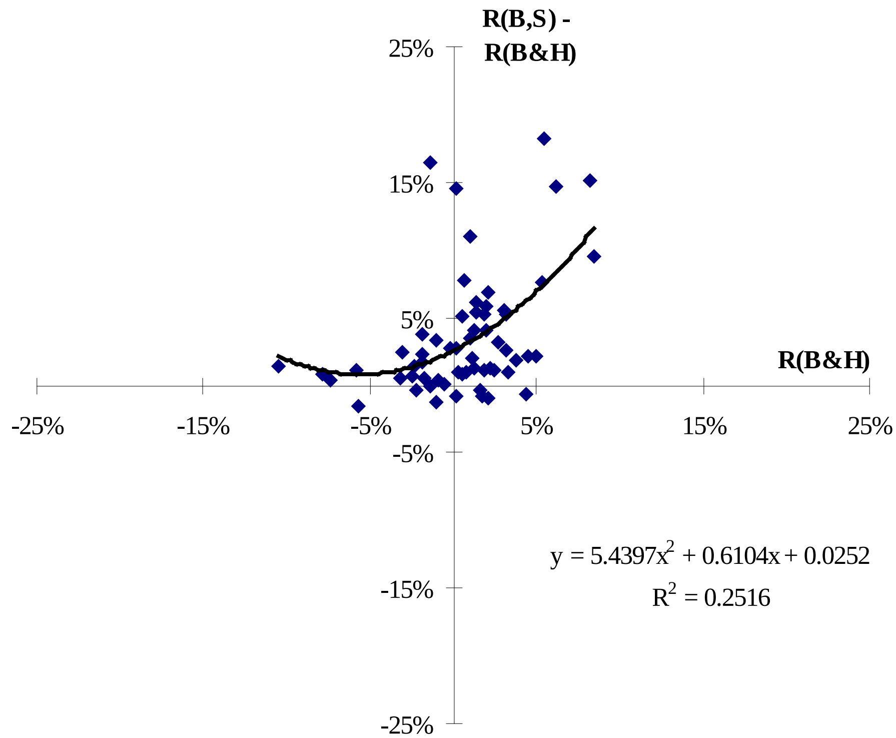


Figure 21 Excess Returns of (B,S)=(0.25%,0.12%) to R(B&H) with quadratic fitted curve

4.1.3 Risk Analysis Using CAPM

To measure the degree to which systematic risk has an impact on the *(B,S)* strategy, this analysis employs tools from the popular Capital Asset Pricing Model (CAPM).

First consider the following notations,

R_F = Return from risk-free asset

$\beta_{(Bi,Sj)}$ = Systematic risk of strategy (B_i,S_j)

$R(B_i,S_j)$ = Returns of strategy (B_i,S_j)

R(B&H) = Market returns, here similar to returns of strategy *(B&H)* of SPDR

CAPM values the expected returns of strategy (B_i,S_j) as a combination of a return from risk-free asset and a product of its systematic risk and market risk premium. Formally,

$$E\left[R\left(B_i,S_j\right)\right] = R_F + \left\{\beta_{(B_i,S_j)}[R(B\&H) - R_F]\right\} \quad (1)$$

where $\beta_{(Bi,Sj)}$ is the beta coefficient from a time series regression of $R(B_i,S_j)$ on market risk premium. From (1), it can be interpreted that the required return on strategy (B_i,S_j) is a linear function of excess market returns.

Re-arrange the terms in (1) with an intercept term.

$$E\left[R\left(B_i,S_j\right)\right] - R_F = \alpha + \left\{\beta_{(B_i,S_j)}[R(B\&H) - R_F]\right\} \quad (2)$$

If CAPM holds, then a regression of $(R(B_i,S_j) - R_F)$ on $(R(B\&H) - R_F)$ will yield an intercept (α) of zero. Any alpha larger than zero implies that the strategy is returning more than is required from it, given its level of systematic risk. In other words, alpha is the additional return unaccounted for by the CAPM.

Since *R(B_i,S_j)* and *R(B&H)* are known from the last section, and we obtain risk free rates from the *3-Month U.S. Treasury bill*, equation (2) can be regressed to determine if there exists such an α term, Here, the analysis continues to use strategy *(B,S)=(0.25%,0.12%)* during 2002-2006.

Table 10 OLS regression analysis of CAPM

Dependent Variable:	R(0.25%,0.12%) - Rf
Constant (Alpha)	0.03342**
	(0.00538)
Returns on B&H - Rf	1.57392**
	(0.15212)
Adj-R2	0.64
Observations	60

Note: ** denotes 1% significance level
Standard errors are given in parentheses.

The estimated systematic risk coefficient (beta) is 1.57, and significantly different from one. This shows that returns from the *(B,S)=(0.25%,0.12%)* strategy are more variable than market returns. The estimated alpha is positive and highly significant, as well. This indicates that the strategy yields returns superior to its systematic risk prediction. The size of the alpha also implies additional 3.3% more monthly returns from the strategy than otherwise anticipated. In annualized terms, this alpha represents over 40% additional returns. A graphic representation of equation (1) is presented in figure 9, below.

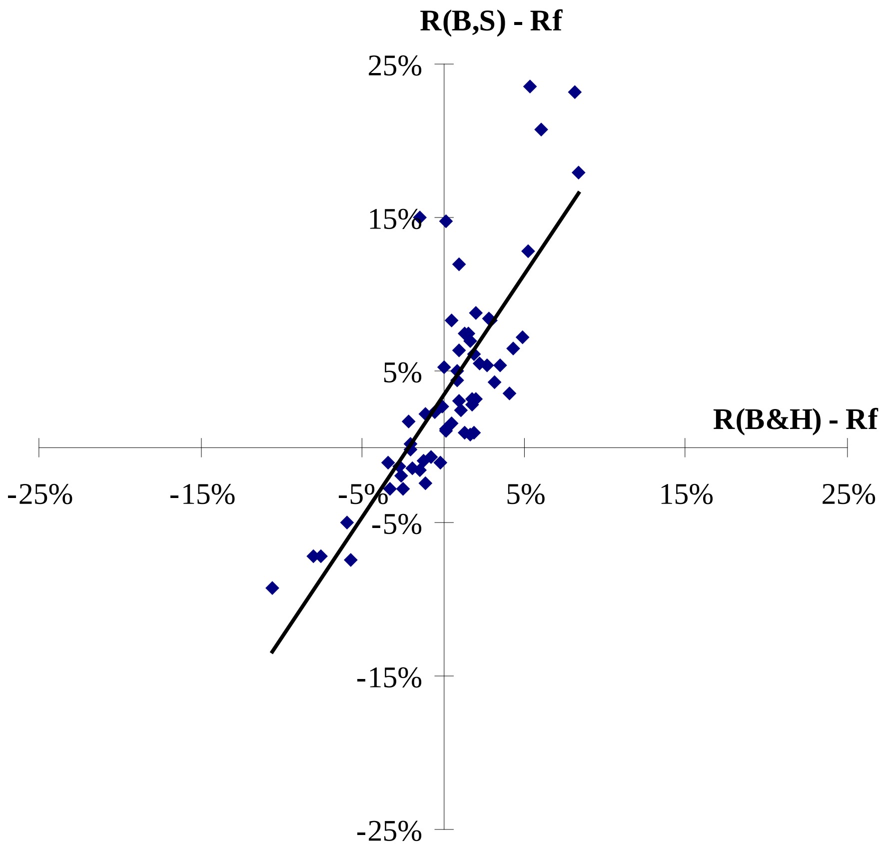


Figure 22 Scatter Plot of Excess R(B,S)=R(0.25%,0.12%) to Rf and Excess R(B&H) to Rf with Plotted Linear Regression line

4.1.4 Downside versus Upside Beta

The discovery of a beta larger than one should not come as a surprise when we recall Figure 6 (distribution of returns from the strategy versus the market). The right tail of the distribution of the strategy's returns was much fatter than that of the benchmark buy-and-hold returns. On the other hand, the left tail of the strategy was not fatter. This suggests that the strategy's returns are higher than those of the buy-and-hold when the market is up but not necessarily lower when the market is down. Therefore, we suspect that beta (systematic risk) is not the same across positive and negative returns. This section addresses asymmetric systematic risk in $R(B_i,S_j)$ by allowing the beta to be different during positive and negative market returns. This is implemented by the introduction of a dummy variable for negative market returns. The specification is the following form

$$R(B_i, S_j) - R_F = \alpha + \left\{\beta^d_{(B_i,S_j)} D[R(B\&H) - R_F]\right\}$$

$$+ \left\{\beta^u_{(B_i,S_j)} (1 - D)[R(B\&H) - R_F]\right\} + \varepsilon \qquad (3)$$

where R_F represents return from risk-free asset, $R(B_i,S_j)$ denotes returns of strategy (B_i,S_j). $R(B\&H)$ is market returns, here similar to returns of strategy $(B\&H)$ of SPDR. D is a dummy variable for the 'downside', 1 if $R(B\&H)\text{-}Rf < 0$. $\beta^d_{(Bi,Sj)}$ represents 'downside' systematic risk of strategy(B_i,S_j) and $\beta^u_{(Bi,Sj)}$ represents ' upside' systematic risk of strategy(B_i,S_j). α denotes additional return unaccounted for in CAPM.

Table 11 OLS Regression Analysis of CAPM

Dependent Variable:	R(0.25%,0.12%) - Rf
Constant (Alpha)	0.02286**
	(0.00799)
[Returns on B&H - Rf] if D=1	1.20311**
	(0.25792)
[Returns on B&H - Rf] if D=0	1.99576**
	(0.28200)
Adj-R2	0.66
Observations	60

Note: ** denotes 1% significance level
Standard errors are given in parentheses

Table 11 reports results from estimating (3) , shown as the systematic risk of *R(B,S),* when *R(B&H)-Rf* is negative, is 1.2, whereas the systematic risk of *R(B,S)* when *R(B&H)-Rf* is positive is much higher, at 2.0. The downside $\beta^d(B,S)$ is statistically significantly larger than the upside $\beta^u(B,S)$. The structure of this asymmetric beta has favored the *R(B,S)* to achieve much higher returns when the market is in an uptrend and also limit the volatility of negative returns when the market is in a downtrend. A graphic representation of equation II is shown in Figure 23. Figure 24 shows a more flexible functional form between *R(B,S)* and the return on the index. This further confirms the above result.

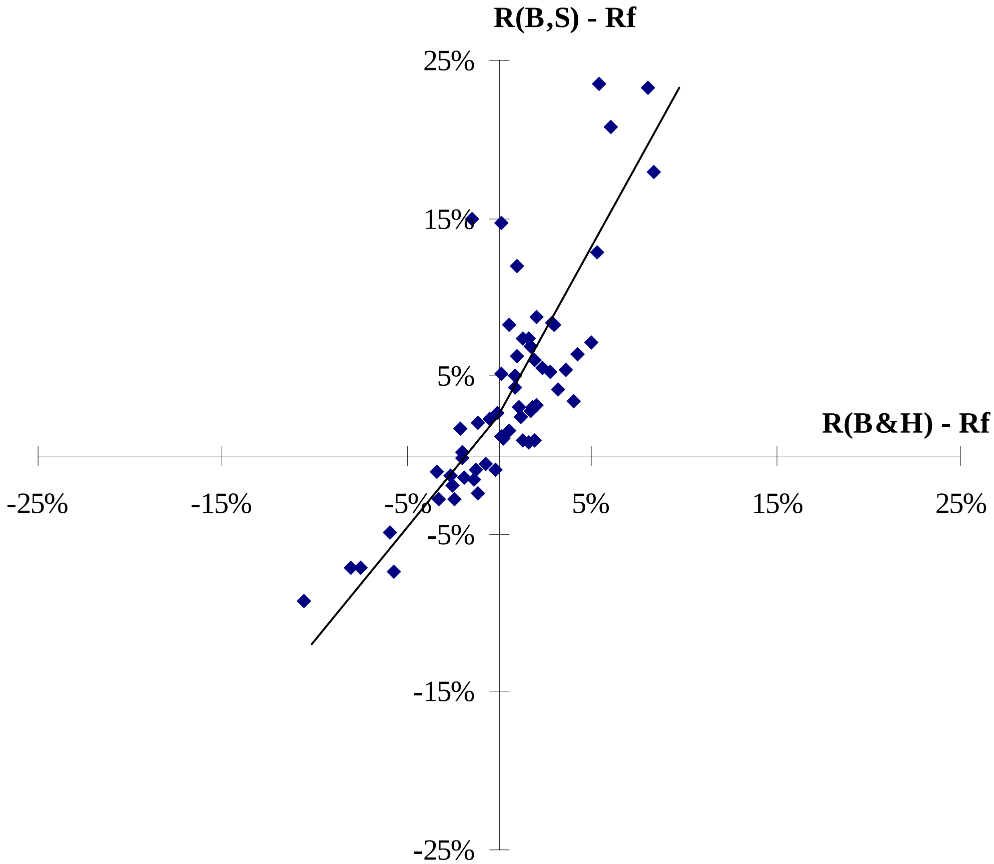


Figure 23 Scatter Plot of Excess R(B,S)=R(0.25%,0.12%) and Excess R(B&H) with Asymmetric Beta

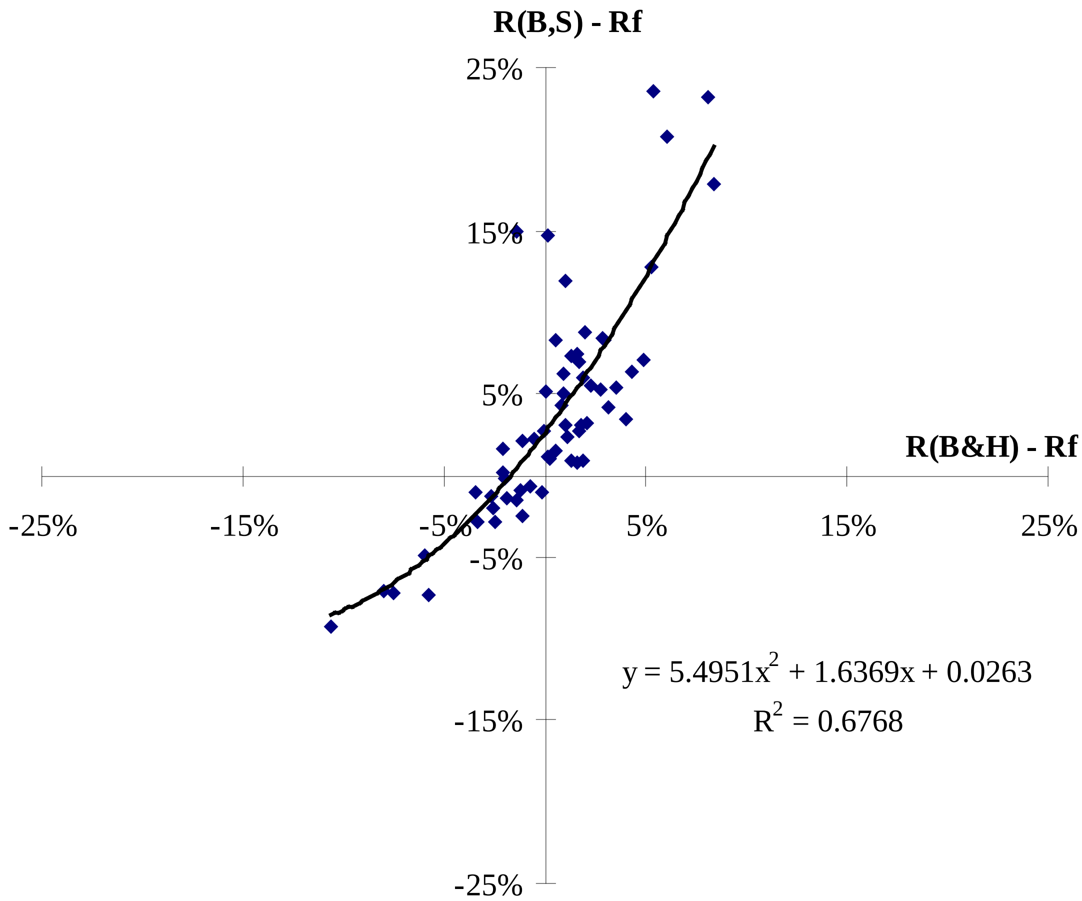


Figure 24 Scatter Plot of Excess R(B,S)=R(0.25%,0.12%) and Excess R(B&H) with Quadratic Fit Curve

4.1.5 Duration of a Trade and Proportion of time in the Market

This section is devoted to the analysis of the proportion of time the strategy holds a long position (in the market) versus not holding a position (out of the market). The impact of this on returns is further analyzed. The analysis in this section relies on *(B,S) = (0.25%,0.12%)* during 2002-2006.

Table 12 Descriptive Statistics of Monthly Transactions and Average Holding time from 2002-2006

	No. of Transactions	*Avg. Holding time / (Avg. Holding time + Avg. Idle time)*
Mean	83.53	0.7813
Standard Error	10.05	0.0175
Median	61.00	0.8021
Mode	18.00	-- *N/A* --
Standard Deviation	77.85	0.1359
Kurtosis	3.02	0.1387
Skewness	1.79	(0.8286)
Range	348	0.5144
Minimum	4	0.4515
Maximum	352	0.9659
Count	60	60

Source: compiled from Stata's output

An average holding time measures a period of time between a buy and sell order, or how long a position is open, or on the market. An average idle time measures a period of time between a sell order and a new buy order, or how long an investor holds no position, or is out of the market. Based on Table 12, on average, investors trading with the *(B,S) = (0.25%,0.12%)* strategy would hold a position in 78% of the time.

According to this construction, an increase in the frequency of transactions would increase the profitability of the strategy. The higher the frequency of transactions, the more favorable the price pattern is in that month. A low frequency of transactions would indicate the opposite. A low frequency of transactions should most likely occur when an investor is unable to close a position and gets stuck in a market downtrend.

Table 13 OLS Regression Analysis of R(B,S) and Market Participation

Dependent Variable: Returns on (B,S)=(0.25%,0.12%)

	Coefficients
Constant	0.04486*
	0.02412
No. of Transactions (per month)	0.00076**
	0.00005
Avg. Holding time /(Avg. Holding time + Avg. Idle time)	-0.08617**
	0.02897
Adj- R^2	0.82
Observation	60

Number in italic is Standard Errors of the estimates
** indicates significant at 10% level*
*** indicates significant at 1% level*
Avg. Holding time = Avg. time from Buy to Sell
Avg. Idle time = Avg. time from Sell to Buy

From Table 13, coefficients of both variables are highly significant. The Coefficient of No. of Transactions has a positive sign, consistent with the prediction. Based on these estimates, one round-trip transaction, Buy and Sell, would increase monthly returns by 0.076 percentage points, or 0.91 in annualized terms. The coefficient for the Avg. Holding time/ (Avg. Holding time + Avg. Idle time) is also significant and negative, revealing a tradeoff between staying in the market, or getting stuck in a price decline and missing out an opportunity to trade. The estimated trade-off implies that if an investor stays in the market for an additional 0.1 of the time, the monthly net returns will be 0.8 percentage points lower. This model captures 0.82 of the variation in monthly net returns of the strategy *(B,S) = (0.25%,0.12%).*

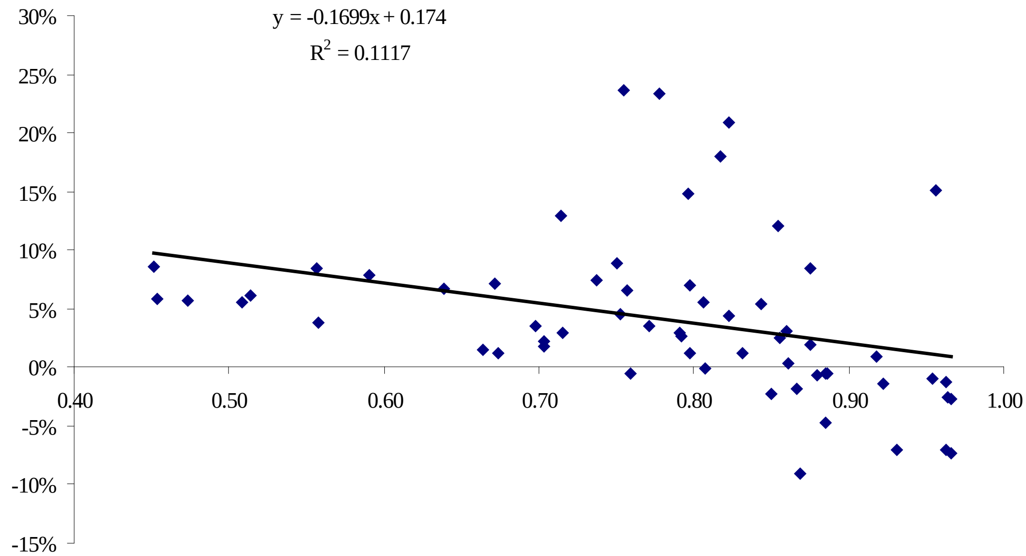


Figure 25 Scatter Plot of Proportion of Time In the Market and R(B,S)=R(0.12%,0.25%)

4.2 RETURNS FROM THE 0.3%, 0.2% AND 0.1% FILTERS

The previous subsections demonstrate the evidence that relies on Hasbrouck's 0.35% filter. In this section, the more appropriate filtering criterion is applied, that is lower than Hasbrouck's 0.35%. Results are then presented to the effect that returns and out-performance are highly sensitive to changes in filtering level. Hence, the decrease of area and size of returns due to lower filters will indicate that previous profitability stems from a trading mechanism that uses outliers, and not only legitimate prices.

The same trading algorithm is re-run, and returns from each combination are compiled and displayed in tables 14 to 16. Excess returns for the buy-and-hold strategy, using holding period returns with dividends as benchmarks are compiled into matrix tables 17 to 19. The darker shade of gray implies higher profitability.

Tables 14 to 16 show that the positive returns region shrinks in both area and size as narrower filters are adopted. Area and size are proportionately reduced in all three studied periods. This hints that outliers exist proportionately in all periods that are covered.

Now when the focus is shifted to comparing returns from trading rules *(Bi,Sj)* to benchmark returns from the buy-and-hold strategy. It is expected that the first and second period results to be similar or worse than those of section (3.1) due to the narrower filters we employ. As expected, tables 17 to 19 show that all strategies from first and second period performed 'poorer' when compared to the buy-and-hold strategy. In the third period, 2002-2006, the previous out-performance region also decreases markedly, both in area and size, as there is a switch from the 0.3% to 0.2% filter, and completely vanishes when we use the 0.1% filter. Based on this specific trading algorithm and results in this section, no evidence can be found to reject *Weak-Form Market Efficiency.*

It is now apparent that with the 0.1% filter, the computerized trading algorithm cannot outperform the buy-and-hold strategy in terms of net returns. Focus is again shifted in order to incorporate risk into the analysis of data with the 0.1% filter, to consider the case of return per unit of risk, or the Sharpe ratio.

Table 21, below, shows the benchmark of ex post Sharpe ratios in all three periods that have been studied (the higher the number, the better reward-to-risk strategy). The results turn out to be generally poorer when compared to the buy-and hold strategy's, and the Sharpe ratios from the buy-and-hold strategy outperform all strategies (Bi,Sj) in all three periods. It is therefore concluded that , even when risk from fluctuation of returns is factored in, the buy-and-hold strategy remains superior to returns from the computerized strategies. A final point to note in this section is that some of the Sharpe ratios are even negative, implying that the strategies had a lower return compared to putting money in a risk-free asset (all three sub-periods).

Table 14 Returns table from 0.3% Filter[23]

Average Annalized Returns form Strategy (Bi,Sj) during January 1993 - december 1996

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-7.7%	1.0%	4.9%	7.6%	8.8%	9.5%	9.9%	10.5%	10.6%	11.6%
0.2%	-5.6%	0.6%	4.2%	6.1%	7.4%	8.5%	8.8%	9.4%	10.4%	11.1%
0.3%	-3.8%	0.5%	3.0%	6.0%	6.9%	8.3%	9.1%	9.5%	10.0%	10.8%
0.4%	-3.8%	1.5%	2.7%	5.5%	7.7%	9.4%	10.4%	11.1%	10.9%	11.3%
0.5%	-2.0%	1.8%	2.7%	4.4%	6.2%	7.5%	8.7%	9.6%	9.2%	10.4%
0.6%	1.0%	2.7%	3.9%	5.2%	7.1%	6.8%	8.6%	9.3%	9.5%	9.3%
0.7%	-0.4%	2.5%	3.6%	4.3%	5.4%	6.5%	7.9%	8.6%	9.9%	10.1%
0.8%	-0.3%	0.5%	3.5%	5.4%	5.6%	6.7%	7.5%	7.8%	8.1%	9.3%
0.9%	0.3%	2.4%	1.1%	5.3%	5.3%	5.9%	7.0%	7.9%	7.9%	9.1%
1.0%	0.4%	1.4%	3.7%	4.3%	3.9%	4.6%	5.0%	6.0%	7.7%	7.7%

Average Annalized Returns form Strategy (Bi,Sj) during January 1997 - december 2001

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-126.9%	-27.9%	-13.6%	-7.8%	-4.3%	-2.1%	-0.7%	0.3%	1.1%	1.5%
0.2%	-96.8%	-23.9%	-11.5%	-5.9%	-2.7%	-0.8%	0.5%	1.0%	1.5%	1.8%
0.3%	-76.3%	-21.0%	-9.5%	-6.5%	-2.1%	-1.0%	0.5%	1.3%	2.3%	3.1%
0.4%	-64.2%	-20.0%	-9.7%	-4.8%	-1.9%	-0.3%	0.8%	2.4%	2.6%	3.6%
0.5%	-52.1%	-18.7%	-7.2%	-5.9%	-3.1%	-2.3%	0.4%	1.7%	0.9%	2.4%
0.6%	-50.7%	-18.5%	-7.2%	-5.2%	-4.7%	-2.9%	-1.1%	0.9%	2.1%	2.3%
0.7%	-41.3%	-18.1%	-10.0%	-5.9%	-3.3%	-3.7%	-2.7%	-0.3%	0.9%	3.4%
0.8%	-32.9%	-14.8%	-9.3%	-5.6%	-2.7%	-2.5%	-2.1%	-0.9%	0.7%	3.1%
0.9%	-29.1%	-11.9%	-8.5%	-4.3%	-2.5%	-1.1%	1.9%	1.9%	2.0%	3.2%
1.0%	-21.8%	-15.9%	-8.9%	-4.2%	-2.2%	-0.6%	1.2%	3.1%	3.7%	3.9%

Average Annalized Returns form Strategy (Bi,Sj) during January 2002 - december 2006

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	11.1%	6.0%	4.4%	3.3%	3.3%	3.2%	3.2%	2.4%	2.6%	2.8%
0.2%	26.7%	11.4%	7.8%	6.0%	4.5%	4.1%	4.7%	3.6%	4.0%	3.8%
0.3%	23.4%	13.9%	8.4%	7.4%	6.3%	5.3%	3.8%	4.0%	4.0%	3.9%
0.4%	12.2%	11.6%	8.4%	5.8%	6.0%	5.6%	4.2%	3.9%	3.7%	3.0%
0.5%	9.8%	7.7%	8.0%	6.1%	6.0%	6.1%	6.1%	4.1%	4.8%	4.5%
0.6%	5.1%	3.9%	5.1%	6.0%	6.0%	6.1%	5.3%	5.6%	4.7%	4.2%
0.7%	7.8%	3.3%	4.1%	4.7%	5.0%	5.2%	6.5%	5.2%	5.6%	4.4%
0.8%	1.5%	3.2%	2.6%	3.7%	4.4%	5.2%	5.5%	6.1%	6.6%	6.8%
0.9%	3.6%	1.9%	3.9%	1.4%	3.4%	3.8%	4.6%	5.4%	6.3%	6.0%
1.0%	-0.3%	-1.3%	1.8%	2.0%	1.7%	3.4%	4.7%	4.8%	4.5%	5.2%

[23] For this and following tables, darker shade of gray indicates higher profitable strategies.

Table 15 Returns table from 0.2% Filter

Average Annalized Returns form Strategy (Bi,Sj) during January 1993 - december 1996

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-8.3%	0.6%	4.6%	7.1%	8.2%	9.2%	9.7%	10.4%	10.5%	11.3%
0.2%	-6.1%	0.2%	3.8%	5.7%	6.8%	8.1%	8.4%	9.7%	10.5%	10.9%
0.3%	-4.7%	-0.1%	2.6%	5.8%	6.4%	8.0%	8.7%	9.5%	10.2%	10.7%
0.4%	-4.7%	0.9%	2.3%	5.3%	7.2%	8.9%	9.7%	10.7%	10.9%	10.6%
0.5%	-2.8%	1.2%	2.1%	3.9%	5.9%	7.3%	8.2%	9.1%	8.9%	9.8%
0.6%	-0.6%	3.0%	3.3%	4.7%	6.8%	6.7%	8.5%	9.1%	9.7%	9.2%
0.7%	-1.6%	2.1%	3.4%	4.0%	5.0%	6.3%	7.7%	8.5%	9.8%	9.7%
0.8%	-1.6%	0.4%	3.6%	5.4%	5.5%	6.6%	7.4%	8.0%	8.4%	9.4%
0.9%	-0.1%	2.2%	4.0%	5.2%	5.3%	5.7%	6.7%	8.2%	8.2%	9.1%
1.0%	0.0%	1.4%	3.7%	4.1%	3.9%	4.4%	4.7%	6.2%	7.9%	7.8%

Average Annalized Returns form Strategy (Bi,Sj) during January 1997 - december 2001

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-126.0%	-29.4%	-14.7%	-8.9%	-5.3%	-2.5%	-1.1%	0.1%	0.9%	0.9%
0.2%	-98.6%	-27.7%	-13.5%	-7.3%	-3.9%	-1.2%	0.0%	0.6%	1.0%	1.6%
0.3%	-78.9%	-26.1%	-13.0%	-8.2%	-4.1%	-1.7%	-0.4%	0.7%	1.4%	2.5%
0.4%	-69.5%	-24.9%	-13.6%	-7.5%	-4.2%	-1.6%	0.0%	1.8%	1.7%	2.7%
0.5%	-56.9%	-22.9%	-11.4%	-8.3%	-5.4%	-3.2%	-0.4%	1.6%	0.7%	2.2%
0.6%	-53.4%	-22.4%	-11.6%	-6.8%	-6.4%	-4.0%	-2.1%	0.7%	1.8%	1.7%
0.7%	-44.8%	-20.8%	-13.5%	-8.4%	-5.0%	-4.6%	-3.4%	-0.4%	0.7%	2.8%
0.8%	-35.7%	-16.6%	-10.8%	-7.7%	-4.5%	-3.4%	-2.8%	-1.4%	0.5%	2.6%
0.9%	-33.0%	-13.4%	-9.1%	-5.9%	-3.5%	-2.5%	0.4%	0.5%	1.0%	2.7%
1.0%	-25.0%	-17.4%	-9.9%	-5.4%	-3.7%	-1.4%	0.0%	1.7%	2.6%	3.1%

Average Annalized Returns form Strategy (Bi,Sj) during January 2002 - december 2006

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	6.8%	4.3%	3.6%	3.0%	2.9%	2.8%	2.8%	2.2%	2.6%	2.4%
0.2%	12.0%	5.8%	4.9%	4.3%	3.4%	3.0%	3.7%	3.1%	3.5%	3.7%
0.3%	6.1%	4.9%	4.1%	4.8%	3.9%	3.8%	2.8%	3.2%	3.2%	3.1%
0.4%	0.9%	4.2%	4.0%	3.6%	4.3%	4.5%	3.2%	2.7%	2.8%	1.9%
0.5%	0.3%	1.5%	4.1%	4.2%	5.1%	5.3%	5.5%	3.8%	4.3%	4.1%
0.6%	-1.4%	-0.8%	2.0%	4.2%	4.3%	5.0%	4.3%	4.7%	3.9%	3.4%
0.7%	3.2%	-0.4%	2.2%	3.6%	4.0%	4.5%	5.9%	5.0%	5.3%	4.3%
0.8%	-3.5%	0.5%	0.7%	3.0%	3.6%	4.5%	4.9%	5.6%	6.3%	6.4%
0.9%	-1.6%	-2.4%	2.4%	1.4%	2.9%	3.7%	4.4%	5.3%	6.1%	5.8%
1.0%	-2.1%	-3.3%	-0.2%	1.3%	1.5%	3.0%	4.7%	4.7%	4.5%	5.2%

Table 16 Returns table from 0.1% Filter

Average Annalized Returns form Strategy (Bi,Sj) during January 1993 - december 1996

	0.10%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%
0.10%	-8.5%	0.8%	5.0%	6.9%	8.2%	9.1%	9.7%	10.4%	10.7%	11.2%
0.20%	-6.9%	0.4%	3.9%	5.3%	7.4%	8.1%	8.5%	9.6%	10.9%	10.8%
0.30%	-5.9%	-0.3%	3.0%	5.1%	6.7%	7.8%	8.5%	9.5%	9.8%	10.6%
0.40%	-4.0%	0.6%	3.0%	4.9%	7.2%	8.9%	9.5%	10.3%	10.6%	10.3%
0.50%	-3.2%	1.2%	4.1%	3.9%	6.0%	7.3%	8.2%	9.0%	8.7%	9.6%
0.60%	-1.3%	3.6%	4.5%	4.5%	6.9%	6.9%	8.6%	9.3%	9.7%	9.3%
0.70%	-2.8%	1.9%	3.4%	3.4%	4.4%	6.1%	7.4%	8.3%	9.5%	9.7%
0.80%	-0.4%	1.2%	3.7%	5.0%	5.7%	6.5%	7.6%	8.0%	8.7%	9.4%
0.90%	-0.5%	1.9%	2.7%	4.9%	5.9%	5.9%	7.0%	8.5%	8.4%	9.0%
1.00%	0.0%	1.5%	3.9%	4.1%	5.0%	4.6%	4.8%	6.5%	8.0%	7.7%

Average Annalized Returns form Strategy (Bi,Sj) during January 1997 - december 2001

	0.10%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%
0.10%	-121.0%	-30.5%	-16.2%	-10.0%	-6.1%	-3.9%	-2.2%	-0.7%	0.0%	-0.3%
0.20%	-98.2%	-30.4%	-15.8%	-9.3%	-5.6%	-3.0%	-1.4%	-0.8%	0.2%	0.1%
0.30%	-83.6%	-29.3%	-15.7%	-10.2%	-6.1%	-3.5%	-2.8%	-1.2%	0.3%	1.3%
0.40%	-72.8%	-27.6%	-15.8%	-9.6%	-5.8%	-3.5%	-1.8%	-0.1%	0.3%	1.2%
0.50%	-59.8%	-25.9%	-13.4%	-9.7%	-6.8%	-5.0%	-1.8%	0.0%	-0.8%	0.6%
0.60%	-55.6%	-24.2%	-13.3%	-7.8%	-7.2%	-5.1%	-3.4%	-0.3%	0.4%	0.4%
0.70%	-46.9%	-22.7%	-14.5%	-9.2%	-5.3%	-5.6%	-4.4%	-0.9%	0.3%	1.9%
0.80%	-37.5%	-18.6%	-13.7%	-9.0%	-5.6%	-3.9%	-3.6%	-2.2%	-0.2%	2.0%
0.90%	-35.8%	-15.2%	-10.8%	-7.1%	-4.7%	-3.7%	-1.0%	-0.3%	0.2%	2.0%
1.00%	-26.3%	-18.5%	-11.2%	-6.5%	-4.9%	-2.9%	-1.4%	0.9%	1.8%	2.3%

Average Annalized Returns form Strategy (Bi,Sj) during January 2002 - december 2006

	0.10%	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%
0.10%	-2.9%	0.2%	1.4%	1.3%	1.9%	2.1%	2.1%	1.9%	2.1%	1.9%
0.20%	-4.6%	-0.4%	1.5%	2.1%	1.6%	1.9%	2.4%	2.0%	2.9%	2.8%
0.30%	-3.9%	-1.2%	1.4%	2.6%	2.8%	2.6%	2.2%	2.3%	2.7%	2.6%
0.40%	-6.4%	-0.1%	1.6%	2.0%	2.8%	3.3%	2.6%	2.0%	1.8%	1.4%
0.50%	-6.0%	-0.8%	2.9%	3.0%	4.1%	4.4%	4.6%	3.6%	3.5%	3.3%
0.60%	-5.3%	-2.3%	1.3%	3.4%	3.5%	4.0%	3.9%	4.3%	3.3%	2.5%
0.70%	-2.8%	-2.4%	1.1%	2.6%	3.7%	3.7%	5.4%	4.6%	4.8%	4.1%
0.80%	-6.2%	-0.6%	-0.2%	1.9%	3.1%	4.0%	4.3%	5.3%	5.7%	5.6%
0.90%	-4.8%	-3.4%	1.7%	1.1%	2.8%	3.3%	3.9%	4.5%	5.5%	5.2%
1.00%	-7.0%	-4.4%	-0.3%	0.8%	1.3%	2.9%	4.2%	4.2%	4.0%	4.8%

Table 17 Out-performance Region from 0.3% Filter

Average Annalized Returns form Strategy (Bi,Sj) during January 1993 - december 1996

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-23.2%	-14.5%	-10.6%	-8.0%	-6.8%	-6.1%	-5.7%	-5.0%	-4.9%	-3.9%
0.2%	-21.1%	-14.9%	-11.3%	-9.4%	-8.1%	-7.1%	-6.8%	-6.1%	-5.2%	-4.4%
0.3%	-19.4%	-15.1%	-12.6%	-9.5%	-8.7%	-7.3%	-6.5%	-6.0%	-5.5%	-4.7%
0.4%	-19.3%	-14.0%	-12.9%	-10.1%	-7.9%	-6.2%	-5.2%	-4.5%	-4.6%	-4.3%
0.5%	-17.5%	-13.7%	-12.8%	-11.1%	-9.3%	-8.1%	-6.8%	-5.9%	-6.4%	-5.2%
0.6%	-14.6%	-12.8%	-11.7%	-10.3%	-8.5%	-8.8%	-7.0%	-6.2%	-6.1%	-6.2%
0.7%	-16.0%	-13.0%	-12.0%	-11.2%	-10.2%	-9.1%	-7.7%	-6.9%	-5.7%	-5.5%
0.8%	-15.8%	-15.0%	-12.0%	-10.1%	-10.0%	-8.8%	-8.0%	-7.8%	-7.4%	-6.3%
0.9%	-15.2%	-13.2%	-11.5%	-10.2%	-10.2%	-9.7%	-8.6%	-7.6%	7.6%	-6.5%
1.0%	-15.2%	-14.2%	-11.8%	-11.3%	-11.6%	-10.9%	-10.5%	-9.6%	-7.8%	-7.9%

Average Annalized Returns form Strategy (Bi,Sj) during January 1997 - december 2001

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-138.4%	-39.5%	-25.2%	-19.4%	-15.9%	-13.7%	-12.2%	-11.2%	-10.5%	-10.1%
0.2%	-108.3%	-35.5%	-23.1%	-17.5%	-14.3%	-12.3%	-11.0%	-10.6%	-10.0%	-9.8%
0.3%	-87.8%	-32.6%	-21.0%	-18.1%	-13.7%	-12.5%	-11.0%	-10.3%	-9.3%	-8.4%
0.4%	-75.8%	-31.5%	-21.3%	-16.3%	-13.5%	-11.9%	-10.8%	-9.2%	-9.0%	-7.9%
0.5%	-63.7%	-30.3%	-18.7%	-17.5%	-14.7%	-13.8%	-11.1%	-9.9%	-10.7%	-9.1%
0.6%	-62.3%	-30.1%	-18.7%	-16.7%	-16.3%	-14.4%	-12.7%	-10.7%	-9.5%	-9.3%
0.7%	-52.8%	-29.7%	-21.6%	-17.5%	-14.9%	-15.3%	-14.2%	-11.8%	-10.6%	-8.1%
0.8%	-44.5%	-26.4%	-20.8%	-17.2%	-14.3%	-14.1%	-13.6%	-12.5%	-10.8%	-8.5%
0.9%	-40.6%	-23.4%	-20.1%	-15.8%	-14.1%	-12.7%	-9.7%	-9.6%	-9.5%	-8.3%
1.0%	-33.3%	-27.5%	-20.4%	-15.7%	-13.8%	-12.2%	-10.3%	-8.5%	-7.8%	-7.7%

Average Annalized Returns form Strategy (Bi,Sj) during January 2002 - december 2006

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	4.3%	-0.8%	-2.4%	-3.5%	-3.5%	-13.7%	-12.2%	-11.2%	-10.5%	-10.1%
0.2%	19.9%	4.6%	1.0%	-0.8%	-2.3%	-12.3%	-11.0%	-10.6%	-10.0%	-9.8%
0.3%	16.6%	7.1%	1.6%	0.6%	-0.5%	-12.5%	-11.0%	-10.3%	-9.3%	-8.4%
0.4%	5.5%	4.8%	1.6%	-1.0%	-0.8%	-11.9%	-10.8%	-9.2%	-9.0%	-7.9%
0.5%	3.0%	0.9%	1.2%	-0.7%	-0.8%	-13.8%	-11.1%	-9.9%	-10.7%	-9.1%
0.6%	-1.7%	-2.9%	-1.7%	-0.8%	-0.8%	-14.4%	-12.7%	-10.7%	-9.5%	-9.3%
0.7%	1.0%	-3.5%	-2.7%	-2.1%	-1.8%	-15.3%	-14.2%	-11.8%	-10.6%	-8.1%
0.8%	-5.3%	-3.6%	-4.2%	-3.1%	-2.4%	-14.1%	-13.6%	-12.5%	-10.8%	-8.5%
0.9%	-3.2%	-4.9%	-2.9%	-5.4%	-3.4%	-12.7%	-9.7%	-9.6%	-9.5%	-8.3%
1.0%	-7.1%	-8.1%	-5.0%	-4.8%	-5.1%	-12.2%	-10.3%	-8.5%	-7.8%	-7.7%

Table 18 Out-performance Region from 0.2% Filter

Average Annalized Returns form Strategy (Bi,Sj) during January 1993 - december 1996

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-23.9%	-14.9%	-11.0%	-8.4%	-7.1%	-6.3%	-5.9%	-5.1%	-5.0%	-4.3%
0.2%	-21.7%	-15.3%	-11.7%	-9.8%	-8.8%	-7.4%	-7.2%	-5.9%	-5.0%	-4.6%
0.3%	-20.2%	-15.7%	-13.0%	-9.7%	-9.1%	-7.6%	-6.9%	-6.0%	-5.4%	-4.9%
0.4%	-20.2%	-14.6%	-13.3%	-10.2%	-8.3%	-6.6%	-5.8%	-4.9%	-4.7%	-4.9%
0.5%	-18.3%	-14.4%	-13.4%	-11.6%	-9.7%	-8.3%	-7.4%	-6.5%	-6.6%	-5.7%
0.6%	-16.1%	-12.6%	-12.2%	-10.8%	-8.7%	-8.8%	-7.0%	-6.5%	-5.9%	-6.3%
0.7%	-17.2%	-13.5%	-12.1%	-11.6%	-10.5%	-9.2%	-7.8%	-7.0%	-5.7%	-5.8%
0.8%	-16.9%	-15.1%	-11.9%	-10.1%	-10.0%	-9.0%	-8.2%	-7.5%	-7.1%	-6.2%
0.9%	-15.7%	-13.4%	-11.5%	-10.3%	-10.3%	-9.8%	-8.9%	-7.4%	-7.3%	-6.4%
1.0%	-15.5%	-14.1%	-11.8%	-11.4%	-11.7%	-10.9%	-10.9%	-9.4%	-7.6%	-7.8%

Average Annalized Returns form Strategy (Bi,Sj) during January 1997 - december 2001

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-137.6%	-40.9%	-26.2%	-20.5%	-16.8%	-14.0%	-12.6%	-11.4%	-10.6%	-10.6%
0.2%	-110.1%	-39.2%	-25.0%	-18.9%	-15.5%	-12.8%	-11.6%	-10.9%	-10.5%	-10.0%
0.3%	-90.4%	-37.7%	-24.5%	-19.8%	-15.6%	-13.3%	-11.9%	-10.9%	-10.1%	-9.1%
0.4%	-81.0%	-36.5%	-25.2%	-19.1%	-15.8%	-13.2%	-11.5%	-9.8%	-9.9%	-8.9%
0.5%	-68.5%	-34.4%	-22.9%	-19.8%	-16.9%	-14.7%	-11.9%	-9.9%	-10.8%	-9.4%
0.6%	-65.0%	-33.9%	-23.2%	-18.3%	-18.0%	-15.6%	-13.6%	-10.8%	-9.8%	-9.8%
0.7%	-56.3%	-32.4%	-25.1%	-20.0%	-16.5%	-16.1%	-15.0%	-11.9%	-10.9%	-8.8%
0.8%	-47.2%	-28.1%	-22.4%	-19.3%	-16.0%	14.9%	-14.3%	-12.9%	-11.0%	-9.0%
0.9%	-44.5%	-25.0%	-20.7%	-17.5%	-15.0%	-14.0%	-11.2%	-11.1%	-10.5%	-8.9%
1.0%	-36.5%	-29.0%	-21.4%	-16.9%	-15.2%	-13.0%	-11.5%	-9.9%	-9.0%	-8.4%

Average Annalized Returns form Strategy (Bi,Sj) during January 2002 - december 2006

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	0.0%	-2.5%	-3.2%	-3.8%	-3.9%	-4.0%	-4.0%	-4.6%	-4.2%	-4.4%
0.2%	5.2%	-1.0%	-1.9%	-2.5%	-3.4%	-3.8%	-3.1%	-3.7%	-3.3%	-3.1%
0.3%	-0.7%	-1.9%	-2.7%	-2.0%	-2.9%	-3.0%	-4.0%	-3.6%	-3.6%	-3.7%
0.4%	-5.9%	-2.6%	-2.8%	-3.2%	-2.5%	-2.3%	-3.6%	-4.1%	-4.0%	-4.9%
0.5%	-6.5%	-5.3%	-2.7%	-2.6%	-1.7%	-1.5%	-1.3%	-3.0%	-2.5%	-2.7%
0.6%	-8.2%	-7.6%	-4.8%	-2.6%	-2.5%	-1.8%	-2.5%	-2.1%	-2.9%	-3.4%
0.7%	-3.6%	-7.2%	-4.6%	-3.2%	-2.8%	-2.3%	-0.9%	-1.8%	-1.5%	-2.5%
0.8%	-10.2%	-6.3%	-6.1%	-3.8%	-3.2%	-2.3%	-1.9%	-1.2%	-0.5%	-0.4%
0.9%	-8.4%	-9.2%	-4.4%	-5.4%	-3.9%	-3.1%	-2.4%	-1.5%	-0.7%	-1.0%
1.0%	-8.9%	-10.1%	-7.0%	-5.5%	-5.3%	-3.8%	-2.1%	-2.0%	-2.3%	-1.6%

Table 19 Out-performance Region from 0.1% Filter

Average Annalized Returns form Strategy (Bi,Sj) during January 1993 - december 1996

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-24.0%	-14.8%	-10.6%	-8.7%	-7.3%	-6.5%	-5.8%	-5.2%	-4.9%	-4.4%
0.2%	-22.5%	-15.1%	-11.6%	-10.2%	-8.1%	-7.5%	-7.0%	-5.9%	-4.7%	-4.7%
0.3%	-21.4%	-15.8%	-12.5%	-10.5%	-8.9%	-7.7%	-7.0%	-6.1%	-5.7%	-4.9%
0.4%	-19.6%	-15.0%	-12.6%	-10.6%	-8.4%	-6.6%	-6.1%	-5.3%	-4.9%	-5.3%
0.5%	-18.8%	-14.4%	-11.5%	-11.6%	-9.5%	-8.3%	-7.3%	-6.5%	-6.9%	-6.0%
0.6%	-16.8%	-11.9%	-11.0%	-11.1%	-8.6%	-8.6%	-7.0%	-6.2%	-5.8%	-6.3%
0.7%	-18.3%	-13.6%	-12.2%	-12.1%	-11.1%	-9.4%	-8.2%	-7.2%	-6.0%	-5.9%
0.8%	-15.9%	-14.3%	-11.8%	-10.6%	-9.9%	-9.1%	-8.0%	-7.5%	-6.8%	-6.1%
0.9%	-14.4%	-13.4%	-11.7%	-9.9%	-9.3%	-9.3%	-7.6%	-6.6%	-7.3%	-6.1%
1.0%	-15.5%	-14.1%	-11.6%	-11.5%	-10.5%	-11.0%	-10.8%	-9.0%	-7.6%	-7.9%

Average Annalized Returns form Strategy (Bi,Sj) during January 1997 - december 2001

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-132.6%	-42.1%	-27.7%	-21.6%	-17.6%	-15.5%	-13.7%	-12.2%	-11.6%	-11.8%
0.2%	-109.7%	-41.9%	-27.4%	-20.9%	-17.1%	-14.5%	-12.9%	-12.3%	-11.4%	-11.5%
0.3%	-95.1%	-40.9%	-27.3%	-21.8%	-17.7%	-15.0%	-14.3%	-12.7%	-11.3%	-10.3%
0.4%	-84.4%	-39.1%	-27.3%	-21.2%	-17.3%	-15.1%	-13.4%	-11.7%	-11.3%	-10.4%
0.5%	-71.4%	-37.4%	-25.0%	-21.3%	-18.4%	-16.5%	-13.3%	-11.5%	-12.3%	-10.9%
0.6%	-67.2%	-35.7%	-24.9%	-19.3%	-18.7%	-16.6%	-15.0%	-11.9%	-11.2%	-11.1%
0.7%	-58.4%	-34.2%	-26.0%	-20.8%	-16.8%	-17.1%	-15.9%	-12.4%	-11.3%	-9.6%
0.8%	-49.1%	-30.2%	-25.3%	-20.5%	-17.2%	-15.5%	-15.1%	-13.7%	-11.8%	-9.6%
0.9%	-35.2%	-23.8%	-20.3%	-17.0%	-16.1%	-14.3%	-12.5%	-12.9%	-10.7%	-8.6%
1.0%	-37.9%	-30.1%	-22.7%	-18.1%	-16.4%	-14.4%	-13.0%	-10.6%	-9.8%	-9.2%

Average Annalized Returns form Strategy (Bi,Sj) during January 2002 - december 2006

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-9.7%	-6.6%	-5.4%	-5.5%	-4.9%	-4.7%	-4.7%	-4.9%	-4.7%	-4.9%
0.2%	-11.4%	-7.2%	-5.3%	-4.7%	-5.2%	-4.9%	-4.4%	-4.8%	-3.9%	-4.0%
0.3%	-10.7%	-8.0%	-5.4%	-4.2%	-4.0%	-4.2%	-4.6%	-4.5%	-4.1%	-4.2%
0.4%	-13.2%	-6.9%	-5.2%	-4.8%	-4.0%	-3.5%	-4.2%	-4.8%	-5.0%	-5.4%
0.5%	-12.8%	-7.6%	-3.9%	-3.8%	-2.7%	-2.4%	-2.2%	-3.2%	-3.3%	-3.5%
0.6%	-12.1%	-9.1%	-5.5%	-3.4%	-3.3%	-2.8%	-2.9%	-2.5%	-3.5%	-4.2%
0.7%	-9.6%	-9.2%	-5.7%	-4.2%	-3.1%	-3.1%	-1.4%	-2.2%	-2.0%	-2.7%
0.8%	-13.0%	-7.4%	-7.0%	-4.9%	-3.7%	-2.8%	-2.5%	-1.5%	-1.1%	-1.2%
0.9%	-8.9%	-7.9%	-5.7%	-4.4%	-3.3%	-3.4%	-2.7%	-2.1%	-1.5%	-1.3%
1.0%	-13.8%	-11.2%	-7.1%	-6.0%	-5.5%	-3.9%	-2.6%	-2.6%	-2.8%	-2.0%

Table 20 Benchmark Ex Post Sharpe Ratios

	Buy&Hold Returns incl. Dividends	Buy&Hold Returns excl. Dividends	Value-weighted Returns incl. Dividends	Value-weighted Returns excl. Dividends	Equal-weighted Returns incl. Dividends	Equal-weighted Returns excl. Dividends	S&P Composite Returns
Jan93-Dec96	0.49	0.42	0.50	0.42	0.39	0.35	0.45
Jan97-Dec01	0.19	0.17	0.17	0.15	0.16	0.14	0.17
Jan02-Dec06	0.16	0.12	0.20	0.16	0.33	0.30	0.12

Table 21 Ex Post Sharpe Ratios based on 0.1% Filter

Shrape Ratios from January 1993 - December 1996

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-63%	-15%	2%	8%	13%	15%	17%	19%	20%	22%
0.2%	-63%	-17%	-2%	3%	11%	13%	14%	17%	21%	21%
0.3%	-63%	-22%	-6%	2%	8%	12%	14%	17%	18%	21%
0.4%	-50%	-20%	-6%	2%	10%	16%	17%	20%	20%	20%
0.5%	-53%	-16%	-2%	-2%	6%	10%	13%	16%	15%	17%
0.6%	-44%	-5%	0%	0%	10%	9%	15%	17%	18%	16%
0.7%	-55%	-15%	-6%	-5%	0%	7%	11%	14%	18%	18%
0.8%	-48%	-20%	-4%	2%	6%	8%	12%	14%	16%	18%
0.9%	-52%	-19%	-10%	2%	7%	6%	10%	15%	15%	17%
1.0%	-42%	-22%	-4%	-2%	3%	1%	1%	8%	13%	12%

Shrape Ratios from January 1997 - December 2001

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-160%	-84%	-43%	-29%	-20%	-16%	-13%	-10%	-9%	-9%
0.2%	-166%	-83%	-42%	-27%	-19%	-14%	-11%	-10%	-8%	-8%
0.3%	-158%	-83%	-44%	-30%	-21%	-16%	-14%	-11%	-8%	-6%
0.4%	-164%	-81%	-45%	-29%	-20%	-15%	-12%	-9%	-8%	-6%
0.5%	-150%	-75%	-38%	-29%	-22%	-18%	-12%	-8%	-10%	-7%
0.6%	-142%	-69%	-39%	-25%	-23%	-19%	-15%	-9%	-8%	-8%
0.7%	-133%	-71%	-43%	-30%	-20%	-20%	-17%	-10%	-8%	-5%
0.8%	-123%	-62%	-43%	-30%	-21%	-17%	-16%	-13%	-9%	-5%
0.9%	-127%	-55%	-37%	-26%	-20%	-17%	-11%	-10%	-8%	-5%
1.0%	-108%	-60%	-38%	-25%	-20%	-15%	-12%	-7%	-5%	-4%

Shrape Ratios from January 2002 - December 2006

	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%	1.0%
0.1%	-13%	-5%	-2%	-3%	-1%	-1%	-1%	-1%	-1%	-1%
0.2%	-18%	-7%	-2%	-1%	-2%	-1%	0%	-1%	1%	1%
0.3%	-16%	-9%	-2%	1%	1%	1%	-1%	0%	1%	1%
0.4%	-25%	-6%	-2%	-1%	1%	2%	1%	-1%	-1%	-2%
0.5%	-24%	-8%	1%	2%	4%	5%	5%	3%	3%	2%
0.6%	-24%	-12%	-3%	3%	3%	4%	4%	5%	2%	0%
0.7%	-17%	-13%	-3%	1%	3%	3%	7%	5%	6%	4%
0.8%	-35%	-8%	-7%	-1%	2%	4%	5%	7%	8%	8%
0.9%	-30%	-18%	-2%	-3%	1%	2%	4%	5%	8%	7%
1.0%	-33%	-22%	-8%	-4%	-3%	1%	4%	5%	4%	6%

5 CONCLUSION

The first section discussed motivation, background, and the data source of this paper. In reviewing the literature, most previous contributions suggest that trading strategies usually fail to outperform the buy-and-hold strategy after factoring in transaction costs.

Data is discussed in section two. Section three describes methodology by, first, addressing components within transaction costs, and, second, by analyzing the Bid/Ask Spread. It is concluded that the spread component has varied widely throughout the period observed. This finding leads to the decision to break the entire trading period into roughly three sub-periods with very different costs structures. Then, a detailed description of a simple technical trading algorithm, comprising a buying rule and a selling rule, is discussed; the calculation of returns is examined, ending section three with a discussion on risk-adjusted returns, and filtering technique.

The main results in section four are threefold. First, it is shown that the profitability of a simple technical trading strategy is critically dependent on filtering of the data. Second, it is shown that the wider 50 cent filter of Hasbrouck (2003) is not appropriate to use because it leaves in too many outliers and mistakes. In fact, using this data filter, it was found to be spectacularly positive out-performance for the simple technical strategy. However, the positive region is not robust as the data is treated with narrower filters of 0.3% and 0.2%. Third, all previously profitable strategies disappear with the 0.1% filter, the smallest filter used, and the most appropriate. This shows that the 0.1% filter is not too wide. When risk is factored in, the buy-and-hold strategy remains superior to any of the discussed technical trading strategies, and this result is robust across all three sub-periods. Having much of profitability hinging on a wide filter level casts a shadow on the validity of results from studies that employ a wider band, such as 50 cents. Short term weak form of market efficiency can never really be tested when there is uncertainty as to what filtering criterion should be used. A wide-filter, and a more conservative one lead to a completely opposite conclusions.

There needs to be more research on the optimal filtering of the TAQ dataset and other ultra-high frequency datasets in general. It was clearly shown in this study that a widely used filter is not appropriate because it is too lax. However, it is not clear what level of stringency should be used for filtering intraday data. A narrow filter will remove outliers and mistakes. However, it will also remove valid prices. This lies within the classical context of a trade-off between a type I error (removing a valid price) and a type II error (keeping an erroneous price). The appropriate trade-off should be examined. Perhaps, more sophisticated filtering rules (than deviations from moving averages) should be explored to reach an optimal trade off.

REFERENCES

Alexander, S. (1961). Price Movements in Speculative Markets: Trends or Random Walks. *Industrial Management Review* , 7-26.

Alexander, S. (1964). Price Movements in Speculative Markets: Trends or Random Walks, No.2. *Industrial Management Review* , 25-64.

Fama, E. (1970). Efficienct Capital Markets: A review of Theory and Empirical Work. *Journal of Finance* , 383-417.

Fama, E., & Blume, M. (1966). Filter Rules and Stock-Market Trading. *Journal of Business* , 226-241.

Farrell, C. H., & Olszewski, E. A. (1993). Assessing Inefficiency in the S&P 500 Futures Market. *Journal of Forecasting* , 395-420.

Hasbrouck, J. (2003). Intraday Price Formation in U.S. Equity Index Markets. *Journal of Finance* , 2375-2399.

Levich, R. M., & Thomas, L. R. (1993). The Significance of Technical Trading Rule Profits in the Foreign Exchange Market: A Bootstrap Approach. *Journal of International Money and Finance* , 451-474.

Markowitz, H. (1987). *Mean-Variance Analysis in Portfolio Choice and Capital Markets.* New York: Basil Blackwell Inc.

Park, C.-H., & Irwin, S. H. (2004). *The Profitability of Technical Analysis: A Review.* Urbana-Champaign: AgMAS Project Research Report 2004-04.

Sharpe, W. (1964). Capital Asset Prices: A theory of Market Equilibrium Under Conditions of Risk. *Journal of Finance* , 425-442.

Sharpe, W. (1966). Mutual Fund Performance. *Journal of Business* , 119-138.

Sharpe, W. (1994). The Sharpe Ratio. *Journal of Portfolio Management* , 49-58.