

University of New Mexico
UNM Digital Repository

Economics ETDs

Electronic Theses and Dissertations

9-5-2013

The Law of One Price and Virtual Worlds

Michael Morrison

Follow this and additional works at: https://digitalrepository.unm.edu/econ_etds

Recommended Citation

Morrison, Michael. "The Law of One Price and Virtual Worlds." (2013). https://digitalrepository.unm.edu/econ_etds/11

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Economics ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Michael R Morrison

Candidate

Economics

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Matías Fontenla, Co-Chairperson

Dr. Christine Sauer, Co-Chairperson

Dr. Alok Bohara

Dr. Christopher Butler

THE LAW OF ONE PRICE AND VIRTUAL WORLDS

by

MICHAEL R. MORRISON

B.S. Economics, Montana State University, 2004
M.A. Economics, University of New Mexico, 2012

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Economics

The University of New Mexico
Albuquerque, New Mexico

July, 2013

Dedication

This dissertation is dedicated to my parents, Paul and Sherry Morrison as well as my Grandparents, Henry and Geraldine Quintero, for showing me what can be accomplished through hard work and dedication. The examples you've set and the support you've given has motivated me throughout my college career.

Acknowledgements

I would like to thank my dissertation committee co-chairs, Dr. Matías Fontenla and Dr Christine Sauer, for all of their help and support throughout this process. If not for Dr. Fontenla I would not have had the courage to study virtual worlds. His support and encouragement from my first class through our first paper and finally my dissertation has been invaluable. I want to thank Dr. Sauer for all of her advice and help in developing my dissertation and early versions of my first publication. Without her aid I would never have accomplished this task. I would also like to thank the rest of my dissertation committee members, Dr. Alok Bohara and Dr. Christopher Butler, for their valuable advice and recommendations for my dissertation.

I would like to thank Erorus, the editor of “The Undermine Journal” for providing the data for my second chapter and the Office of Graduate Studies for the Research and Travel Projects Grant for funding Erorus’ efforts.

I would like to thank the librarians at Parish Library, especially Susan Awe, and the Data Dissemination and Client Services Team at the International Monetary Fund, for going above and beyond to provide the data I needed despite numerous technical problems.

I would also like to thank the Southern Economics Association and the Economics Department at UNM for the Sponsored Graduate Student Presentation in New Orleans. The presentation provided many useful comments and criticism of my work.

Finally I would like to thank all of my friends and family for supporting me throughout this process. Most especially I would like to thank, Jeff Felardo, Dennis Barber III, Brad Bergsbaken, Lenny Raper, Annie Theodoropoulos, and Terry Kelly. I would also like to thank the makers of Red Bull for giving me the energy I needed when I had none.

THE LAW OF ONE PRICE AND VIRTUAL WORLDS

BY

MICHAEL R. MORRISON

B.S. Economics, Montana State University, 2004
M.A. Economics, University of New Mexico, 2012
Ph.D. Economics, University of New Mexico, 2013

Abstract

Price convergence and purchasing power parity are two important foundations to international trade theory. Yet there is still no consensus on their existence, despite numerous studies. While the debate over price convergence and purchasing power parity continues, the last several years have seen a substantial rise in the use of virtual worlds. Virtual worlds are interactive digital environments where individuals with an internet connection can come together to communicate, cooperate and trade. Virtual worlds create smaller copies of the real world. These copies lack many of the complexities found in the real world. This feature makes virtual worlds a unique setting for studying price convergence and other economic phenomena.

This dissertation begins by surveying the extant research on virtual worlds throughout the social sciences to illustrate the usefulness of virtual worlds in economics research. In the second chapter price convergence is tested in one of these virtual worlds. Insight gained from the second chapter is used in the final chapter to alter an existing purchasing power parity model. This altered model is used to estimate the effect transportation costs on purchasing power parity equilibrium.

TABLE OF CONTENTS

LIST OF FIGURES	viii
LIST OF TABLES	ix
Introduction.....	1
Chapter 1 Virtual World Research in the Social Sciences: Opportunities for Economic Research.....	6
1.1 Introduction	7
1.1.1 Demographics	9
1.2 VW Research.....	11
1.2.1 What’s Been Done in Economics?	13
1.2.2 Research Outside of Economics	15
1.2.2.1 Motivation	15
1.2.2.2 Identity.....	23
1.2.2.3 Behavior.....	27
1.2.2.4 Social Networks and Culture.....	34
1.3 Conclusion.....	43
Chapter 2 Price Convergence after Cataclysm	46
2.1 Introduction	47
2.1.1 Virtual Worlds and Economic Research	49
2.1.2 Price Convergence	53
2.2 Environment and Agents.....	56
2.2.1 Environment	56
2.2.2 Agents.....	57
2.3 Data	58
2.3.1 Data Collection and Cleaning.....	60
2.4 Models and Results	60
2.4.1 Engle–Granger Cointegration.....	61
2.4.1.A The Model	61
2.4.1.B Results	63
2.4.2 Exponential Smooth Transition Autoregressive (ESTAR)	64
2.4.2.A The Model	64

2.4.2.B Results	66
2.4.3 Phillips and Sul	68
2.4.3.A The Model	68
2.4.3.B Results	69
2.4.4 Band of Inaction (BofI) test.....	72
2.4.4.A The Model	72
2.4.4.B Results	75
2.5 Conclusion.....	76
Chapter 3 Variable Band of Inaction: Transaction Costs and the Size of the Band of Inaction	79
3.1 Introduction	80
3.2 Purchasing Power Parity and Arbitrage	81
3.2.1 Purchasing Power Parity Model and Estimation Method.....	82
3.3 Data	87
3.3.1 Transportation Costs	89
3.4 MP-ESTAR model of PPP and the BofI.....	94
3.5 Conclusion.....	98
Conclusion	100
References	105
Appendix A	111
Appendix B	116
Appendix C	120

LIST OF FIGURES

Figure 2.1 Cross-sectional variances for three indices in World 1 (Azjol-Nerub)	71
Figure 2.2 Price index values and BofI for World 5 Frostmane.	75
Figure 3.3 Graph of generated transportation cost values over time.	92
Figure 3.4 Transition Functions from the First and Last 10Years of the Sample.....	96
Figure 3.5 Transition functions, $F(y(t-d), \tau(1,t))$ for Model 1 and Model 2	97

LIST OF TABLES

Table 2.1 Results of the Engle-Granger Cointegration Test.....	63
Table 2.2 Results from ESTAR Model Tests	67
Table 2.3 Results from Phillips and Sul tests	70
Table 2.4. T-values for BofI test.....	76
Table 3.5 Summary CPI and RPI data from 1974 to 2010	89
Table 3.6. Summary of Transportation Cost data.....	93
Table 3.7 Panel ESTAR regression of non-Euro countries in the first and last 10 years of the sample.	95
Table 3.8 Modified panel ESTAR regression results for Model 1 and Model 2	97

Introduction

Price convergence and the related concept purchasing power parity (PPP) are two of the most important theoretical foundations of international trade theory. There is, however, little consensus on their existence. Price convergence states that, in the absence of transaction costs, the price level in any two markets will be equal. If there are any price differentials, the difference will be eliminated through arbitrage. PPP holds when this convergence occurs across currencies. There have been several recent studies that fail to find support for either price convergence or PPP (de Carvalho & Júlio, 2012; Jacobson & Nessén, 2004). Others have found support, but differ on the duration of price deviations (Fischer, 2009; Kim & Kim, 2011; Alan M. Taylor, 2001). Reasons for these inconsistent results include capital market imperfections, different productivity growth rates (also called the Balassa-Samuelson effect) (Balassa, 1964), transaction costs, and several others (Hallwood & MacDonald, 2000).

The introduction of virtual worlds (VWs) creates a unique opportunity study price convergence, and gain a greater understanding of how it works. VWs are three dimensional interactive environments constructed within a computer. People from all over the world with an internet connection can interact with each other in these VWs using a personal computer. Through VWs, individuals are able to trade, cooperate and communicate with each other (Bainbridge, 2007). All of this activity occurs in a digital space, making it easy to collect high quality data on many activities that would be unobservable in the real world.

The contribution of this dissertation is threefold. First it uses data from a specific VW to test for price convergence in a quasi-lab setting. Second it analyses the rate of transportation costs in finding empirical support for convergence and PPP. Initially it illustrates the usefulness of VWs in research. The support for price convergence and PPP is found, by developing a greater

understanding of the band of inaction (BoI). The BoI is the range of prices where the cost of trading is greater than the benefit. Within this range of price differentials there is no incentive for arbitrage, and therefore no further convergence pressure. Finally the usefulness of VWs is demonstrated through its use in studying price convergence as well as providing a survey of other social science research using VWs.

The first chapter provides a survey of the VW research conducted throughout the social sciences. It also includes an overview of the demographics of VW users. The first chapter begins with a discussion of the demographic characteristics of the two major varieties of VW. This overview is followed by a quick overview of the current economics research using VWs. The VW research outside of economics is grouped into four main categories, motivations for use, identity, specific behaviors, social networks and culture formation.

If used correctly, the research shows that VWs are a powerful tool for research into human behavior. The extant economics research into VWs supports this finding by demonstrating that agents do act in a manner consistent with previous experimental economics research (Chesney, Chuah, & Hoffmann, 2009; Fiedler, Haruvy, & Li, 2011). The breadth of VW research outside of economics leads to several interesting future lines of research using VWs. The first chapter highlights some of these research possibilities, while surveying the other social science VW literature.

The second chapter of this dissertation tests for price convergence in one of the largest VWs, World of Warcraft (WoW). WoW is a VW with over 12 million users worldwide (Blizzard Entertainment, 2010). Within the game there are two countries, each with an internal goods market as well as one international market. All markets (domestic and foreign) use the same

currency. This format allows for both domestic and international trade. There is a small fee for international trade, which creates a BofI. In such an environment convergence into the BofI is expected to occur.

Several tests for price convergence are used including the Engle-Granger Cointegration test (Engle & Granger, 1987), the exponential smooth transition autoregressive (ESTAR) test (Terasvirta, 1994), the Phillips and Sul Log t Convergence Test (Phillips & Sul, 2007) as well as, the BofI test.

The BofI Test is an expanded version of a new test developed in Morrison and Fontenla (2012). The original BofI test required several simplifying assumptions due to the low quality data in Morrison and Fontenla. The high quality and high frequency data available for this chapter allowed the elimination of many of the simplifying assumptions; creating a much more robust convergence test.

I find price convergence using the Engle-Granger and BofI test, however, not with ESTAR and Phillips/Sul test. This may be due to the fact that both the Engle-Granger and BofI test look for price series that have already converged, while the other two tests look for series that are converging. As the price series within WoW converge very rapidly after a shock, the majority of observations are within the BofI. Because of the sudden convergence, there were too few observations outside of the BofI for the ESTAR and Phillips/Sul test to detect the converging price series.

These results as well as the development of the BofI test both demonstrate the usefulness of VWs in economic research. VWs provide large quantities of high quality and high frequency data, they also lack many of the complicating factors that prevent detection of the price convergence in the

real world, such as the Balassa-Samuelson effect, imperfect capital markets, and hidden transaction costs among others. The high quality data and lack of complicating factors allow the development of the BofI test. Without the data available in a VW accurate measurement of the BofI would be impractical. In the real world accurate measurement of transaction costs is difficult due to the various methods of transportation, complex tariffs and quotas, uncertainty over price change while goods are in transit, hidden trade barriers, and many other trade issues (Hummels, 2007).

The third chapter finds a practical method of measuring the BofI in the real world. This was achieved by altering the ESTAR model to measure the effect of transportation costs on the BofI in the real world. The ESTAR model estimates two separate autoregressive models with a function to control the transition between each model, also called a regime. In the context of price convergence and PPP studies, the autoregressive term is the price differential. The two regimes are inside and outside of the BofI.

This chapter uses price level and real exchange rate data from fourteen developed countries, as well as derived transportation cost data to test the effect of transportation costs on the BofI within a PPP framework. The estimated models show that the BofI narrows as transportation costs decrease. The model can be used to measure the effect of many other trade barriers on the BofI, with the ultimate goal of measuring the BofI in the real world.

Taken together these three chapters demonstrate usefulness of VW research. Chapter one surveys the VW literature within economics and the other social sciences. Chapter two demonstrates the usefulness of VWs by approaching a difficult real world issue within the international trade

literature. Chapter three shows how research in VWs can generate new research questions outside of VWs through development of a new model to help measure the BofI.

Chapter 1 Virtual World Research in the Social Sciences: Opportunities for Economic Research

Abstract: The field of economics, unlike other social sciences, has largely overlooked virtual worlds (VWs) as a setting for research. An established body of literature throughout the social sciences provides background and information for economists. The literature also hints at possible economic research questions. The object of this paper is to survey the extant social science literature on VWs, identify the important categories with brief descriptions of the research, and point out areas where economic theory could improve the literature.

1.1 Introduction

Virtual worlds (VWs) in economic research have been largely overlooked despite their great potential. This chapter outlines this potential, by highlighting similar work from the other social sciences as well as by providing general information regarding VWs and VW users. Communications, psychology, sociology, and anthropology have conducted research in different aspects of VWs. Many of the questions addressed are related to economics. These questions apply to many subdisciplines including behavioral economics, industrial organization, experimental economics, macroeconomics, and others. By examining the history of VW research in the other social sciences, new and useful avenues for economic research can be uncovered.

Many definitions of VW can be found throughout the literature. The most useful defines VWs as a digital environment characterized by the presence of interactivity, physicality, and persistence (Castronova, 2005). Interactivity means the world can be accessed by many users at the same time and that the actions of one user can affect other users. Physicality is defined as generally conforming to the physical laws of Earth, including the scarcity of resources. Persistence means the world continues to exist after a user has left the world and that any changes to the world remain. In other words, the world is not recreated the next time the user connects to the VW. Within VWs people can communicate, trade, and achieve shared goals, while separated by large distances in the real world

Users of a VW interact with this virtual environment through an avatar, which is a digital representation of themselves. Individuals control their avatar and observe the VW through a personal computer or video game console. Because many VWs are video games, the avatar control and VW observation techniques mirror the interface and control scheme common to video games.

The first VWs created were marketed and set up as games where users were given specific goals (kill that monster, explore that cave, etc.) that required group effort to achieve. Through group effort, users received their individual reward. Users were given strengths and weaknesses designed to complement each other within a team framework. This led to the implementation of a variety of strategies and social norms by the users to ensure effective teamwork. An example is the development of dragon kill points (DKP), which reward players for consistent participation towards achieving long-term goals (Reeves, Malone, & O Driscoll, 2008).

As the games developed in sophistication additional forms of interaction were gradually introduced, such as the use of virtual goods as inputs in the production of other goods. Users then were forced to make decisions to produce a desired good, or to purchase the good in whole or in part. Purchases were made using virtual currencies that, like the virtual goods, existed entirely within the VW. This led to development of simple in-world economies, where users specialized in the production of specific goods, selling the goods produced and using the proceeds to advance personal goals. This use of comparative advantage is illustrated in the production and sale of goods such as magic swords. Typically the production of a magic sword requires several skills including, mining, blacksmithing, herbalism, and magic. The maximum number of skills any one avatar can have is two. Therefore a user wishing to own a magic sword must purchase the product of two skills and use knowledge of the remaining two skills to produce the magic sword.

As the number of goods and complexity of production increased the economies became larger and more robust. These worlds experienced many of the issues and problems associated with the real-world economy; including inflation, supply and demand shocks, and financial crises. In effect these video games became virtual copies of the real-world economy.

In addition to the increased complexity of VWs, a split occurred among VWs. Some new VWs continued to follow the video game paradigm. These game-type VWs became known as massively multiplayer online games (MMOG) or massively multiplayer online role playing games (MMORPG). Other VWs began to appeal to a different market. These new VWs focused on the social networking aspects of VWs, becoming digital spaces for socializing rather than a structured space with specific goals to achieve. The most popular example of this type of social VW is Second Life (SL).

By removing the goals imposed by the game developers, new issues regarding motivation and personal interaction were added to this new type of VW. This in turn raised many new research questions, such as, how do individuals respond to bullying in a social VW (Chesney, Coyne, Logan, & Madden, 2009), or how does the presence of a virtual observer affect an individual's behavior in the VW (Hayes, Ulinski, & Hodges, 2010). However, before any of these research questions are addressed, it is important to know the general demographic makeup of the VWs studied. Without demographic information it is impossible to understand how the results of VW research can be applied to the real world.

1.1.1 Demographics

The general demographic makeup of each VW varies widely based on the VW type and target audience. Unlike the stereotype the average VW user is not a 20-year-old male living in his parent's basement. The Entertainment Software Association (2012) reports that average age of game players is 30, with 32% of the population younger than 18. It is estimated the average U.S. household owns at least one dedicated game console, PC, or smartphone capable of playing games. 47% of game players are female, and 11% of all online video games are persistent multiplayer universes, also known as VWs (Entertainment Software Association, 2012). These

demographic statistics show that video game players in general do not fit the stereotype. Among VWs the demographics are less balanced in age and gender, with different genders and age ranges gravitating to different VWs.

The VW population mirrors the racial demographics of the real world, and is typically older than the stereotype. Williams et al. (2008) obtained demographic information for users of the MMOG Everquest 2. Surveying 7,000 users the authors found that the average age was 31.16 with a median of 31. Some 80% of the respondents were male as opposed to the average among all game players of 53%. The Everquest 2 population is 87% White, 3% Hispanic/Latino, 2% percent Black/African American, 3% Asian/Pacific Islander, and 2% Native American, compared to the general population which is 75% White, 13% Hispanic/Latino, 13% Black/African American, 4% Asian/Pacific Islander and 1% Native American. This means that Whites and Native Americans are overrepresented in Everquest 2 while Asians, Latinos, and African Americans are underrepresented. The Williams et al. survey was only for one MMOG, but the results qualitatively match survey results from other MMOGs. A general finding is that MMOG users are primarily white males, with an average age of about 30.

Social VWs by contrast are generally much more balanced in their demographic makeup. The gender split for SL is estimated to be 58% male and 42% female based on a report issued by Linden Labs (Linden Labs, 2008). The average age of SL users is 25-34. Assuming SL is representative of other social VWs, the population of social VWs more closely matches the real world than MMOGs. However, both types of VWs are populated by more than the stereotypical video game player. The users are from all walks of life, all ages and with many more women than is assumed. Thus the demographic makeup of both VW types matches the general

population better than the typical college-age subject pool used in traditional economics experiments.

1.2 VW Research

Research into VW occurred throughout the industry's development and evolution of such worlds. Research began primarily within the fields of computer science and engineering, focusing on the development and improvement of three-dimensional environments and game mechanics. Over time, researchers from education and business began to discover applications within their respective fields.

This research focused on finding applications for improving the achievement of field-specific goals, such as improving language skills or improving sales. In education, this included improving educational attainment and teaching effectiveness. VWs enable the creation of custom-built environments that users can navigate and observe to learn new things. VWs have been used to improve teaching in the fields of linguistics, biology, business, and many others. It has proven so effective that many universities maintain online campuses in SL, including Tulane University, University of Chicago, and 150 other universities.¹

In business, questions were related to improving production processes and marketing techniques. As the VWs became more complex and greater in number, several questions arose regarding the marketing and sale of VWs. Several new business models were developed or adapted from other industries, including subscription fee based models and so called freemium models. In freemium business models, access to the VW is free, but enhancements and additional tools are sold separately (Lyons, Playford, Messinger, Niu, & Stroulia, 2009). All of this research from

¹ http://wiki.secondlife.com/wiki/Second_Life_Education_Directory

computer science, engineering, education, and business was primarily normative in nature, finding the best process for achieving a specific goal.

By examining the nature of phenomena specific to VWs, and shared between the virtual and real worlds social scientists began positive VW research. These lines of research encompassed a wide range of user/player characteristics and interactions, including general demographic makeup of the VWs, the formation of identity, and definition of culture in the VW.

VWs grant a degree of anonymity for users. The users' digital identity and personality need not mirror their real world selves. Researchers began to examine the differences between a person's real world self and their digital self. What implications does the choice of specific physical traits have for the user? How and why are they treated differently because of this choice, and how has this changed with experience in the VW?

Beyond the research into individuals, researchers began to study the formation of cultures within VWs. Questions include: How do VW cultures differ from the real world cultures? What are the implications associated with the large real world distances between members of these virtual societies? What do these virtual cultures tell us about how individuals define culture and society?

The majority of research into VWs has focused on anthropological, sociological, psychological, or communication questions. Little economic research has been conducted into VWs and even less macroeconomic research involving VWs. This chapter highlights the many opportunities for economic research in VWs. After a review of the extant VW economic research, the VW research conducted elsewhere in the social sciences is presented, with a focus on illustrating the opportunities for economic research. The social science research will be grouped into four broad categories: motivations, identity, behavior, and social networks and culture.

1.2.1 What's Been Done in Economics?

The bulk of economic research into VWs can be grouped into two categories. The first category of research attempted to prove the ecological validity of VW experiments (S. A. Atlas, 2008b). Ecological validity refers to the parallelism and saliency of VW studies. This is achieved by using tools from experimental economics, repeating standard economic games within the VW, and comparing the results to traditional experiments. One example of this type of paper is the research conducted by Chesney et al. (2009). The authors repeated the dictator and ultimatum games within the SL. The authors found that participant responses were qualitatively the same as the standard lab games. However, they did receive much better demographic representation than the traditional college subject pool. Recent papers within this category have begun to take the rationality of participants as given and have focused more on transferring typical lab experiments into the VW with similar outcomes (S. Atlas & Putterman, 2011; S. A. Atlas, 2008a; Fiedler & Haruvy, 2009; Fiedler, et al., 2011; Haruvy, 2011; Nicklisch & Salz, 2008).

The second category treats the VW as natural experiments. Early on, these papers described the opportunities for research and calculated preliminary statistics for these worlds. As VWs developed, this category of paper began to explicitly test economic theories using data from VWs. Castronova (2001) is a good example of this descriptive type of research. Castronova describes the MMOG of Everquest (EQ) from a new user's perspective. In the process, he constructed a gross national product (GNP) value for the VW using a simple survey mechanism. Castronova's GNP estimates displayed a wide margin of error, the values were still surprising. His most conservative estimate found the GNP of EQ equal to 2001 Bulgaria. His most liberal estimates found the GNP equal to Russia in the same year. This is a wide range of values; at any estimation, the VW economy was comparable to real world economies. The results are even

more surprising given that the population of EQ was estimated at 400,000 users as the time of the study (Castronova, 2001).

Some of the papers that begin to use VWs as natural experiments include Morrison and Fontenla (2012), which used World of Warcraft (WoW) price data to test price convergence. Additionally, the study by Bloomfield and Cho (2011) examined the rise and fall of an unregulated stock market in Second Life (SL), among many others (R. Bloomfield & Rennekamp, 2009; Chandra & Leenders, 2012; Fiedler, 2011; Guo & Gong, 2011; Malaby, 2006). Several of these studies were included in a special issue of Southern Economics Journal (S. Atlas & Putterman, 2011; R. J. Bloomfield & Cho, 2011; Duffy, 2011; Fiedler, 2011; Harrison, Haruvy, & Rutström, 2011; Haruvy, 2011).

This binary categorization of economic research can be useful, but it is not the only way to categorize the research. Robert Bloomfield (2008) has grouped economic research in VWs into three main categories: immersionist, augmentationist, and experimentalist. The immersionist studies examine questions and issues that occur within the VW and have little or no connection to the outside world. Augmentationist studies view VWs as an augmentation of the real world and study the relationship between virtual and real world institutions. Finally, experimentalist studies view the VW as an extension of the traditional laboratory. These studies create labs within the VWs that mirror the real world labs used in traditional experimental economics. In these labs, the researchers conduct controlled experiments similar to the experiments performed in traditional economics labs. All studies, whether immersionist, augmentationist, or experimentalist, examine the behavior and choices of human beings individually or in groups. Therefore, the results of these studies provide insights that can be used to better understand

human behavior, individually and collectively. This implies the distinction between all categories is based on the tools used, rather than on the research goal.

Under Bloomfield's categorization, the majority of studies that focus on establishing the ecological validity VWs (the first category) would be described as experimentalist studies. However, the studies in the second category that treats VWs as natural experiments would be spread throughout Bloomfield's three categories.

Using Bloomfield's categories, it is apparent that the majority of economic research is focused within the experimentalist section. Few economic studies examine the emergent behavior of economies within VW. This is an area of research that has been well studied outside of economics and is ripe for research within economics. The following sections will describe this research and point out prospective lines of economic research.

1.2.2 Research Outside of Economics

The noneconomics body of work on VWs is best grouped by research question rather than by discipline, because the subjects tend to draw from many disciplines, including anthropology, sociology, and communications. Four general groups can be identified: motivation, identity, behavior, and social networks and culture.

1.2.2.1 Motivation

The studies examining motivation use a variety of established sociological and IT use models to discern why people use VWs, focusing on a variety of social VWs, such as Habbo and SL as well as MMOGs such as WoW. One branch of this research has used formal theoretical models from the field of information systems to examine what influenced the purchase and continued use of a specific VW (Guangying & Haughton, 2008; Mäntymäki & Merikivi, 2010; Mäntymäki

& Salo, 2010; Matti Mäntymäki & Jari Salo, 2011; Verhagen, Feldberg, van den Hooff, Meents, & Merikivi, 2012).

In general, these studies regard motivations behind VW use as either extrinsic or intrinsic.

Extrinsic motivations are defined as coming from outside. A person is motivated to use the VW to achieve some outside goal, or the motivation comes from the results of using a VW rather than the process of using the VW. An example of this extrinsic motivation would be the acquisition of social capital through social VW use (Matti Mäntymäki & Jari Salo, 2011). Intrinsic motivations are motivations that arise from the process, rather than from the results of using a VW. This type of motivation would include deriving enjoyment from in-world achievement or conversation (Matti Mäntymäki & Jari Salo, 2011).

These intrinsic and extrinsic motivations affect a person's intention or desire to purchase a VW, which is viewed as predictor of VW purchase. The goal of much of this motivation research is to find what aspects of the VW influence the intrinsic and/or extrinsic motivations for VW use.

Many aspects found to influence motivation elsewhere in the IT literature also affect motivations in VW use. These aspects include perceived ease of use, perceived enjoyment, perceived network size, escapism, self-efficacy, and availability (Matti Mäntymäki & Jari Salo, 2011).

A field of this research explicitly examines the combined effect of intrinsic and extrinsic motivations on a person's intention to purchase a VW, also referred to as purchase intent. Some papers explicitly model the different motivations as a mediating factor between system characteristics and purchase intent (Guangying & Haughton, 2008; Verhagen, et al., 2012). For example, Verhagen et al. (2012) surveyed 846 Dutch SL users regarding their motivation for and use of SL. The authors identified four prime VW specific characteristics: economic value, perceived ease of use, escapism, and visual attractiveness. The authors then identified one

extrinsic motivation, perceived usefulness, and one intrinsic motivation, entertainment value.

The authors hypothesized that each of the VW specific characteristics would influence both the extrinsic and intrinsic motivation. The two motivations would influence a person's attitude toward use, which in turn would influence a person's probability of using SL.

The authors used the survey results in a structural equation model to test their hypothesis. The structural equation model determined the connections between motivators and use that created the best fit. The authors found that their model explained 39% of the observed variance in attitudes toward SL use. The intrinsic motivation (entertainment value) was the strongest predictor of attitude toward use. The extrinsic motivation (perceived usefulness) influenced both attitude toward use and the intrinsic motivation. Of the four VW-specific characteristics, only visual attractiveness did not influence both motivations. Visual attractiveness influenced only the intrinsic motivation. The intrinsic motivation was influenced most strongly by visual attractiveness and escapism, while economic value had the least influence. The extrinsic motivation was most strongly influenced by escapism; perceived ease of use had the weakest influence.

Mäntymäki (2011) used methods similar to those of Verhagen et al. (2012) to examine the connection between VW characteristics and VW purchase. However, Mäntymäki did not look only at VW purchases. He included the user's intention to continue using the VW. Mäntymäki theorized that a VW purchase did not involve a one-time purchase or one-time use. Using a VW involves an investment and continuous use to realize any enjoyment from the initial purchase. Therefore, anyone thinking about purchasing or joining a VW also must consider the likelihood of continued use. As a result, Mäntymäki included continuous use in his analysis.

In the model, Mäntymäki separated the VW-specific characteristics into four factor groups: motivational factors, social factors, interface factors, and facilitating factors. The motivational factors were perceived enjoyment and perceived usefulness. The social factors included perceived network size and subjective norm. Subjective norm is the need to comply with the desires of others who are important to the user. Perceived ease of use and social presence made up the interface factors, and self-efficacy and availability were facilitating factors. Each factor influenced continuous-use intention, which then influences purchase intention as well as motivational factors and social presence.

Mäntymäki surveyed 3,265 Finnish teenage users of the VW Habbo Hotel, an international VW that is targeted to teenagers. Mäntymäki specifically used this VW because no studies of this type specifically examined teenage populations. Teenagers, while no longer the primary VW demographic, do represent future VW users. Therefore, understanding the factors that influence VW use would be of interest to anyone looking to develop a new VW or to continue the use of an existing VW.

To test the VW use and continued-use theories, Mäntymäki used structural equation modeling to test his theories. The model explained almost 60% of the variance in continued use intention and almost 50% of the variance in purchase intention. Mäntymäki found that all factors were important in a person's decision to continue using the VW, except the factor perceived usefulness, subjective norms, and availability. Perhaps not surprisingly, perceived enjoyment had the strongest influence on continuous use and purchase intention. Social presence had the weakest influence on purchase intention but a stronger influence on continued-use intention. Social presence also had the weakest influence on continuous-use intention.

Two papers listed here examined the connection between VW characteristics and purchase intent. Some studies, such as Verhagen et al. (2012), explicitly modeled the connection between intrinsic and extrinsic motivations and purchase intent (Guangying & Haughton, 2008). Others implicitly combined the two motivations when examining the connection between system characteristics and purchase intent (Fetscherin & Lattemann, 2008; M. Mäntymäki & J. Salo, 2011; Shin, 2009). Several opportunities may be possible to construct hedonic-type models to deconstruct the values associated with each VW feature. A hedonic-type model should be able to determine how much the inclusion of a VW feature will change the willingness to pay for the VW.

Beyond this research into the extrinsic and intrinsic motivations in selecting a specific VW, an additional line of research seeks to explain the motivations driving general VW use. While the work by Mäntymäki and Salo (2011) and Verhagen et al. (2012) examined the motivations in relation to VW-specific characteristics, this other line of research seeks to explain the motivation for VW use in general. One could think of the first line of research as an attempt to explain why a user chooses a specific VW, while the second line of research attempts to explain why a user chooses to use a VW in the first place. The most prolific writers in this area are Yee and Ducheneaut. Together, the two published a significant amount of work in this area from the Palo Alto Research Center (Ducheneaut, Yee, Nickell, & Moore, 2006; Shen, Brdiczka, Ducheneaut, Yee, & Begole, 2012; Yee, 2006a, 2006b; Yee, Ducheneaut, & Nelson, 2012; Yee, Ducheneaut, Shiao, & Nelson, 2012).

The first article is the work by Yee (2006b), where he developed and empirically tested a set of 10 subcomponents of motivation that form three larger motivation components. The components were adapted from previous work on the motivations for play in MUDs (Multi-User Dungeons).

MUDs were an early text-based precursor to online games and VWs. Yee's three major components were achievement, social, and immersion.

The Achievement component is made up of three subcomponents: advancement, mechanics, and competition. Advancement refers to "the desire to gain power, progress rapidly, and accumulate in-game symbols of wealth or status" (Yee, 2006b, 773). Mechanics focuses on the desire to optimize character performance through an understanding of the underlying rules and mechanics of the system. Competition is the desire to compete with others.

The Social component also has three subcomponents: socializing, relationship, and teamwork. Socializing refers to the desire to communicate or chat with other users. Relationship goes beyond socializing in that the desire is not for chatting but for more meaningful long-term relationships with other users. Teamwork is focused on the achievement of goals through group action, or being part of a team.

The Immersion component has four subcomponents: discovery, role-playing, customization, and escapism. Discovery refers to the desire to learn new things and discover knowledge unknown to other players. Role-playing is associated with immersion. It is related to creating a background story for a character and acting as the character. Customization relates to the desire to customize the appearance of an avatar. Finally, escapism refers to the desire to use the VW as a tool to avoid real-life issues and problems.

Yee surveyed 6,675 users from four VWs: Ultima Online, EQ, Dark Age of Camelot, and Star Wars Galaxies. Yee then used an exploratory factor analysis to find the representation of the factors, using 31 of the 40 motivation questions in the survey. The representation connects the 31 survey items to the individual factors and to their factor categories.

In another paper, Yee (2006a) examined five motivational factors and compared them to demographic characteristics. The motivating factors were relationship, manipulation, immersion, escapism, and achievement. The Immersion factor describes users who are motivated by a desire to enter a fantasy world and be someone else for a while. Closely related, but separate from the Immersion factor, is the Escapism factor. The Escapism factor measures a user's desire to escape from the real world and its problems. These users are looking for a chance to temporarily forget about their stress and problems. Users who indicated a high Relationship factor are motivated by a desire to interact with others and form meaningful relationships. The Escapism and Relationship factors together are roughly analogous to the Social component in Yee's other 2006 paper (2006b). The Manipulation factor indicates a desire to objectify the environment and other users for the purpose of personal gain. Finally, the Achievement factor measures a user's desire to achieve goals and overcome challenges within the game environment. The Achievement component from Yee's other 2006 paper roughly corresponds to the Achievement and Manipulation factors.

Yee (2006a) found that male users were more likely motivated by achievement and manipulation, while female users were motivated by the relationship, immersion, and escapism factors. These results confirm previous research into the interpersonal interaction each gender prefers.

Yee, Ducheneaut, and Nelson (2012) revised and streamlined the process for determining the important motivational factors for play. The authors reduced the 39-item survey to a 12-item survey and also standardized survey responses to a 5-point Likert scale. This work used the three main components--socializing, immersion and achievement--from Yee's earlier work (2006b). The authors also eliminated all the subcomponents, thus producing a more parsimonious model.

The authors then validated this method through a survey. In the survey, the authors also were able to show that the model is consistent across real world cultures, by comparing the results from respondents in the United States, Hong Kong, and Taiwan.

Finally, the authors compared the results to available revealed preference data from the online WoW database. WoW database is public store of in-game metrics for every avatar. Included in this database is a list of all achievements completed by each avatar. Achievements are small in-game goals². Hundreds of achievements are categorized into six main categories: quests, exploration, player vs. player (PvP), dungeons/raids, professions, and World Events. These six categories can be roughly matched with the three main user goals. Using regression analysis, the authors were able to show that the self-reported survey data was highly correlated with the revealed preference achievement data, which provides support for the three-factor motivation model.

Yee et al. (2012) expanded on these findings using the same survey results. The authors added basic demographic data to their analysis of the revealed preference data and motivation. The results were consistent with the literature on demographics and motivations for play. Male users and younger users preferred the more “hack and slash” type play such as user versus user combat, while female and older users preferred more noncombat activities such as exploration and crafting. The younger and male motivations roughly correspond to the Achievement categories. The older and female motivations correspond to the Social and Immersion factors.

² Quests achievements are awarded for completing individual relatively quick and easy missions. Exploration achievements are awarded for geographic exploration within the game. PvP achievements are awarded for combat between users. Dungeons/Raids achievements are from completion of difficult team-based challenges. Professions achievements are awarded for completion of noncombat crafting challenges. World Events Achievements are awarded for participation in seasonal story-based events held within the game.

Motivation for using a VW has been studied by looking at the aspects of VW use that appeal to each user. These aspects appear to fall into three major factors: achievement, social, and immersion. These factors were developed and tested through various studies by Yee, Ducheneaut, and other co-authors (Debeauvais, Nardi, Schiano, Ducheneaut, & Yee, 2011; Ducheneaut, et al., 2006; Shen, et al., 2012; Yee, 2006a, 2006b; Yee, Ducheneaut, & Nelson, 2012; Yee, Ducheneaut, Shiao, et al., 2012).

The motivations for selecting specific VWs have been studied using very different tools. These studies have looked at the intrinsic and extrinsic motivations for users. The studies show that users examine perceived enjoyment, ease of use, perceived network size, and subjective norms. Papers by Guangying and Haughton, Mäntymäki and Merikivi, Mäntymäki et al., and Verhagen et al. (2008; 2010; 2010; 2011; 2011; 2012) used established information technology models to find the aspects of VWs that users look for when choosing a VW. In general, perceived enjoyment and ease of use are among the most important factors in choosing a VW. Both groups of studies provide important information to researchers as well as to VW developers.

1.2.2.2 Identity

While many researchers are interested in the motivations for VW use, others are interested in the formation of VW identities. Identity studies examine the formation of a user's identity in a VW and how various aspects of the VW impact that identity.

Yee is also a prolific author in this category. Yee et al. (2009a) studied identity formation when the authors introduced the concept of the Proteus Effect. The Proteus Effect is defined as the effect of an avatar's appearance on individual behavior; for example, a male user may behave differently within a game setting if his avatar is female (Yee, Ducheneaut, Yao, & Nelson, 2011).

This hypothesized difference in behavior is driven by the user's expectations of his avatar, rather than by any difference in the way he is treated by other users.

Yee et al. (2011) published a paper attempting to test this hypothesis, using surveys tied to publicly available data in the VW WoW. The authors examined the differences in the level of support given to others by users based on the gender of their avatar. The authors found that male users whose avatar is female do tend to engage in more stereotypically female activities. These results established a correlation. However, there is no proof of causation for the Proteus Effect.

Yee, Ducheneaut, Nelson, and Likarish (2011) attempted to predict the personality of users based on measurable behavior metrics from the WoW's online database. The authors used the same survey data presented in the motivation study by Yee et al. (2012). The authors used the Big Five Personality Traits model to measure the personality of each respondent, then used achievement data for each respondent's character to predict their personality. The Big Five Model categorizes respondents on five broad personality categories: extraversion, agreeableness, conscientiousness, neuroticism, and openness (Goldberg, 1999). The Big Five Model was developed to create a standardized method of describing personality traits within the psychology literature. The first of the Big Five, extraversion, implies positive emotions toward social interaction and sociability. Agreeableness refers to a person's propensity to cooperate with and show compassion for others. Conscientiousness refers to an individual's self-discipline and sense of duty, which aid in tasks and goal achievement. Neuroticism refers to an individual's emotional stability. Those with a high score in neuroticism are more nervous, easily upset, and more emotionally reactive. The final component, openness, describes the general appreciation for adventure, new experiences, and originality. Together, these five components have been shown to explain most of the variation between individual personalities (Goldberg, 1999). The authors estimated models for

each of the Big Five components, with specific behavior metrics drawn from the WoW online database. All of the results were significant but with medium to small effect sizes.

Shen, Brdiczka, Ducheneaut, Yee, and Begole (2012) expanded on this work by including social network data from the guild membership of each character as well as from textual analysis of avatar and guild names.³ The textual analysis counted the number of positive, negative, and neutral words within avatar and guild names. From the combination of the three data types, a more reliable personality prediction was formed. These predictions can be used for targeted advertising and in tailoring user experiences to the personality type. The authors also point out that this type of research could be adapted to the work environment, allowing employers to form more efficient teams and to identify those employees most likely to release sensitive information or damage company networks. This last application is most likely included due to the funding support from DARPA's Anomaly Detection at Multiple Scales (ADAMS) program.

Bessière, Seay, and Kiesler (2007) used a different approach to examine identity in a VW. Their paper looked at the difference between a user's impression of the difference between themselves and their ideal self, and a user's impression of the difference between their avatar and their ideal self. Using survey results, they found that people with low self-esteem or depression reported a significant difference between themselves and their ideal self. But the difference between their avatars and their ideal selves was not significantly different from the people without depression or low self-esteem. These results indicate that individuals may be using VWs to hide or escape from shameful personal aspects, suggesting that preference structures in a VW may be different from the real world. This difference may be amplified as the difference increases between a person's real world self and their avatar.

Guilds are semipermanent associations of users, supported by the VW's developers. Guilds typically have a set structure with a formal leadership and privileges for membership levels.

Wadley, Gibbs, and Ducheneaut (2009) used qualitative methods to investigate the use patterns of voice communication in SL. The authors found that voice communication generally is rejected when there is a distinct difference between a user's online and offline identity. The users generally wish to preserve the difference. Because voice is not the default form of communication, as in the real world, the use of the voice tool has become much more specialized than in the real world. These results support the theories that users' online identities are bridging the gap between their real and ideal selves (Bessièrè, *et al.*, 2007).

Nagy (2010) outlines a way to approach this issue. He looks at how an individual's VW identity influences their online purchasing behavior. Nagy outlines a framework for research into this area. Although no results are presented, the paper illustrates the importance of examining the connection between purchasing behavior and identity in a VW. The article by Bessièrè, *et al.* (2007) demonstrates that individuals alter the appearance of their avatar to match their ideal self. Research into the Proteus Effect by Yee *et al.* (2009a) indicates that the appearance of an individual's avatar affects their behavior, either directly (the Proteus Effect) or indirectly (through the perceptions of others). These changes in behavior may extend to purchasing behavior. If purchasing behavior is affected, then there are many implications for online sales and marketing. Economic tools and research methods could be especially useful in understanding these connections and finding the appropriate response for online businesses.

The effect of identity on online sales already has received attention within academia. In 2010, Krasonikolakis *et al.* (2010) presented an article at the International Conference on E-Business related to the research question posed by Nagy. The authors were interested primarily in what influences an individual's likelihood of purchasing goods through a VW, as opposed to traditional websites.

Krasonikolakis et al. conducted a survey of Greek SL users to find correlations between the method of online purchasing, demographics, and SL use patterns. Though the sample size included only 104 respondents, the authors were able to find several correlations between purchasing behavior and SL use. The authors found that individuals were less likely to purchase goods through a VW than through a website regardless of their motivations for VW use, use patterns, or demographics. This article was unable to discover any difference between purchasing behavior and VW use patterns, but the article does highlight the potential difficulties that need to be addressed.

Identity in VWs can influence and be influenced by several things. The Proteus Effect described and tested by Yee et al. (2009b) shows that individuals may change their identity to fit the mold in which they are placed. An example of this could be the tendency of male users to perform more stereotypically female behaviors while controlling a female avatar (Yee, Ducheneaut, Yao, et al., 2011). Or individuals may use the freedom of VWs to shape themselves into a form closer to their ideal, as shown by Bessièrè et al. (2007). These changes then may impact any number of other aspects in a user's virtual life, including purchasing behaviors (Krasonikolakis, et al., 2010; Nagy, 2010). The effect of VWs on identity, as well as the effect of VW identity on behavior, is a field with many more research questions to ask. The findings could have tremendous implications on the marketing and development of VWs. While these studies show that identity can affect VW, identity is not the only thing that affects behavior. The following section looks at VW behavior and its causes other than from identity formation.

1.2.2.3 Behavior

Behavior studies attempt to answer questions such as; what explains the behavior observed within VWs, and are they caused by the VWs? These questions look at the behaviors of

individuals within the VW rather than within the behavior of groups or subcultures. Behavior studies have examined a variety of issues, such as the effect of using an avatar on productivity and visual cues on behavior as well as the effectiveness of plans to induce interaction among users.

The first major type of behavior study looks at the mediating effect of using an avatar to represent oneself. Outside of a VW, a person's behavior often is affected by how others perceive the person. Another person's body language and mode of speech can affect an individual's behavior in many ways (Hayes, *et al.*, 2010). However, in a VW, body language and many other cues to a person's thoughts and opinions are absent. How then does that affect an individual's behavior within the VW?

Hayes et al. (2010) attempted to answer this question by looking at how the presence of an avatar and an individual's perspective affect behavior. The aim of the study was to see if interactions through an avatar affected social inhibitions and task performance. The authors performed a psychological experiment in the social VW, SL. Results from the experiment demonstrated that individuals performed simple tasks better when observed, or in front of an audience, and performed complex tasks worse when under the observation of a virtual agent.

The authors conducted several experimental treatments. The first varied the presence and gender of the observer, i.e., no observer, male observer, and female observer. The second treatment varied the perspective used in the VW, first-person versus third-person views. The results weakly supported the idea that observation does affect task performance, though the authors acknowledge that it may have been because of low sample size (Hayes, *et al.*, 2010).

While the results from this paper were not as strong as expected, the authors added an important contribution to the literature. They discussed some of the issues peculiar to VW research that

needed to be overcome. These issues are relevant to social science experiments, including economic experiments, conducted in VWs. One major issue is the believability of the observer in the experiments. Researchers must consider the effect of any avatar observers who are present within a VW study. The presence of an avatar observer could skew the results, as the consensus theory in the study suggests. However, the avatar also may have a different effect than predicted depending on the realistic nature of the observer.

Tied to this examination of the effect of another avatar on an individual's behavior are two studies by Moore et al. (2007; 2007). These studies evaluate the effectiveness of noncombat, visual cues on user behavior in the VW.

In the first study, ethnomethodology and conversation analysis is used in the VW, *City of Heroes* (Moore, Gathman, et al., 2007). This study was driven by the observation that avatars are unable to project many of the nonverbal cues used in everyday life. By incorporating some of these visual cues into the animation of avatars, the authors were able to reduce confusion, or "slippages," in users' behavior. The authors found that the introduction of visual cues improved coordination and reduced slippages. Slippages were defined as delays in conversation, users talking over each other, or general confusion due to mistimed communications. The visual cues included a text alert and animation associated with checking a map, inventory, or text chatting with a third party. These cues in turn provided more information to surrounding users. This allowed the users to better coordinate their actions and reduce confusion.

In the second study (Moore, Ducheneaut, et al., 2007) the authors continue the discussion of slippages in conversation and coordination among users. The authors also suggest several possible solutions and examine the way users inserted cues into their play. The authors also examine the mechanisms users have developed to cope with the lack of cues.

These studies could be used as a motivation for economic research in this area. A concrete metric for team performance and slippage could be developed, based on the completion of high-level tasks. One possible metric could be the completion of large, complex raids or dungeons found in WoW and other MMOGs. Then the metric could be measured with and without visual and vocal cues in a variety of situations. The results could be useful especially in solving coordination problems in large teams across large geographic distances.

Another area of behavior research in VWs is the work by Debeauvais et al. (2012). The authors surveyed WoW users to help determine the factors that influence the decision to engage in real money transactions. Real Money Transactions (RMT) are purchases of virtual currency denominated in real currency. In WoW, this is the purchase of gold, the in-game currency, using U.S. dollars or other real world currency. RMTs are expressly forbidden in WoW's End User License Agreement (EULA). If a user is caught conducting an RMT, their access to WoW is revoked. Despite the stiff penalties RMT still occur, 14% of the respondents reported having engaged in RMT at some point.

The authors used logistic regressions as well as examined correlations between various survey responses and whether the respondent engaged in RMT. Additionally, the authors used a partial correlation matrix in the analysis. Most of the results were as expected; users who had more time commitments outside of the VW were more likely to engage in RMT

This survey is useful for economic analysis. However, the study could be improved by including an income question, as well as more information regarding the timing and quantity of gold purchased. This would allow for examination of external factors, such as price levels and inflation within WoW as well as the general level of production and output within the VW.

Enacting these changes would require the implementation of a new survey. This study would be

of interest to VW developers. Some developers generate revenue from officially sanctioned RMTs; therefore, any information on how to boost RMTs would be important. Other developers, such as WoW, the VW used in this study, forbid RMT and rely on subscription fees to generate revenue. Developers would be interested in these results to help prevent RMT and destabilizing effect RMTs could have on the in-world economy.

The last behavior study is Ducheneaut, Moore, and Nickell (2004), which looked at the VW Star Wars Galaxies (SWG). The authors used ethnographic studies and data collected from public conversations to understand the nature of user interactions. SWG was designed to encourage social interaction among users. Results from the Ducheneaut et al. (2004) study show that this attempt largely failed. Much of the interaction was automated through the use of macros⁴ by the users. Nevertheless, some true interaction occurred, but we do not know if this was the result of users who truly wanted to interact or because there was no efficient way to use macros to interact.

This study highlights the dangers of trying to force behaviors from users, as well as assuming a traditional form of human interaction in this nontraditional environment. The researchers commented repeatedly on the instrumental nature of interactions. When conversations were observed in specific locations in the VW, users would typically interact only as an instrument to complete a transaction.

This is not surprising, given the users' ability to converse with friends and acquaintances through a private "/tell" channel that has no restrictions of distance or location. It seems obvious that

⁴ Macros are short computer scripts designed to repeat a complex action on command or automatically. In the context of this study, the macros were designed to advertise for a service through repeated shouting and then providing said service to users who responded.

users interested in conversation would converse using this channel either with known friends or in an attempt to create new friends.

As a design aspect, the emergence of macros is not surprising. The design forced users to act against their primary motivations for play. The users soon found ways to avoid what was viewed as tedious activities needed to satisfy arbitrary conditions. These conditions and activities prevented many from enjoying the aspects of play that originally attracted the users.

Questions exist about the reliability of studies in an environment such as SWG. If the game is explicitly designed to create a certain behavior, then observation of that behavior may not provide insightful results. This assumes the game design successfully induces the desired behavior. However, if it is unsuccessful, important insights may be gained from failure. But in either case, successful or unsuccessful, the behaviors observed must be acknowledged as a response to specific stimuli and not as a natural outgrowth or evolution of users in the VW.

Despite attempts to force behavior in users, SWG does contribute important insights to VW studies. Within VWs, the important unit of observation is the user, who often is not directly correlated with the character or avatar. Because users often can have multiple characters tied to one account, it is perhaps more useful to view users as individual firms, with each character or avatar representing a user or part of the firm.

The goal of these “firms” is to maximize the users’ utility. Maximizing utility often requires the use of several characters. An example: Imagine a user driven by achievement goals. The user increases utility by achieving set combat-based goals. Early on, this can be accomplished easily with only one character, specializing in combat. However, as the user progresses, the requirements for each goal become more onerous, perhaps requiring greater monetary resources or the production of specific goods. At this point, the user may create a second character who has

a greater comparative advantage in wealth creation. Now the user has two characters, one who is engaged in the primary activity of achievement and the second in a support role, providing the necessary resources for the primary character.

In this scenario, if we examine the behavior of any one character, we will not gain a full understanding of the user. Each character reveals only a part of the whole. When examining the behavior or makeup of a user, the appropriate unit of measurement is the account or firm housing all the separate characters. Without cooperation from the users or the game's developers, this level of observation is difficult to achieve. However, studies of the aggregate behavior of an economy are not necessarily restricted by this problem.

In relation to the problems observed in this study, it may be that many of the characters using macros are controlled by the achieving characters and as such have little interest in personal interaction. Because of macros, this achieving user does not even need to take direct control of this character. The achiever simply may set a macro as a support character and then accomplish some other task outside of the VW.

One last lesson hinted at by Ducheneaut et al. (2004), and present in other studies, is that just because the game environment imposes a cultural importance or interpretation of an item does not mean that that interpretation always will transfer to the user. An example within SWG: Users gradually ally themselves with one of the two factions present in the game The Rebel Alliance or The Empire. Presumably, the environment and lore of the VW and the Star Wars franchise assign specific moral and ethical values to each side. However, it may be that a user's alignment to either side has less to do with the superficial cultural characteristics and more to do with tangible, practical rewards associated with each side. Users may not be influenced by the cultural setting and view the game in pure mechanical terms.

From these behavior studies, specifically examining user behavior in VWs, it is apparent that individuals may behave differently in the real world. However, the magnitude and direction of behavior change is uncertain. It is apparent that the inability to display visual cues can impair coordination and conversation among users, as demonstrated by Moore et al. (2007; 2007). Users also have been observed to treat computer-generated avatars as real individuals (Hayes, *et al.*, 2010). Yee et al. (2011) demonstrated that the gender of an avatar may cause users to engage in behaviors stereotypically associated with that gender. It also has been shown that the overriding motivations for play also can impact an individual's decision to engage in RMT. However, these results are tempered by the results from Ducheneaut and Moore (2004), showing that attempts to induce user behaviors can fail and lead to unintended consequences, such as the creation and use of macros. Finally, Ducheneaut and Moore (2004) also highlight the fact that several avatars can be used by a single user. If the researchers do not account for this fact, then the behavior of individual avatars may be inexplicable.

1.2.2.4 Social Networks and Culture

Several studies investigate group dynamics and social networks within VWs. These studies have looked at groups in MMOGs, such as WoW, as well as in social VWs such as SL.

Groups in MMOGs generally take on two structured forms. First is the more permanent group structure called guilds. These are semipermanent associations of users, supported by the VW's developers. Guilds typically have a set structure with formal leadership and privileges for various membership levels. The second type is known as a pickup group. These are temporary groupings of two to five users and typically are formed to achieve a set goal and then dissolved. Both the temporary pickup group and the more permanent guild have unique dynamics and can reveal much about the formation of social networks and culture within a VW. The structured nature of

groups in social networks has driven much of the research using game-type VWs. These studies generally have relied on automated data collection paired with surveys. They have examined the determinates of guild survival, the likelihood of a positive or negative response conditional on previous interactions, and the homogeneity of race and class within a guild.

Groups in the social-type VWs are less structured and typically have no formal support from the VW's developers. The lack of a formal group structure has influenced the type of studies conducted in social VWs. These studies tend to look at how groups work together to solve complex problems or deal with disruptive behavior of individuals. The researchers rely on formal experiments and interviews.

First among the MMOG studies is the paper by Williams et al. (2006) that examined the social structures of WoW guilds primarily through ethnographic interviews. The authors reported that the motivations for guild membership, the degree of socialization, and the importance of the social bonds within guilds varied with guild size. A general finding was that small- to medium-sized guilds were driven primarily by social motivations, while the larger guilds were driven by achievement motivations (completion of endgame content or PvP). In general, users rated their social connections within guilds as less important than real life connections, though a small minority (~5%) reported that the relationships formed within the game were stronger than real life relationships. The authors also reported that as the goals of guilds shifted toward achievement, the organization and social structure of the guild shifted to a much more structured and organized form. The authors noted increased use of websites and third-party voice over internet protocol (VOIP) applications to organize group activities. These general findings from the ethnographic research highlight many avenues for additional empirical research related to guild performance and structure. One major line of research would be to expand on the studies of

centrality within a guild, or social connectedness, by examining the effect of voice on connectedness.

Ducheneaut and Yee (2008) used various metrics from WoW to examine the social networks of users. They examined how often and with whom users were forming pickup groups as well as more permanent associations through guilds. The authors also examined the connection between their social networks and level advancement. Using available metrics, the authors found that social networks in WoW were more sparse than expected and that users tended to play in a form of “collective solitude,” meaning that users, while not forming varied and strong social networks, still enjoyed being around others even if they apparently were not creating or strengthening connections in their social network. The article addressed many connections from the data. However, it appears the authors ignored one complicating factor when presenting their conclusions. When users are within the same guild, they share a channel of communication that is constant regardless of the location or group status of the users. When using this channel, many users could be forming connections through conversation while still never physically forming a pickup group. If this is a significant occurrence, then many of the conclusions drawn in this paper are faulty. Additional research is necessary to account for this shared channel of communication.

As an extension to this line of research, the authors presented a conference paper in 2007 (Ducheneaut, Yee, Nickell, & Moore, 2007) that specifically looked at the social network aspects of a guild that affected the survival of the guild.

The authors found the major factors in guild survival to be class balance, guild size, level standard deviation, maximum subgraph size, time in instances, and density. At the higher levels of WoW, proper coordination of users and combining skills and abilities become much more

important. This in turn requires a relatively equal distribution of classes within the guild. Guild size or the number of guild members is important because success in the endgame content requires as many as 40 participants. Therefore, the greater the guild size, the greater the probability enough guild members will be available to participate in the endgame content. Level standard deviation is a measure of the distribution of guild members across levels of advancement (60 levels at the time of publication). If distribution is relatively even, then there are greater opportunities for guild members to form groups and assist each other in level advancement. This then improves the incentives for remaining in the guild and therefore guild survival. Maximum subgraph size refers to the size of subgroups within the guild. Smaller subgroups are hypothesized to avoid coordination issues. Finally, density measures the level of connectedness among the members. More connections between guild members mean greater peer pressure to remain in the guild and contribute to guild success (measured in endgame content completion, level advancement, etc.). The results from this paper provide insight not only into the social dynamics of WoW, but also may help in the understanding of organizations in general. The authors note that the form of guilds mirrors the organization trends observed in business. Therefore, insights into effective organization and leadership could be applied to the world of business leadership and organization.

The authors point out that large, successful guilds tend to reduce many transaction costs for guild members. A natural extension would be to investigate the effects of guild size/success on the performance of the in-world economy. Are there economies or diseconomies of scale for guilds with respect to efficient markets? Do large, stable, and successful guilds (i.e., firms/corporations) lead to greater market efficiency? Perhaps large guilds could lead to decreases in efficiency, as guild members begin trading only within the guild?

As a final extension to their work, Ducheneaut, Yee, and Bellotti, (2010) described the potential for combining qualitative ethnography and quantitative methods in studying VWs. The authors suggested the wealth and ease of data collection in VWs could be used to gain insight into the generalizability of ethnographic results. In the process, the authors described the quantitative tool they developed in their earlier work, the Social Dashboard. This tool automatically collects real time data on activities within a VW, for use in collaboration with ethnographies. The data collected through the Social Dashboard also could be used in independent quantitative studies. It may become a valuable tool in the realm of quantitative rather than qualitative studies.

Johnson et al. (2009) examine the common group dynamics between WoW guilds and street gangs. The authors develop a simple group formation model using techniques from physics. The authors compared group formation in both WoW guilds and Los Angeles street gangs. The results show that group formation follows a similar pattern in both environments. Beyond the results, this paper is also noteworthy in that it represents a merger between physics methodologies and social science questions. This interdisciplinary study highlights the usefulness of looking for models and research methods beyond typical discipline boundaries.

Kobayashi (2010) examined the effect of online communities on offline behavior, focusing on the effect of heterogeneity (of online communities) on social tolerance online as well as offline. Social tolerance within this study was restricted to interpersonal social tolerance, or the ability to tolerate differing opinions in other people.

The online communities studied were guilds within the MMOG Lineage. Lineage is a Korean and Japanese MMOG with a similar structure to WoW. Kobayashi (2010) found that social tolerance online and offline increased as the degree of perceived heterogeneity increased within a

user's guild. This result demonstrates that online cultures can affect the offline world and that the effect can be positive. In this instance, it increased tolerance of others opinions.

Golub (2010) provided a brief ethnographic outline of a "medium-core" raiding guild within WoW, presenting the ethnography to demonstrate that the cultures based on the content of VWs is not restricted to the VW. He argued that researchers must look beyond the confines of the VW when studying a VW's culture. This is driven by the fact that a VW's activities often spill into other media and aspects of a user's life. The author specifically refutes the idea of culture existing within a single place but rather believes it exists across places. The authors show that users communicated across several mediums including internet forums, VOIP services, and others, as well as in-world communication channels. This article reveals few avenues for future research; it does describe the level of coordination and teamwork necessary for raiding, and the expansion of VW activities outside of the VW, in addition to addressing the primary argument of the paper that culture is something that exists in the connections between individuals rather than in a specific place. This is made more apparent with the creation of VWs where physical space, even within a VW, is less important to maintenance of social relations and development of culture. In the past, the primary way for individuals to form connections was face-to-face connections. This was a restriction imposed by technology. However, as technology has improved, the need for face-to-face interaction in the creation of social networks and culture has become less important. Seeing that culture can form in the absence of shared physical spaces demonstrates that it is the interconnection of complex social networks that creates culture, not the physical venue needed in the past for social connection.

Lortie and Guitton (2011) looked at the degree of in-world racial and gender homogeneity within WoW guilds. Using WoW's online database, the authors found that increased homogeneity

appeared only among human-like avatars and not among other groups. This occurs despite a programmed need for heterogeneity in classes, which is determined somewhat by racial choice. The data for this paper was collected from the publically available database of WoW characters. It highlights some of the useful applications of this robust data set. However, the paper failed to take into account the issue of selection bias. Users are not randomly assigned a race and gender when entering the game. The users select the race, gender, and many other characteristics when they create the character. This could mean that while there is increased homogeneity among human-like guilds, it could be driven by the preferences peculiar to users that select a human avatar, rather than by any causal relationship with the physical characteristics of the characters. If a study could account for any selection bias, then this could be an important contribution to the literature with applications to economics.

The final social network study to use MMOGs is the paper by Thurner et al. (2012). In this paper, the authors examined a user's probability of engaging in a positive or negative action, conditional on the receiving either a positive or negative action. The study used data from the VW Pardus, a MMOG. The authors examined the behavior of 34,055 users over 1,238 consecutive days. The game allowed users to engage in eight actions at any time. Four actions are positive actions: communication, trade, setting a friendship link, and removing an enemy link. The other four are negative actions: attack, placing a bounty on another user, removing a friendship link, and setting an enemy link. The authors then examined the actions of users with respect to the previous action or received action. The authors found that users were more likely to engage in a positive action if their previous action was positive or if they had received a positive action. However, the results were much stronger for negative actions. Users were much more likely to engage in a negative action if their previous action was negative or if they had

received a negative action. This would indicate a strong persistence in user behavior, based on prior behavior. The article did not speculate on why this may be or what the consequences of this finding may be. The study does, however, provide another example of the information that can be gleaned from VW studies.

The social VW studies begin with a paper by Chesney, Coyne, Logan, and Madden (2009). The authors examined the phenomenon of “griefing,” a form of bullying or harassment within the VW SL. The authors described the patterns of griefing and how it relates to other forms of harassment and bullying. The authors also discuss the potential consequences for victims, potential motivations for perpetrators, and where responsibility for prevention should lie from the perspective of users. The information presented was collected from observation and focus groups. Griefing is viewed as equal to bullying in other environments, with similar motivations and consequences, and should be addressed in a similar fashion.

Wadley and Ducheneaut (2009) also looked at collaboration within SL. The authors found that participants regularly took advantage of their ability to separate perspective from that of their avatars when constructing items within SL. The participants quickly found ways for referencing items and locations in a way that made sense to all parties, even though participants were unable to discern the perspective of their collaborators. Earlier research suggested users needed many more tools and cues anchored to the avatar, such as avatar gaze, pointing, etc., in order to coordinate actions. However, the participants in this study eschewed such tools when they were permitted to decouple their perspective from that of their avatar. In fact, participants suggested many additional cues that were entirely separate from the avatar. The study did not compare the effectiveness of collaboration using an avatar-centric perspective, versus the decoupled perspective from SL, but it does seem apparent that the entire approach to collaboration is

fundamentally different. Therefore, this difference requires different tools and collaboration styles as well. It is up to future research to determine when one approach is more appropriate than others and under what circumstances.

Social networks formed within VWs have tremendous impacts on a user's behavior, both in the VW and in the outside world. Whether the social networks are formed in a MMOG or social VW, both have the capacity to alter a user's life as well as to provide insights into team building in social networks in all settings. Johnson et al. (2009) showed this connection between online and real world groups when they compared the group formation patterns of WoW guilds and Los Angeles street gangs. If street gangs form in similar fashions, it is not unreasonable to assume other real world groups and social networks mirror VW group formation as well.

Using the Social Dashboard, Ducheneaut, Yee, Nickell, and Moore (2007) were able to show the determinants of guild survival that could provide useful insights into effective team organization in the workplace. Kobayashi (2010) showed how our interactions online can change our interactions with people in our everyday life. His study showed that diverse populations encountered online can make individuals more accepting of diversity elsewhere in their lives. The works by Thurner et al. (2012) and Chesney et al. (2009) hint at ways to combat and prevent bullying in VWs and in the outside world.

All of these studies demonstrate ways in which VW research can produce useful results, both for VWs and the outside world. There are ample opportunities for economists to be part of this research, whether through applying economics techniques to the research questions already asked or through research into related economics research questions.

1.3 Conclusion

VWs are a digital environment where people from all over the world interact with each other on a daily basis. VW users represent a wide range of demographic characteristics and preferences. The studies presented in this chapter show that users are investing significant amounts of time and effort into their VW lives. Users both affect and are affected by the VW. While some may use VWs as an opportunity to reportray themselves, all users behave as rationally in the VW as they would in the real world.

It was shown that the motivations for VW use in general varies across both age and gender and can be driven by a variety of factors, including achievement, immersion, escapism, and others (Ducheneaut, et al., 2006; Yee, 2006a, 2006b). Studies also showed that the decision for selecting a given VW includes many personal and VW-specific factors. These factors could affect a person's utility from using a VW as well as any outside gains from VW use (Guangying & Houghton, 2008; Mäntymäki & Merikivi, 2010; Mäntymäki & Salo, 2010; M. Mäntymäki & J. Salo, 2011; Verhagen, et al., 2012).

After a user has chosen to use a VW, the identity the user presents in a VW is partially determined by their self-view (Bessièrè, *et al.*, 2007) and by how much of their real world selves they must reveal (Wadley, *et al.*, 2009). The user's identity also may be affected by the characteristics of their avatar, known as the Proteus Effect (Yee, et al., 2009a; Yee, Ducheneaut, Nelson, et al., 2011; Yee, Ducheneaut, Yao, et al., 2011). It also was shown that users' personality can be inferred from broad measures of their VW use (Shen, et al., 2012; Yee, Ducheneaut, Nelson, et al., 2011; Yee, Ducheneaut, Shiao, et al., 2012).

The behavior section showed that users sometimes behaved as if they were the avatar. This was shown in the study by Hayes et al. (2010) where users performed worse on difficult tasks in a

VW when they were observed, just as they would in the real world. While users sometimes may behave as if they were the avatar, they still were stymied by the lack of nonverbal cues we rely on in the real world. Moore et al. (2007; 2007), in their two studies, showed how the absence of nonverbal cues can impede the flow of conversation and productivity. By comparing their observations from VWs to the outside world, they were able to draw many conclusions regarding communication in the outside world. Similar conclusions could be drawn in the field of economics when examining the assessment of risk and trust, when dealing with others.

These studies highlighted many things that could be done to improve the construction of VWs. However, the study by Ducheneaut et al. (2004) showed in the VW SWG how attempting to improve a VW can have unintended consequences. The developers of SWG tried to build in the need to socialize, but this only caused problems as users found ways around the forced social interaction.

The studies have shown that individuals form VW groups in the same way they form groups in the real world (Johnson, *et al.*, 2009). The groups and social networks formed within VWs often expand beyond the boundaries of the VW as well as affect the way individuals see others in the real world (Golub, 2010; Kobayashi, 2010). These results demonstrated the need for understanding the structure of groups within VWs as well as the determinants of a persistent group (Ducheneaut & Yee, 2008; Ducheneaut, *et al.*, 2010; Ducheneaut, *et al.*, 2007; Lortie & Guitton, 2011; D. Williams, *et al.*, 2006).

Beyond formal group and social network creation, the cultures that emerge within VWs also have provided insight into issues in the real world, such as how to combat bullying. Chesney et al. (2009) demonstrated the similarities between grieving in the VW and bullying in the real world. Looking at the coping and prevention strategies adopted in the VW may provide ways to

address bullying in the real world. Finally, methods developed for communication in a VW can provide insight into solving difficulties in communication through other mediums, such as email (Wadley & Ducheneaut, 2009).

The users respond to incentives and find ways to achieve their goals despite technical problems and interference from others. The unique way they solve problems within the VW can help researchers understand the dynamics of many things taken for granted in the real world. VWs represent a new and useful field of research for economics. Whether the studies are immersionist, experimentalist, or augmentationist, the results can provide useful insights in both VWs and the outside world. The research conducted in the other social sciences demonstrates many useful avenues for research, including studies into the value of specific VW features, the effect of guild size on the performance of the VW economy the effect of the VW economy on RMT?

Chapter 2 Price Convergence after Cataclysm

Abstract: This paper tests for the existence of price convergence after a major shock in the largest online game, World of Warcraft. We collect data from eight worlds, or copies of the virtual world. World of Warcraft provides a unique setting for the study of price convergence. This virtual world is a controlled setting without nontariff trade barriers, transportation costs, difference in productivity growth, and other barriers to price convergence found in the real world. The period of study covers a major expansion to the game, called Cataclysm. In this expansion, several new goods are introduced, and new uses are found for many pre-existing goods. We use the Engle–Granger Cointegration test (1987), ESTAR (Terasvirta, 1994), Phillips and Sul Log t convergence test (2007) and the band of inaction test we developed in Morrison and Fontenla (2012). We find convergence in all eight worlds using the band of inaction test and Engle–Granger test. We fail to find convergence using the ESTAR test and Phillips and Sul tests.

2.1 Introduction

The Law of One Price (LOP) is the simple and important idea that prices of goods across countries should be similar, or else arbitrage opportunities would generate pressure toward price convergence. Directly related to LOP; price convergence states that price levels across countries should be the same. Despite the intuitive nature of these theories, their empirical validity has remained a contested matter. A large, ongoing literature has been gathering high-quality long-run data and has taken advantage of more powerful econometric tools, especially nonlinear unit root tests (de Carvalho & Júlio, 2012; Alan M Taylor, 2002). Still, no agreement seems to have been reached, even when studies have been run on essentially the same data and models (Alan M. Taylor, 2001).

Our paper takes a different approach, in that we test for price convergence in a natural, macroeconomic experiment using the largest online game, World of Warcraft (WoW). This virtual world is a two-nation, controlled setting void of nontariff trade barriers, transportation costs, difference in productivity growth, and other barriers to price convergence found in the real world. Further, our data covers a major shock to the game, called Cataclysm. In this expansion, several new goods are introduced, and new uses are found for many pre-existing goods. We use the Engle-Granger cointegration test, exponential smooth transition autoregressive (ESTAR) model, Phillips and Sul log t convergence test (2007), and the band of inaction (BofI) test we developed in Morrison and Fontenla (2012). Our test is unique in its ability to empirically model the BofI. We find convergence in all eight worlds using the BofI test, as well as convergence in most worlds using the Engle-Granger test. The ESTAR and the Phillips and Sul tests failed to find convergence. The majority of observations are within the BofI, which would prevent the

Phillips and Sul and ESTAR test from finding the mean reverting behavior that indicates convergence.

Price convergence and the closely related purchasing power parity (PPP) have been studied for decades. Originally, tests used linear cointegration analysis to find evidence of price convergence and PPP. Tests such as the augmented Dickey-Fuller test suffered from low power and provided inconclusive results. To overcome these issues, many studies used panel Dickey-Fuller tests and large data sets (Funke & Koske, 2008). This increased the power of the tests and provided more conclusive results. Taylor (2001) pointed out that many of these studies falsely assumed a linear convergence path and lacked the necessary precision to accurately capture price convergence. Taylor also argued that the monthly, quarterly, or even annual data used in the studies would be unable to capture convergence if it occurred on a daily or weekly time scale.

The realities of data collection in the real world prevented many researchers from addressing the precision argument, but several researchers began developing nonlinear tests for price convergence. These tests began to incorporate the BofI concept in their theoretical framework. Tests using the transition auto-regressive (TAR) and the Smooth-TAR (STAR) models were much more successful in finding price convergence (Baum, Barkoulas, & Caglayan, 2001). These tests inherently assumed a BofI and incorporated the change in series behavior when entering the band. However, these models assume that all price series display a similar convergence path.

Phillips and Sul (2007) developed their log t convergence test, allowing for nonlinear as well as heterogeneous convergence paths. The Phillips and Sul test looks for converging behavior and is unable to detect convergence in series that already have entered the BofI.

Morrison and Fontenla (2012) developed the framework for a new convergence test, based on accurately measuring the BofI. The new test compared the benefits and costs of trade given the current price levels in two countries. If the costs consistently outweigh the benefits of trade, then the series must lie within the BofI. Unfortunately, because of a paucity of data, the authors were forced to make many simplifying assumptions to find results.

The next section will discuss the extant economic research using virtual worlds. The third section discusses price convergence in greater detail with references to previous work. The fourth section will discuss the environment and agents of WoW, and that will be followed by a discussion of the data and data collection. Then the price convergence tests and their results will be presented. First, we will present the Engle-Granger cointegration test, followed by the ESTAR test, then the Phillip and Sul test followed by our BofI test. The conclusion will follow the presentation of the BofI test and results.

2.1.1 Virtual Worlds and Economic Research

Data for this study come from the virtual world WoW. Virtual worlds are computer generated reflections of the real world. Players experience the virtual world through an avatar, or virtual representation of themselves. The players are able to interact, affect, and be affected by other players and computer generated agents. These effects are persistent; they affect future as well as present game play. The virtual world follows the same general physical rules as the real world, including the presence of scarcity. Players must make trade-offs in time and virtual currency to achieve goals. Players are required to purchase or construct goods and services to achieve their long-term goals within the game.

There exists some skepticism that players in a video game can or will act rationally. Players do not interact in the virtual world like they do in the real world. Players can experience the world only through a computer screen; nothing is tangible. The virtual world is a fantastical environment where the impossible is possible. However, in a virtual world, players are forced to make trade-offs in the face of scarce resources, just like in the real world. Presumably, players maximize their utility from the game. In the process of maximizing utility, the players must choose how to allocate their scarce resources in the most efficient way possible, much like in the real world.

Ultimately the scarcity of resources is controlled by the game developers. The game developers work hard to ensure there is a scarcity of resources to provide challenge to the game. In the face of this scarcity, players must act rationally to maximize their utility. Persistent irrationality would result in disutility until the player becomes discouraged and self-selects themselves out of the game.

Demographic information for WoW users is not readily available. However, demographic information for a similar VW Everquest 2 has been collected by Williams et al. (2008). The demographics in WoW are believed to be similar to that of Everquest 2. Williams et al. (2008) found that the median age of users was 31 with the average age of 31.16, and 80% of users are male. The population was 87% White, 3.34% Hispanic/Latino, 1.55% Black/African American, 2.68% Asian/Pacific Islander, and 1.74% Native American. The VW population does not perfectly mirror the real world demographics but is close.

This is not the first time online games have been used for economic research. A previous study by Morrison and Fontenla (2012) used less precise data collected from WoW to examine price

convergence and found that prices had converged using an early version of the BofI test used here and the Phillips and Sul log t convergence test.

Numerous other studies have been conducted in the virtual world, Second Life (SL). One study by Chesney et al (2009) repeated several standard economic games in the virtual world and then compared the results to traditional labs. Chesney et al. found the results were unchanged from the standard lab results, though they did get much better demographic representation. Fiedler and Haruvy (2009) conducted a similar study, with several trust games in SL, then conducted the same games in a traditional lab setting. They found that interaction through a virtual world increases the amount of money sent relative to laboratory results. However, participants recruited through the virtual world gave and returned less than control groups.

Perhaps the first study using virtual worlds for economic research was Castronova's (2001) paper on the virtual world Everquest. Along with describing the general nature of virtual worlds, he generated several macroeconomic variables. He estimated that the GNP of Everquest was on par with several Eastern European nations. Castronova since has published several other studies for economic research using virtual worlds (2005, 2008).

Haruvy (2011), in a special issue of the Southern Economic Journal, presents a summary of several economic studies using virtual worlds. In this special issue, several traditional economics experiments as well as one natural experiment within virtual worlds are conducted. The first article in the special issue by Duffy (2011) uses a personal anecdote to illustrate some of the practical issues that must be addressed and overcome when conducting online experiments--specifically regarding identity and honesty when conducting traditional lab-type experiments.

The next article by Atlas and Putterman (2011) in the special issue conducts a trust game in the virtual world SL. The authors find high levels of trust and reciprocity. Additionally the authors discuss many of the practical issues encountered while conducting experiments in SL, including communication between participants, unbalanced randomization between treatments, individuals participating multiple times, among many others. The issues raised and their solutions to these problems make this article a useful read for anyone replicating traditional lab-based experiments.

Also in the special issue, Fiedler (2011) looks at the effect of experience on subject behavior in an asset market experiment. Fiedler conducted the experiment using the software zTree. She recruited participants through Second Life and then had them remotely log into a university computer to conduct the experiment. This article clearly demonstrated how traditional experiment software and methods can be adapted for use online. This method preserves the benefits of traditional experimental methods while still using the benefits associated with virtual world experiments.

Bloomfield and Cho (2011) use an asset market within Second Life to conduct a natural experiment. They examine the firm's performance, trading volume, cash flow, and market performance in an unregulated asset market. The authors document financial meltdowns, fraudulent activities, and the resulting regulations. This study illustrates the possibilities of using virtual world phenomenon for natural experiments.

The final article in the special issue discusses the issues and opportunities available from the virtual reality nature of virtual worlds. This paper by Harrison, Haruvy, and Ruström (2011) points out that in virtual worlds a new environment is created. This new environment can be designed to set up associations in the minds of participants that would not be possible in a real

world lab setting. These associations, if used correctly, can be used to achieve greater saliency in economic experiments.

2.1.2 Price Convergence

Price convergence is an extension of LOP. LOP describes the effects of arbitrage. By purchasing goods in one market and selling the goods for a profit in a separate market the prices in both markets are driven to equality. LOP describes this process for individual prices, while price convergence applies to general price levels. LOP is a basic assumption of many trade models and has been a part of economics since the 16th century. However, empirical confirmation of LOP or its extensions have provided conflicting results. Until recently, the consensus was in favor of price convergence in the long run but with much short-run variation. Previous studies had found three-year to five-year half-lives for convergence (Froot & Rogoff., 1996; Funke & Koske, 2008; Wolszczak-Derlacz, 2008). These studies used simple linear tests for convergence and may have overlooked many of the complicating factors associated with price convergence. Recent studies using nonlinear tests have found half-lives significantly less than three years (Chong, Jordà, & Taylor, 2012; Fan & Wei, 2006; Kim & Kim, 2011; Nejib, 2008; Alan M. Taylor, 2001; Alan M Taylor, 2002).

Our study adds to the literature by examining price convergence in a new setting using two relatively new tests in addition to more-established tests. One benefit from our setting is the quality of data. We are able to use much finer data than many previous studies. This allows us to avoid temporal aggregation. Taylor (2001) describes temporal aggregation as one of the major pitfalls in LOP studies. In examining the potential biases introduced by temporal aggregation, Taylor references the seminal work by Working (1960). Working demonstrates that using first differences in temporally aggregated data can lead to underestimated variances, for even minor

temporal aggregations (Working, 1960). Taylor (2001) expands on this work by demonstrating that in AR(1) processes, Engle-Granger tests, as well as half-life calculations used in PPP and price convergence studies, the temporal aggregation leads to upwardly biased estimations (Alan M. Taylor, 2001). Our data from WoW is temporally aggregated to the daily level, as opposed to the monthly, quarterly, or annual aggregation used elsewhere (de Carvalho & Júlio, 2012). This level of aggregation avoids any of the biases discussed by Working (1960) and Taylor (2001).

In addition to addressing the temporal aggregation issues raised by Taylor (2001), we explicitly model the BofI. The second major pitfall identified by Taylor is a failure to account for the BofI (2001). The BofI is the range of price differentials where the benefit of arbitrage is less than the cost. While inside the BofI, each price is governed by separate market and macroeconomic forces, implying no convergence. Traditionally, the price differential is modeled as a random walk, bounded by the BofI edges. Outside of BofI, arbitrage profits are available, and therefore prices should converge. The BofI implies that prices will converge only to equality out of pure chance. Instead, prices will converge into a specific range of price differentials. Once inside the BofI, prices have converged as much as possible. Therefore, entering the BofI must be considered equilibrium, rather than price equality (Alan M. Taylor, 2001).

Assuming the price differential is x_t , then the BofI can be modeled as a two-regime threshold autoregression

$$x_t = \begin{cases} +c + \gamma(x_{t-1} - c) + e_t & \text{if } x_{t-1} > c, \\ x_{t-1} + e_t & \text{if } c \geq x_{t-1} \geq -c, \\ -c + \gamma(x_{t-1} + c) + e_t & \text{if } -c > x_{t-1}, \end{cases}$$

Where $e_t = N(0, \sigma^2)$, γ is the autoregressive coefficient for deviations from the edge of the band, and c is the amplitude of the edge of the band. This representation of the BofI with a

discrete edge would be estimated with a TAR model. The TAR model is most appropriate for LOP studies using individual goods. However, when examining the LOP in context of price levels, as in price convergence or PPP studies, then a STAR model is more appropriate. With price levels, each good in the basket can enter the good specific BofI at separate price differentials. As the price level differentials decrease in size, more and more goods enter their BofI, and therefore, the price level is closer to entering the price level specific BofI.

This nonlinear approach into the BofI of price levels prevents the use of traditional tests. Traditional linear tests, such as the Engel-Granger cointegration test, among others, assume a linear convergence path. This linear assumption means these tests are more likely to reject convergence as well as to produce biased coefficient and half-life estimations. Taylor (2001) and de Carvalho and Júlio (2012) provide excellent summaries of the causes and nature of these biases.

Many other price convergence tests account for the BofI, including the Phillips and Sul test and the ESTAR model tests. However, we find that in some circumstances these tests are inappropriate. Tests such as the Phillips and Sul test look for decreases in cross-sectional variance over time to indicate convergence. However, if the majority of observations are within the BofI or if observation begins before a shock that drives price differentials outside the BofI, then Phillips and Sul will not find decreasing cross-sectional variance. ESTAR-type tests model inside and outside the BofI as two separate regimes. However, if there are not enough observations in one regime, then the ESTAR model would be unable to estimate the model correctly, increasing the probability of Type I or Type II errors.

The cointegration-type of tests look to see if prices mirror each other. The Phillips and Sul and ESTAR tests look for converging behavior. There exists another test that instead looks to see if prices are within the BofI. The BofI test, developed in our earlier work (Morrison & Fontenla, 2012) overcomes many of the pitfalls described by Taylor (2001). This test examines the price differentials relative to a known BofI. If the price differentials lie within the BofI, then prices are at equilibrium and have converged as much as possible. If prices are outside the band, then it is necessary to employ other tests to see if prices are approaching the BofI.

Since the BofI test looks to see if prices are converged, rather than converging, the selection of an initial time period is less important. Additionally, it does not matter if the convergence path is linear or nonlinear, because the BofI test does not explicitly consider the convergence path. The only complexity comes from the calculation of the BofI. The unique setting for this study enables accurate measurement of the BofI.

2.2 Environment and Agents

2.2.1 Environment

WoW is one of the most popular games of its genre, with more than 12 million subscribers worldwide (Blizzard Entertainment, 2010). To comfortably accommodate so many players, WoW has been divided into more than 200 unique stand-alone copies of the game. These copies, also called worlds, realms, or servers, are identical in the rules of the game; the only differences lie in the populations and rules for interaction.

Each realm is separated into two approximately equal countries, “Horde” and “Alliance.” While different in appearance, the citizens of each country face the same physical and social rules and interact in the same environment with similar endowments.

There are four world types within WoW: normal, player versus player, role-playing, and role-playing player versus player. The player-versus-player worlds allow for unrestricted combat between citizens of different countries; Horde users can fight Alliance users and vice versa. The role-playing worlds encourage a more immersive experience. Users are encouraged to communicate as their avatar would communicate, and discouraged from referencing the outside world. In the normal worlds, players from the two countries cannot fight unless both formally agree, and there are no restrictions on the subjects discussed by users.

There is both international and national trade between agents. Both avenues of trade use a simple auction framework. A virtual currency called gold is used in both countries as the medium of exchange. Agents may offer items for sale, specifying a minimum bid, maximum buy-out price, and the duration of the auction. When placing an auction, a deposit based on the value and duration of the auction is required. If the auction is successful, the deposit is returned with the auction's earnings, minus a small fee. If the auction is unsuccessful, the deposit is forfeit. Agents purchasing from the auction are given the choice of bidding on an item and waiting to see if their bid is the winning bid or the agent may pay the buyout price and receive the good immediately. While both the national and international markets have the same framework, both the deposit and the fee schedule are higher in the international market.

2.2.2 Agents

Price convergence implicitly assumes the presence of arbitrageurs. An arbitrageur is an agent who exploits price differentials between two markets--purchasing goods from the lower-price market and selling the goods in the higher-price market. By purchasing in the lower-price market, the arbitrageur increases the demand for goods and raises the price. When the arbitrageur sells goods in the higher-price market, the supply is increased and the price is decreased. In the

absence of transaction costs, arbitrageurs will continue this process until the price in both markets reaches equality. At this point, the cost of trading equals the benefit from trading. Transaction costs increase the overall cost of trading, which halts the arbitrage process before prices have reached equality.

The presence of transaction costs adds an additional complication to the arbitrageur's profit maximization. Now the arbitrageur not only must observe the price differential between the two markets, the arbitrageur also must be cognizant of the transaction costs associated with trade. The problem moves from looking for a price differential to looking for a price differential greater than the transaction cost.

By changing the arbitrageur's profit maximization condition, the BofI is created. This is a range of values where the price differential is less than the transaction cost. Outside this band, there is an incentive to arbitrage between the two markets. This causes the two markets to converge into the band. Once inside the band, there is no further incentive to arbitrage. Then, both markets behave as separate entities. The arbitrageur always will trade when outside the BofI and never when inside the band. The process of arbitrage will drive the price differential back inside the BofI with relative speed.

2.3 Data

The raw data was collected from the online service, The Armory.⁵ This service is offered by Blizzard Inc., the publishers of WoW. Among other things, The Armory provides a real-time listing of all posted auctions in all countries on all worlds. For each auction, the name of the good on sale, the quantity of the good offered for sale, the name of the seller, the current or

⁵ <http://us.battle.net/wow/en/game/armory>

minimum bid, and the buyout price (maximum bid) is provided. However, this service does not display historical records; only the current auctions are shown.

A website, the Undermine Journal,⁶ does scan each auction house at least once an hour and provides a historical record of all auctions posted in each server. The website's developer was employed to track nine specific goods across eight worlds, in all auction houses.⁷ For each auction, all available information was included: name of good, quantity, seller's name, minimum bid, and buyout price.

Data was collected for eight worlds or server types. Two worlds were chosen randomly from each of the four world types: normal, player versus player, role-playing, and role-playing player versus player. World choices were made to ensure representation from each world type.

The data collection began on Dec. 4, 2010, three days before the release of the expansion Cataclysm on Dec. 7, 2010. Data collection ended on Jan. 31, 2011. Cataclysm added several new items, new territory, and modified existing territory, as well as made several changes to the game play. This represented a large shock that would affect all worlds and countries. The magnitude of the shock would vary based on the individual characteristics of each country and world.

⁶ <https://theunderminejournal.com/>

⁷ There goods tracked were (unless there is some reason I do not know about, I would lowercase these names) Netherweave Cloth, Frostweave Cloth, Embersilk Cloth, Greater Celestial Essences, Heavenly Shards, Infinite Dust, Heartblossom, Obsidion Ore, and Saronite Ore. The eight worlds studied were Azjol-Nerub, Draka, Earthen Ring, Firetree, Frostmane, Garrosh, Khaz'goroth, and Thrall.

2.3.1 Data Collection and Cleaning

The data were grouped by world and aggregated into daily per-unit observations. The data was presented each day with one file per world, per country, per item, per hour, for a total of 846 files each day. The files were condensed into daily, world/country/item files.

In each hour, sales were sorted by sale price. The highest sale price in that hour was recorded as the market-clearing price for that hour. The average daily price was calculated from hourly market prices. At this point, files were aggregated to one file per world per country.

At this level, three price indices were generated from the nine goods tracked in each country. The first index, “pre,” included only those items that existed before Cataclysm. The second index, “post,” included only new items introduced with Cataclysm. The final index, “index,” included all items. The indices were created as an average of all included goods weighted by the item’s sales volume as a fraction of the total index volume. The daily value for each index was used as the unit of observation for analysis in the later models. A full description of the data collection and preparation is available in Appendix A.

2.4 Models and Results

Two broad categories of tests are used in this chapter, mean and variance convergence. The Engle-Granger and BofI tests are the two mean convergence tests used in this chapter. The mean convergence tests focus on measuring the price levels and their behavior relative to each other. Each test determines convergence based on the price-level values. The Engle-Granger test looks for a common trend between price series, while the BofI test finds convergence when the price series are within an exogenously determined band. This allows both tests to find series that have converged; the Engle-Granger test also is able to find series that are converging.

The variance convergence tests measure changes in the dispersion or difference between the price levels. These tests find convergence when the variance between price levels consistently decreases after a deviation from price convergence. The ESTAR and Phillips and Sul tests are the two variance convergence tests used in this chapter. By focusing on the dispersion of prices, the variance convergence tests are able to detect price series that are converging, after many deviations from convergence. While these tests are able to find converging price series, they are unable to find converged price series. After series have converged, the cross-sectional variance would remain fairly stable or show no downward trend, giving the tests nothing to detect.

2.4.1 Engle–Granger Cointegration

2.4.1.A The Model

Early in LOP studies, the common method was to look for cointegration. If a linear combination of two or more nonstationary series is stationary, then the series are cointegrated. In other words, the two series follow a common trend. In the context of LOP studies, this corresponds to relative price convergence. That is, the price series do not reach equality, but they do not drift too far from each other either (Greene, 2003).

In their seminal paper, Engle and Granger (1987) introduced their two-step test for cointegration. First, a unit root test, such as the Augmented Dickey-Fuller (ADF) test, is conducted to determine if there is a unit root for each price series. The test also will determine the order of integration for each price series if a unit root exists. The ADF test has a null of a random walk, or no unit root. The test begins with an auxiliary regression of the series.

$$\Delta y_t = \mu + \beta t + (\theta - 1)y_{t-1} + \sum \delta_i \Delta y_{t-i} + \varepsilon_t \quad (1)$$

Where θ is the sum of all δ_i , and ε_t is an i.i.d. error term. A true random walk form is obtained by restricting $\mu = \beta = 0$. A random walk with drift has the restriction $\beta = 0$, while the trend stationary model allows μ and β to be free. The null hypothesis of $\theta = 1$ is tested with the following test statistic

$$DF_t = \frac{\hat{\theta}-1}{Est.Std.Error(\hat{\theta})} \quad (2)$$

The test statistic is not normally distributed. Special critical values tabulated in Fuller (1976) and by Dickey and Fuller (1981) are required. Both price series must reject the null and be integrated to the same degree.

If this occurs, then an OLS regression is run on the two price series and the residuals are extracted.

$$y_t = \beta x_t + \varepsilon_t \quad (3)$$

Where y_t and x_t are the two price series, ε_t is the residual to be extracted. The Engle-Granger test then is conducted on ε_t . The Engle-Granger test is similar in form to the ADF; however, in this case the null hypothesis is no cointegration. In other words, the null hypothesis of the Engle-Granger test is the residual is nonstationary.

While the Engle-Granger test is similar in form to the Dickey-Fuller test, the critical values are not the same. The standard critical values from the Dickey-Fuller test bias the results to rejecting the null. To compensate for this, Engle and Granger (1987) suggest using special critical values based on the sample size and number of cointegrating variables. MacKinnon (2010) has provided the special critical values in a table for several combinations of sample size and cointegrating variables. If the residual is stationary (i.e., reject the null), then the series are cointegrated.

2.4.1.B Results

Engle-Granger tests were conducted for each index in each world separately. A total of 24 convergence tests were conducted. Of the 24 tests, 10 indices had price series that did not contain a unit root or were not integrated to the same degree. In the table below, a zero is used to indicate these tests.

Among the remaining tests, each has a slightly different number of observations, which means that critical values for each test are different. In the table below, the test statistic is reported for each Engle-Granger test. Asterisks indicate the level of significance for tests that reject the null of no cointegration.

Table 2.1 Results of the Engle-Granger Cointegration Test

	Pre	Post	Index
Azjol-Nerub-1	-7.283***	-6.930***	0
Draka-2	0	-5.083***	0
Earthen Ring-3	-8.465***	0	0
Firetree-4	-4.850***	-6.074***	-4.975***
Frostmane -5	-5.530***	0	0
Garrosh-6	-4.973***	-5.173***	-5.240***
Khaz'goroth-7	0	0	-4.774***
Thrall-8	-3.419*	0	-4.794***

*, **, ***, indicate significance at the 10%, 5%, and 1% level, respectively. The 0 values indicate that the price series were not integrated to the same degree.

From the table above, it can be seen that 14 of the 24 indices had series that were integrated of the same degree. All indices with common degrees of integration were cointegrated as well.

Cointegration implies that the series tested already were converged, with a linear reversion after a shock. These results, therefore, mirror the results found using the BofI test that also looks for converged behavior. The next two tests conducted look for mean-reverting behavior. By looking

at mean-reverting behavior, these tests are able to account for nonlinear reversion, something the Engle-Granger test is not able to do.

2.4.2 Exponential Smooth Transition Autoregressive (ESTAR)

2.4.2.A The Model

The Engle-Granger test supports price convergence only if the series are converged for a majority of the time and the reversion is linear. The presence of the BofI implies nonlinear reversion. This means the standard cointegration analysis is biased against price convergence.

The ESTAR model uses the residual from an OLS regression of both price series, as does the Engle-Granger test. However, the ESTAR model estimates two separate regimes for the residual, within and outside of the BofI. The ESTAR model test focuses on the behavior in the second regime outside of the BofI. Outside of the BofI, arbitrage profits are possible. Therefore, the error term should show mean-reverting behavior and support for price convergence. These two separate regimes are modeled as a function of the error term, y_t , as shown below (Granger & Terasvirta, 1993):

$$y_t = AR(p) + AR(p)^* \times F(y_{t-d}) \quad (4)$$

or,

$$y_t = k + \sum_{j=1}^p \pi_j y_{t-j} + (k^* + \sum_{j=1}^p \pi_j^* y_{t-j}) \times F(y_{t-d}) + \varepsilon_t \quad (5)$$

Where

$$AR(p) = k + \sum_{j=1}^p \pi_j y_{t-j}; AR(p)^* = (k^* + \sum_{j=1}^p \pi_j^* y_{t-j}); F(y_{t-d}) = 1 - \exp [-\psi(y_{t-d} - c^*)^2] \quad (6)$$

π_j^* and π_j are the autoregressive parameters for each regime, k and k^* are constant terms for each regime, and ε_t is an i.i.d. error term. The function $F(y_{t-d})$ dictates the transition from one regime to another. It is a U-shaped function bounded between 0 and 1. The term c^* dictates the center, and ψ controls the width of the transition function. $F(y_{t-d})$ is a positive function of y_{t-d} .

When $F(y_{t-d})$ approaches 0, it is near the bottom of the U. In this area, the dominant regime is inside of the BofI ($AR(p)$).

$$y_t = k + \sum_{j=1}^p \pi_j y_{t-j} \text{ when } F(y_{t-d}) = 0 \quad (7)$$

This regime shows no mean-reverting behavior and is expected to be governed by individual market forces. The behavior of the residual is assumed to follow a random walk when within the BofI. As $F(y_{t-d})$ approaches 1, the values approach the outer edge of the U. Then the dominant regime shifts to outside the BofI ($AR(p) + AR(p)^*$).

$$y_t = (k+k^*) + \sum_{j=1}^p (\pi_j + \pi_j^*) y_{t-j} + \mu_t \text{ when } F(y_{t-d}) = 1 \quad (8)$$

This regime exhibits mean-reverting behavior. The key factor in determining the regime, therefore, is the transition function. The wider the transition function or greater distance horizontal distance between $F(y_{t-d}) = 0$ and $F(y_{t-d}) = 1$, the larger the BofI.

Following Michael et al. (1997), the ESTAR model was adjusted so each regime mirrors the ADF test. This allows for convergence testing after first testing for nonlinearity and homogeneity in the price series.

$$\Delta y_t = k + \lambda y_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta y_{t-j} + (k^* + \lambda^* y_{t-1} + \sum_{j=1}^{p-1} \psi_j^* \Delta y_{t-j}) \times F(y_{t-d}) + \mu_t \quad (9)$$

Like the ADF test, the important parameters for convergence are λ and λ^* . If $\lambda^* < 0$ and $\lambda + \lambda^* < 0$ are true, then price convergence holds outside of the BofI. As the behavior of price levels inside the BofI is effectively a random walk $\lambda \geq 0$ is possible.

2.4.2.B Results

As with the Engle-Granger test, each index in each world was tested for convergence separately. Values for λ^* and $\lambda + \lambda^*$ are listed below, with asterisks to indicate the level of significance. Only when λ^* and $\lambda + \lambda^*$ reject the null of ≥ 0 is there price convergence. The last row for each world indicates if the index shows price convergence or not.

Table 2.2 Results from ESTAR Model Tests

		Pre	Post	Index
Azjol-Nerub-1	λ^*	22.49	-98.83**	-0.17
	$\lambda + \lambda^*$	-45.24+22.49	197.30-98.83	-5.47e-12-0.17
	Converge/Fail	Fail	Fail	Fail
Draka-2	λ^*	-3.41	-0.49	7.05
	$\lambda + \lambda^*$	6.09-3.41	6.56e-12-0.49	-14.48-7.05
	Converge/Fail	Fail	Fail	Fail
Earthen Ring-3	λ^*	-46.94	-0.07	5.69
	$\lambda + \lambda^*$	93.6-49.94	-0.92-0.07***	-12.24+5.69
	Converge/Fail	Fail	Fail	Fail
Firetree-4	λ^*	2.79	29.2	9.07
	$\lambda + \lambda^*$	-6.35+2.79	-59.38+29.2*	-18.82+9.07
	Converge/Fail	Fail	Fail	Fail
Frostmane -5	λ^*	3.28e-24	-32.32	0.03
	$\lambda + \lambda^*$	-9.16-3.28e24***	63.56 -32.32	-0.46+0.03
	Converge/Fail	Fail	Fail	Fail
Garrosh-6	λ^*	-6.15	-0.84**	-50.57
	$\lambda + \lambda^*$	11.63-6.15	0.02-0.84***	100.35-50.57
	Converge/Fail	Fail	Fail	Fail
Khaz'goroth-7	λ^*	-0.02	-28.23	4.14
	$\lambda + \lambda^*$	-0.67-0.02	55.53-28.23	-8.74+4.14
	Converge/Fail	Fail	Fail	Fail
Thrall-8	λ^*	21.59	-0.49	-0.08
	$\lambda + \lambda^*$	-44.18+21.59***	-5.61e-12-0.49	-0.06-0.08
	Converge/Fail	Fail	Fail	Fail

*, **, ***, indicate significance at the 10%, 5%, and 1% level, respectively. Indices that had significant values for λ^* and $\lambda + \lambda^*$ indicate convergence, indicated by “converge.” Those that did not achieve significance in either or both λ^* and $\lambda + \lambda^*$ did not converge, which is indicated by “fail.”

All indices failed to show mean-reverting behavior outside of the BofI. This could be driven by the relatively low number of observations outside of the BofI, as shown by the BofI test. With few observations outside of the BofI, the ESTAR model would attempt to force an edge to the transition function within the BofI, rather than at the true edge of the BofI. The pattern is consistent with the other price convergence tests used in this paper.

2.4.3 Phillips and Sul

2.4.3.A The Model

Given the initial market shock described above, it is natural to look for convergence behavior rather than for a presence inside the BofI. The Phillips and Sul Log t convergence test is a flexible test for convergent behavior. Rather than looking for a unit root as in the Engle-Granger test or looking for convergent behavior outside the BofI as in the ESTAR test, the Phillips and Sul test looks for decreases in cross-sectional variance over time. The test assumes the price series are in disequilibrium at the initial observation, then tracks the movement back to equilibrium. Equilibrium is defined as within the BofI.

The Phillips and Sul test is a modified sigma convergence test. The test uses decreasing cross-sectional variance over time as a measure of price convergence. Phillips and Sul is different from other tests in that the convergence need not be linear. Each series can converge at a separate and varying slope. This makes the Phillips and Sul test one of the more flexible convergence tests available.

The Phillips and Sul test assumes that each price series is composed of a common, unobservable component, μ_t , multiplied by a stochastic factor loading coefficient, δ_{it} . The null hypothesis of convergence for the test is defined as

$$H_0: \delta_{it} \rightarrow \delta \text{ for all } \delta.$$

A nonparametric method is used to estimate the time-varying loading coefficients. A relative transition parameter h_{it} is derived that is equivalent to a relative δ_{it} .

$$h_{it} = \frac{X_{it}}{\frac{1}{N} \sum_{i=1}^N X_{it}} = \frac{\delta_{it}}{\frac{1}{N} \sum_{i=1}^N \delta_{it}} \quad (10)$$

If all the δ_{it} converge to δ , then the relative transition parameter, h_{it} , will converge to one, as the cross-sectional variance converges to zero. From the relative transition parameter, the cross-sectional variance, H_{it} , is computed for use in the final test.

$$H_{it} = \frac{1}{N} \sum_{i=1}^N (h_{it} - 1)^2 \quad (11)$$

The cross-sectional variance over time is compared to the original cross-sectional variance at time $t=1$. If the series are converging, then the cross-sectional variance will be decreasing. As the variance decreases, the difference between the current variance and the variance at $t=1$ will become more pronounced. Using this pronounced difference between variances over time, Phillips and Sul perform an OLS regression of the following equation:

$$\text{Log} \left(\frac{H_1}{H_t} \right) - 2\text{Log}(L(t)) = \hat{a} + \hat{b} \log(t) + \hat{u}_t \quad (12)$$

Where H_1 is the cross-sectional variance at time $t=1$, and $L(t) = \log(t+1)$. Phillips and Sul recommend estimating the equation after dropping the first 30% of the observations. Dropping the observations is necessary for the limit distribution and power properties of the test. The series are said to be converging if \hat{b} is greater than zero. A simple one-sided t-test of \hat{b} is performed using heteroskedastic and autocorrelation consistent (HAC) standard errors.

2.4.3.B Results

Convergence between the two countries in each world is tested separately for each index, leading to a total of 24 convergence tests. The \hat{b} t values for the 24 indices tested are listed in the table below. A t-value greater than -1.65 indicates failure to reject the null hypothesis of convergence.

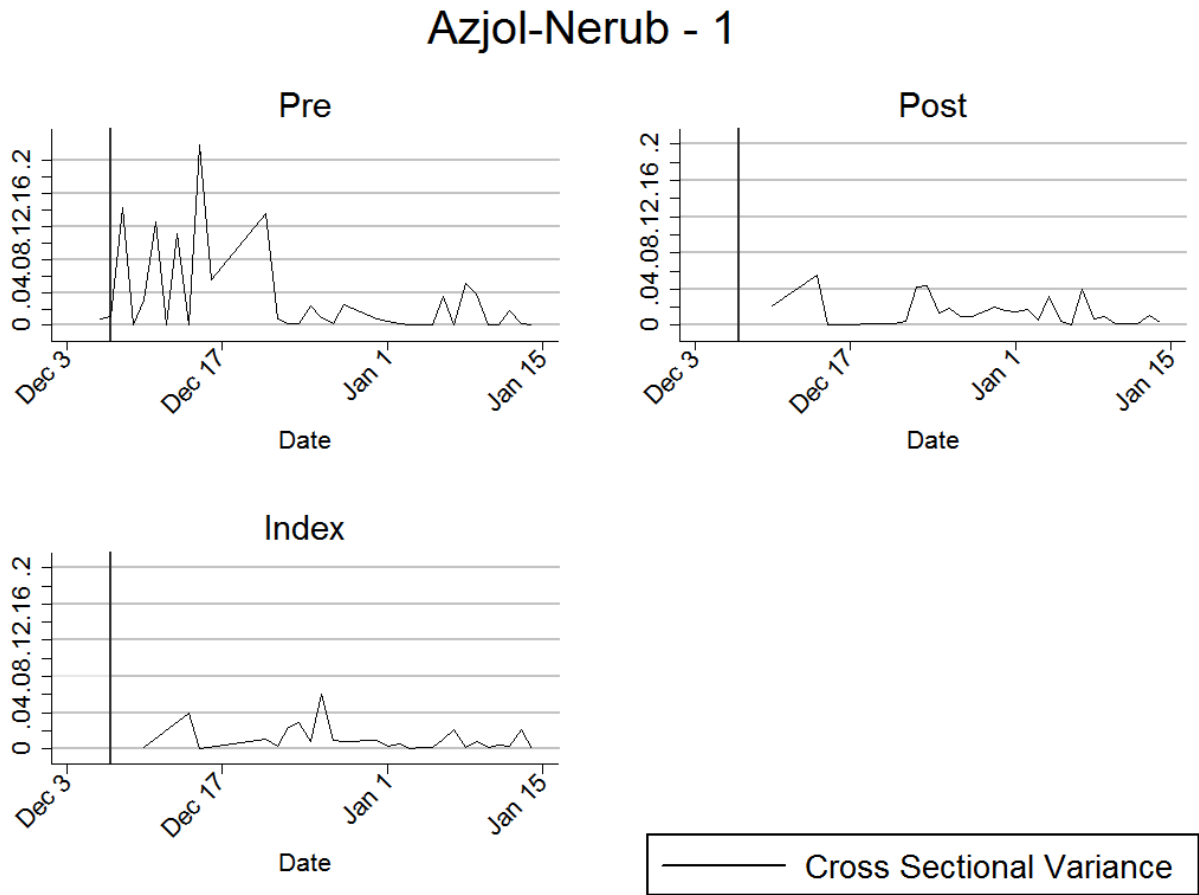
Table 2.3 Results from Phillips and Sul tests

	Pre	Post	Index
Azjol-Nerub-1	-52.86	-119.04	-43.25
Draka-2	-41.36	-147.62	-27.49
Earthen Ring-3	59.73	6.53	-58.12
Firetree-4	-27.16	-8.87	-26.99
Frostmane -5	-57.30	13.38	-107.06
Garrosh-6	-28.86	59.57	-4.97
Khaz'goroth-7	80.99	-5.86	-87.95
Thrall-8	12.56	-17.28	-76.19

T-values for b . T-Values greater than -1.65 indicate convergent behavior.

As can be seen, most of the series reject the null hypothesis of convergence. In other words, the cross-sectional variances of many of the series are not decreasing overtime. It was hoped that by looking at the cross-sectional variances right after a major shock to the market (introduction of Cataclysm), it would be possible to capture convergent behavior. Figure 1 presents the cross-sectional variances for each index over time for a representative world. The variance decreases quickly after the shock, usually in matter of days or weeks. Often the convergence after the shock occurs during the period dropped from analysis by the Phillips and Sul test. In some cases, there is little or no apparent reaction to the shock. It could be the magnitude of the shock was approximately equal between the two countries. If this were true, both countries would see similar price movements and no real change in variance.

Figure 2.1 Cross-sectional variances for three indices in World 1 (Azjol-Nerub)



Figures for all eight worlds are in Appendix B.

By looking only at the change in cross-sectional variance of series, Phillips and Sul are able to capture only series that are outside of the BofI and are in the process of converging. If the price series are within the BofI, then each series would be governed by market forces in each country. A decrease in cross-sectional variance would be only coincidental. Because the Phillips and Sul test is useful only in testing convergence among series outside of the BofI, a different test must be used instead. This different test would indicate if the price series were inside the BofI. If the series are inside the BofI, then the series have converged as much as possible and the Phillips

and Sul test is unnecessary. However, if the series lay outside the BofI, then the Phillips and Sul test would demonstrate the convergent behavior.

2.4.4 Band of Inaction (BofI) test

2.4.4.A The Model

Taylor (2001) defines the BofI in a TAR-like form below

$$x_t = \begin{cases} +c + \gamma(x_{t-1} - c) + e_t & \text{if } x_{t-1} > c, \\ x_{t-1} + e_t & \text{if } c \geq x_{t-1} \geq -c, \\ -c + \gamma(x_{t-1} + c) + e_t & \text{if } -c > x_{t-1}, \end{cases} \quad (13)$$

In this form, the x_t is the price differential or price gap, $x_t = P_{1,t} - P_{2,t}$, where $P_{i,t}$ is one of the price series converging, and c is the amplitude or half the width of the BofI. Therefore the BofI is defined as $c \geq x_{t-1} \geq -c$. The edges can be redefined when considered in absolute value terms as

$$|x_{t-1}| \leq |c| \quad (14)$$

Or

$$|P_{1,t-1} - P_{2,t-1}| \leq |c| \quad (15)$$

If we define $P_{1,t}$ as the highest price level at time t and $P_{2,t}$ as the lowest value at time t , then the BofI edges can be redefined as

$$P_{H,t-1} - P_{L,t-1} \leq |c| \text{ where } P_{H,t-1} = P_{1,t-1} \text{ and } P_{L,t-1} = P_{2,t-1} \quad (16)$$

Previous work has treated c as an endogenous term to be estimated based on the autoregressive behavior of x_t . In this work, however, we are able to empirically estimate the values c independent of the behavior of x_t .

Theoretically, c is the sum of transaction costs. The definition of the Boff's width then can be restated as the following inequality

$$B \leq C \quad (17)$$

Where B is the benefit from trade or price differential, $B = P_H$, and $C = |c| + P_L$.⁸ Within WoW, international trade is conducted through auctions. Because the costs of initiating an auction and the success rate of auctions are known, accurate measurement of the cost of trade can be determined, though some restructuring of the benefits from trade are required.

The benefit is the price received from a sale in the higher-priced market if the auction is successful. If the auction is unsuccessful, then the benefit is zero. There is a probability ρ that the auction will be successful, and therefore the benefit can be re-expressed as $B = \rho P_H + (1 - \rho)0$ or $B = \rho P_H$.

The cost of arbitrage is more complicated. First is the cost purchasing the good in the lower-priced market, P_L . Additionally, there is the cost of transport as well as any tariffs, fees, and other costs associated with trade. Within the virtual world, the transportation cost is negligible.

The transaction costs are incurred from the use of the international auction. If a good is successfully sold at auction, the seller is charged 15% of the sale price as an auction fee, or $0.15 * P_H$. If the good is not sold, the seller surrenders the deposit. The deposit is based on the duration of the auction and the minimum value of the good. The minimum value of the good, known as the merchant sale value (MSV), is the price received if the good were sold to a computer-generated agent. The MSV is adjusted by a factor, τ , based on the auction's duration. τ takes the values 75%, 150%, or 300% for a 12-, 24-, or 48-hour auction, respectively. As defined

⁸ The time subscript is dropped for simplicity.

above, the probability of a successful auction is ρ . Therefore, the total cost of arbitrage can be represented by the following equation:

$$C = \rho(.15 * P_H) + (1 - \rho)(\tau * MSV) + P_L; \text{ where } C = |c| \quad (18)$$

By inputting the equations for benefit and cost into the inequality above, the BofI is now defined as

$$\rho(.15 * P_H) + (1 - \rho)(\tau * MSV) \geq P_{H,t-1} - P_{L,t-1} \quad (19)$$

Using simple algebra, the inequality can describe the BofI for any one good.

$$P_L + (1 - \rho)(\tau * MSV) - \rho * .85P_H \geq 0 \quad (20)$$

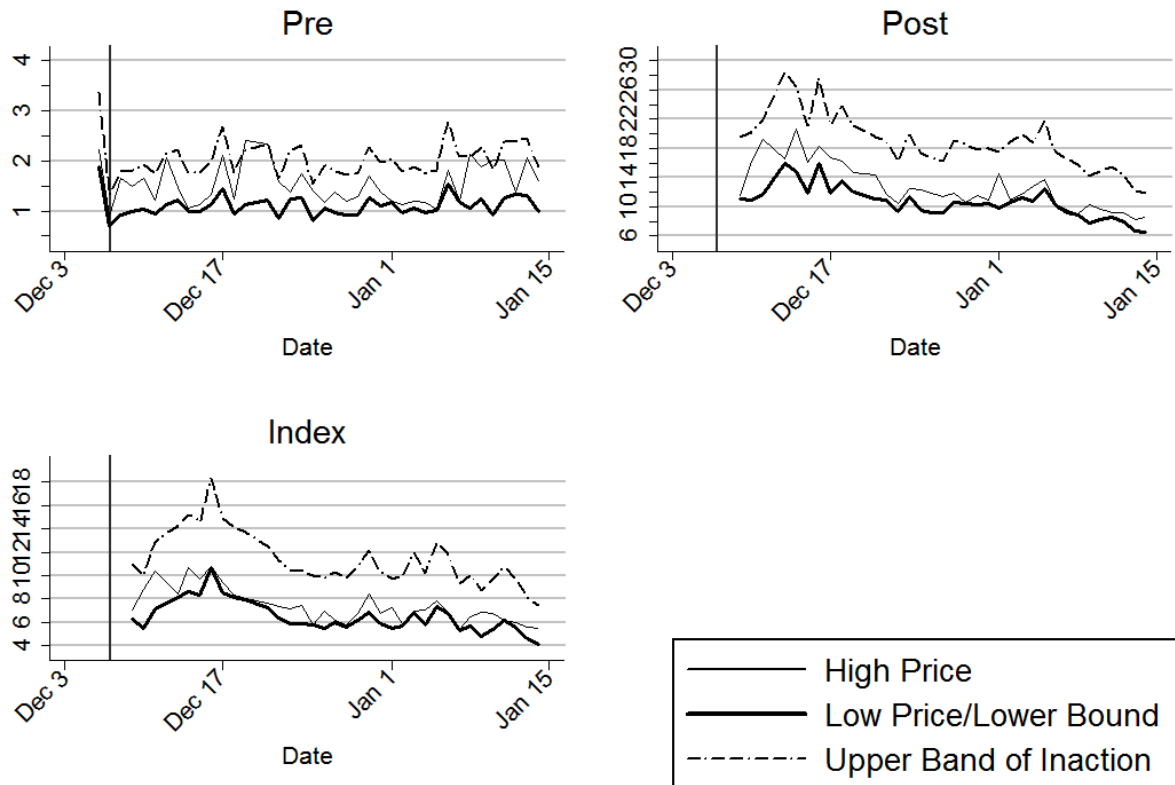
To test for price convergence, an aggregate price-level condition must be derived. Aggregate prices must be used rather than individual prices. The μ must be expanded into the respective share of sellers who chose a 12-, 24- or 48-hour auction, defined as Θ_i . ρ now represents the sample's share of successful auctions, rather than the arbitrageurs' probability. The economy-wide condition becomes:

$$P_L + (1 - \rho)(\theta_1 * .75 + \theta_2 * 1.5 + \theta_3 * 3)MSV - \rho * .85P_H \geq 0 \quad (21)$$

After inputting values for ρ , Θ , MSV , P_L , and P_H , a value greater than 0 indicates the cost of arbitrage is greater than the benefit. Therefore, the prices lie within the BofI, and no convergent behavior should be observed. Figure 2 below shows the price levels and BofI for all indices in a sample world.

Figure 2.2 Price index values and BofI for World 5 Frostmane.

Frostmane - 5



Figures for all eight worlds are in Appendix C.

2.4.4.B Results

To provide a more formal proof of this, a simple one-sided t-test was performed on the inequality above. The t-values are reported below. A t-value greater than -1.65 indicates the inequality holds and that prices have converged as much as possible.

Table 2.4. T-values for BofI test

	Pre	Post	Index
Azjol-Nerub-1	15.68	28.08	26.34
Draka-2	17.72	29.95	18.49
Earthen Ring-3	20.62	31.00	24.66
Firetree-4	21.05	31.39	24.18
Frostmane-5	30.48	29.44	24.31
Garrosh-6	21.61	25.23	26.07
Khaz'goroth-7	26.36	36.33	27.83
Thrall-8	13.17	31.65	22.66

T-Values greater than -1.65 indicate a converged realm.

As can be seen, all t-values are significantly greater than -1.65. This means that for all indices, the cost of arbitrage is greater than the benefit. In other words, the indices lie within the BofI. It is not surprising that so many of the indices rejected the null of convergence in the Phillips and Sul test and that all indices rejected convergence using the ESTAR test. Each index spent the majority of the time observed within the BofI. Therefore, there were few observations outside the BofI to define the second regime in the ESTAR models. Additionally, within the BofI, each index responded to individual market forces. Any decrease in cross-sectional variance would be coincidental.

This mirrors the results from the Engle-Granger test, which found the index differential was stationary for most indices and worlds. The failure of the Engle-Granger test to find convergence in all worlds and indices could be attributed to the random walk-like behavior predicted within the BofI. This behavior could have falsely given the impression of nonstationarity.

2.5 Conclusion

The results are consistent among the two types of convergences tests. The mean convergence tests find convergence, while the variance convergence tests fail to find convergence. Using the expanded BofI test, we show that all indices lie inside the BofI. This result is supported by similar though weaker results using the Engle-Granger cointegration test. As the indices

remained primarily within the BofI, each index was determined by market forces in each individual country. This could give the appearance of a random walk and could explain the weaker results from the Engle-Granger test.

Because the majority of observations were within the BofI, few observations could be used by the variance convergence tests. This meant that the ESTAR model was unable to define the regime outside of the BofI. The model then would attempt to force the creation of a second regime from observations within the BofI. This explains the failure of the ESTAR model to find convergence in any of the indices tested. The paucity of observations outside the BofI also would negatively affect the performance of the Phillips and Sul test as well. Each price series quickly entered the BofI and followed a random walk-like path. Therefore, any sustained reduction in cross-sectional variance would be coincidental.

These results highlight the need for a portfolio-type approach when testing for LOP and price convergence. It is possible to falsely reject convergence; if only one type of convergence test is used, both mean and variance convergence tests should be used. Even after a major shock such as Cataclysm in WoW, price levels could remain within the BofI so consistently as to be unobservable with variance convergence tests. It's possible that even in relatively calm periods without easily identifiable shocks, price levels may be so volatile as to prevent detection using mean convergence tests. This problem is compounded with temporal aggregation and linear specifications, as Taylor (2001) points out. A price convergence test should search for converged price series as well as for converging price series.

Using virtual worlds for macroeconomic research has the advantage of allowing for more precise data collection. Within WoW, we were able to conduct a simple, natural experiment overcoming

many of the issues faced by other LOP studies. We were able to collect accurate daily data. This enabled us to graphically observe the rapid price convergence, though we were unable to observe the mean-reverting behavior with either the ESTAR test or Phillips and Sul tests.

Online games such as WoW, which create a robust virtual world, are becoming increasingly popular and provide numerous opportunities for natural economic experiments, as demonstrated above. There is a great potential for approaching old questions in new ways, leading to unique answers and solutions to real world problems.

Chapter 3 Variable Band of Inaction: Transaction Costs and the Size of the Band of Inaction

Abstract: The presence of the band of inaction (BofI) in purchasing power parity (PPP) studies is well established. However, few studies have modeled the behavior of the BofI in response to changes in transaction costs. The BofI is the range of price differentials where the cost of trade is greater than the benefit. In this paper, an econometric model is developed to model changes in the BofI in response to changes in transportation costs. Price levels and exchange rates for 14 developed countries in relation to the United States are used in the analysis. The econometric model is based on the smooth transition autoregressive model used by Michael et al. (1997) to test for PPP. The model is further expanded to accommodate a panel data set following Gonzalez et al. (2005). The results demonstrate that the width of the BofI does move with changes in transportation costs.

3.1 Introduction

In the presence of transaction costs, purchasing power parity (PPP) models imply the existence of a band of inaction (BofI). Theoretically, the BofI describes a range of price differentials where the cost of arbitrage is greater than the benefit. Several studies have found empirical support for the BofI, such as those by Taylor (2001) , Nakagawa (2002), Heimonen (2006), and Pesaran et al. (2009), to name a few. Because the BofI depends on transaction costs, it should vary with transaction costs. The degree of variation would indicate the relative importance of transaction costs in determining the BofI. However, no studies have used empirical tests to confirm the connection between the BofI and transaction costs. In this paper, a variation of the smooth transition autoregressive (STAR) model is used to model this relationship between the BofI and transaction costs, specifically transportation costs.

Many authors, such as Michael et al. (1997), Baum et al. (2001), and Bahmani-Oskooee et al. (2007), use variations of the STAR model to account for nonlinearity in PPP. The STAR model in essence estimates two different autoregressive models, one for each regime. In the context of PPP, the two regimes correspond to inside and outside of the BofI. The transition between regimes is described by a transition function. This is an exponential function of deviations from PPP, bounded between zero and one. Small deviations from PPP remain within the BofI and the inner regime. Meanwhile, large deviations from PPP exit the BofI and enter the second, outer regime.

Transaction costs are difficult to measure, but it is possible to measure some of the individual components. By looking at the components, it is possible to determine the relative importance of each component in the BofI. The most recognized components of transaction costs include transportation costs, price uncertainty, tariffs, quotas, and other nontariff trade barriers. In this

analysis, transportation costs are examined as transportation costs have fallen dramatically over the past 30 years. Hummels (2007) calculates that transportation costs have fallen more than 25% from 1974 to 2004.

Using a panel data set of price levels and exchange rates, the model estimates two regimes governed by an exponential transition function. The transition function is modified to be a function of two separate variables. This allows the transition function to define the boundaries of the BofI as well as allows the boundaries to fluctuate with transportation costs. The model is further modified to use panel data. The model estimated is therefore a modified panel exponential STAR model.

3.2 Purchasing Power Parity and Arbitrage

PPP can exist only in an environment that supports arbitrage. Arbitrageurs exploit the price differentials between markets by purchasing goods from a low-price market, then transporting and selling the goods in a high-price market. By purchasing in the low-price market, the arbitrageur increases the demand for goods and raises the price. When the arbitrageur sells goods, the supply is increased and the price decreases in the high-price market. In the absence of transaction costs, arbitrageurs continue this process until the price in all markets reaches equality. At this point, the cost of trading equals the benefit. Transaction costs increase the cost of trading, which halts the arbitrage process before prices have reached equality.

Transaction costs add a complication to the arbitrageur's profit maximization function. Now the arbitrageur also must observe the transaction costs associated with trade. The problem is altered from looking for a price differential to looking for a price differential greater than the transaction cost.

By changing the arbitrageur's profit maximization condition, the BofI is created. Outside this band is an incentive to arbitrage between markets. Once inside the band, there is no further incentive to arbitrage. Both markets are driven by internal forces effectively following a random walk.

When studying single-goods markets, the transition from inside to outside the BofI is discrete. When price levels are considered, as in PPP and price convergence studies, the transition into the BofI becomes smooth. Each good in the basket faces different transaction costs; therefore, goods arbitrage ceases at a different point for each good. This necessitates a smooth rather than discrete transition function.

3.2.1 Purchasing Power Parity Model and Estimation Method

The traditional PPP condition is written as

$$E_t - P_t^* + P_t = b + y_t \quad (1)$$

Where E_t is the log of the nominal exchange rates⁹; P_t^* and P_t are logs of the foreign and domestic price levels, respectively; y_t is an error term measuring deviations from PPP; and b is a constant term representing differences in the basket of goods and units of measurement. This is commonly restated for regression analysis as

$$E_t = b + \alpha P_t^* + \beta P_t + y_t \quad (2)$$

With the restriction $\alpha = -\beta = 1$ imposed. Linear PPP can be tested with this equation using the cointegration framework demonstrated by Engle and Granger (1987). This requires that the error

⁹ Quoted as foreign exchange per dollar (FE/\$)

term, y_t , follow a stationary process, given a nonstationary process followed by the other variables. In this scenario, any deviation from PPP reverts to the mean on a linear path.

This standard cointegration analysis supports PPP only if the reversion is linear. The presence of transaction costs implies nonlinear reversion. This means the standard cointegration analysis is biased against PPP. When inside the BofI, no arbitrage occurs and the error term would show unit-root behavior and fail to support PPP. Outside the BofI where arbitrage profits are possible, the error term shows mean reverting behavior and support for PPP. These two separate regimes can be model with the exponential STAR (ESTAR) model as a function of the error term, y_t (Granger & Terasvirta, 1993):

$$y_t = AR(p) + AR(p)^* \times F(y_{t-d}) \quad (3)$$

or,

$$y_t = k + \sum_{j=1}^p \pi_j y_{t-j} + (k^* + \sum_{j=1}^p \pi_j^* y_{t-j}) \times F(y_{t-d}) + \mu_t \quad (4)$$

Where

$$AR(p) = k + \sum_{j=1}^p \pi_j y_{t-j}; \quad AR(p)^* = (k^* + \sum_{j=1}^p \pi_j^* y_{t-j}); \quad F(y_{t-d}) = 1 - \exp [-\phi(y_{t-d} - c^*)^2] \quad (5)$$

The transition function $F(y_{t-d})$ is an exponential cumulative distribution function. It is a U-shaped function bounded between 0 and 1. The term c^* dictates the center of the transition, and ϕ controls the width of the transition function. $F(y_{t-d})$ is a positive function of y_{t-d} , the lagged error term from equation 2. When $F(y_{t-d})$ approaches 0, it is near the bottom of the U. In this area, the dominant regime is the inner BofI ($AR(p)$). This regime shows no mean reverting behavior and is expected to follow a random walk. As $F(y_{t-d})$ approaches 1, the values

approach the outer edge of the U. Then the dominant regime shifts to outside the BofI ($AR(p) + AR(p)^*$). This regime exhibits mean reverting behavior. The key factor in determining the regime is the transition function. The wider the transition function or greater the horizontal distance between $F(y_{t-d}) = 0$ and $F(y_{t-d}) = 1$, the larger the BofI.

ϕ controls the width of the transition function and BofI. If $\phi = 0$ then the model collapses to a simple $AR(p)$ process with no outer regime. As $\phi \rightarrow 0$, the BofI widens, and only very large deviations from PPP leave the BofI. If $\phi \rightarrow \infty$, then the model becomes a discrete threshold autoregressive (TAR) model with a narrow BofI dominated by the outer regime. As ϕ approaches infinity, the BofI narrows and only small deviations from PPP fail to leave the BofI.

Before estimation of the nonlinear STAR model, the series first must be tested for nonlinearity. Teräsvirta (1994) suggests a Lagrange Multiplier (LM)-type linearity test. If the delay parameter, d , is fixed, then the linearity test for an ESTAR model consists of an OLS estimation of the following artificial regression

$$y_t = \beta_{00} + \sum_{j=1}^p (\beta_{0j}y_{t-j} + \beta_{1j}y_{t-j}y_{t-d} + \beta_{2j}y_{t-j}y_{t-d}^2) + \epsilon_t \quad (6)$$

Then testing the null hypothesis of linearity

$$H_{0L}: \beta_{1j} = \beta_{2j} = 0 \quad (j = 1, \dots, p)$$

Michael et al. (1997) suggest using an ordinary F-test as an approximation of the LM-type test, because the F-test has improved size and power properties.

The approach suggested by Tsay (1989) is used to determine the delay parameter. For this approach, the linearity test is repeated for several plausible values of d . The delay parameter used in the final model is the value of d that minimizes the P-value of the linearity test.

Assuming the series rejects the null of linearity and a delay parameter is determined following Tsay (1989), then the ESTAR model can be used. Following Michael et al. (1997), the STAR model was adjusted to allow for convergence testing. This new form contains the Augmented Dickey-Fuller (ADF) model within each regime.

$$\Delta y_t = k + \lambda y_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta y_{t-j} + (k^* + \lambda^* y_{t-1} + \sum_{j=1}^{p-1} \psi_j^* \Delta y_{t-j}) \times F(y_{t-d}) + \mu_t \quad (7)$$

Like the ADF model, in this form the important parameters for convergence are λ and λ^* . If $\lambda^* < 0$ and $\lambda + \lambda^* < 0$, then PPP holds outside of the Bofl. As the behavior of price levels inside the Bofl is effectively a random walk, $\lambda \geq 0$ is possible.

Given the complexity of the ESTAR model, there was some doubt that the price series used in this analysis contained enough observations to estimate each country individually. Also, the large number of countries included in this paper could make reporting results difficult and confusing. For these reasons, the model was transformed into a panel model following Gonzalez et al. (2005). Prior to reformulation as a panel model, the series first must be tested for homogeneity. If the data generating process is homogeneous, then the panel ESTAR model is unidentified. Therefore, a homogeneity test is needed to avoid creating an unidentified model. The homogeneity test suggested by Gonzalez et al. (2005) takes a form similar to the linearity test in Teräsvirta (1994). The null hypothesis of homogeneity is represented as $H_0: \phi = 0$ or $H_0: \beta_j = 0$ in the auxiliary regression below. The auxiliary regression is derived from a first-order Taylor series expansion around $\phi = 0$.

$$y_t = \mu_{it} + \beta_0 y_t + \beta_1 y_{i,t} y_{i,t-d} + \beta_2 y_{i,t} y_{i,t-d}^2 + \mu_{it}^* \quad (8)$$

Where β_1, β_2 are multiples of ϕ , and $\mu_{it}^* = R_M \beta_1 y_{it}$. R_M is the remainder of the Taylor series expansion. This means testing the null hypothesis of $\beta_1 = \beta_2 = 0$ is equivalent to the null hypothesis of $H_0: \phi = 0$. The null hypothesis can be tested using a LM test (Gonzalez, *et al.*, 2005). Assuming the series rejects the null of homogeneity, equation 7 can be rewritten as

$$\Delta y_{it} = u_i + \beta' x_{it}(y_{t-d}: \phi, c^*) + u_{it} \quad (9)$$

Where

$$\beta' x_{it}(\phi, c^*) = k + \lambda y_{t-1} + \sum_{j=1}^{p-1} \psi_j \Delta y_{t-j} + (k^* + \lambda^* y_{t-1} + \sum_{j=1}^{p-1} \psi_j^* \Delta y_{t-j}) \times F(y_{t-d}: \phi, c^*) \quad (10)$$

The individual means then were subtracted from the equation, yielding

$$\widetilde{\Delta y}_{it} = \beta' \widetilde{x}_{it}(y_{t-d}: \phi, c^*) + \widetilde{u}_{it} \quad (11)$$

Where $\widetilde{\Delta y}_{it} = \Delta y_{it} - \overline{\Delta y}_i$, $\widetilde{x}_{it} = (x_{it} - \bar{x}_i, x_{it} F(y_{t-d}: \phi, c^*) - \overline{w}_i(\phi, c^*))$, $\widetilde{u}_{it} = u_{it} - \bar{u}_i$. The variables $\overline{\Delta y}_i$, \bar{x}_i , and \bar{u}_i are individual means and $\overline{w}_i(\delta \tau_t) \equiv T^{-1} \sum_{t=1}^T x_{it} F(y_{t-d}: \phi, c^*)$. After this transformation, the vector of independent variables depends on ϕ via the individual means and the levels. Because of this, Gonzalez et al. (2005) suggest recomputing \widetilde{x}_{it} at each iteration of the nonlinear least squares (NLS) optimization.

In practical terms, this means estimating the linear components of the model using an ordinary least squares (OLS) estimation after each iteration of the NLS estimation. The NLS estimation determines the nonlinear components of the model, which are fed into the next OLS estimation. This process is repeated until the NLS models reaches the minimized concentrated sum of squared errors.

If transaction costs influence the width of the BofI, then ϕ can be thought of as an additively separable function of transaction costs

$$\phi_t = k + \sum_{i=1}^n \delta_i \tau_{i,t}^{-1} \quad (12)$$

Where τ_{it} , is a vector of all inputs to transaction costs, n is the number of inputs, k is a constant term representing all transaction costs not explicitly considered, and δ_i is a vector of coefficients measuring the relative strength of each input to transaction costs. $\tau_{i,t}$ enters in the denominator so that larger values create a wider BofI. In this analysis, transportation costs are the only input used, so ϕ_t can be restated in equation 13 below.

$$\phi_t = k + \delta_1 \tau_{1,t}^{-1} \quad (13)$$

Where $\tau_{1,t}$ represents transportation costs at time t , δ_1 describes the effect of $\tau_{1,t}$ on ϕ_t . The transition function then can be modified to reflect this relationship.

$$F(y_{t-d}, \tau_{1,t}) = \left\{ 1 - \exp \left[- \left(k + \frac{\delta_1}{\tau_{1,t}} \right) (y_{t-d} - c^*)^2 \right] \right\} \quad (14)$$

The combination of δ_1 , k , and $\tau_{1,t}$ describes the width of the BofI. If δ_1 is significant and positive, then transportation costs affect the BofI in the expected direction. Decreases in transportation costs narrow the BofI, while increases widen the BofI.

3.3 Data

Price level and exchange rate data were collected from the International Monetary Fund's (IMF) International Financial Statistics. Monthly data were collected for 14 developed countries and the

United States from 1974 to 2008¹⁰. Price level data are measured as CPI for all countries except the United Kingdom. The United Kingdom is measured as Retail Price Index (RPI) due to issues of data availability. Data for those countries that adopted the Euro ended when the Euro was created. The base country used for comparison was the United States. The year 2005 is the base year for all CPI and RPI values. Summary price level statistics for each country are provided below.

¹⁰ The countries are Canada, Denmark, Finland, France, Germany, Greece, Italy, Japan, Netherlands, Norway, Portugal, Sweden, Switzerland, and the United Kingdom.

Table 3.5 Summary CPI and RPI data from 1974 to 2010

Country	Obs	Mean	Std. Dev.	Min	Max
Canada	456	72.98164	25.58315	23.08	113.02
Japan	456	90.45294	14.36585	45.81	104.07
Switzerland	456	81.22252	18.65799	45.84	105.61
Denmark	456	73.17921	26.14986	21.61	114.77
Finland	456	75.04882	26.93796	19.04	114.5
France	456	74.70524	25.69974	21.17	111.34
Germany	456	79.32401	19.51121	42.66	111.9
Greece	456	53.23382	40.61243	2.4	122.82
Italy	456	65.39456	32.28258	8.62	114.41
Netherlands	456	76.28162	20.24781	35.13	111.45
Norway	456	71.62366	28.06583	20.02	114.12
Portugal	456	58.82384	37.0284	2.93	114.42
Sweden	456	72.74643	28.92979	18.82	112.26
U.K. (RPI)	456	68.67282	30.24321	13.1788	124.7037
U.S.	456	71.19081	26.23434	23.88	116.18

Data were collected from the International Monetary Fund's (IMF) International Financial Statistics (IFS).

3.3.1 Transportation Costs

The measurement of transportation costs has become more complicated since the 1970s. Because of advances in technology in air and ocean freight, the introduction of containerized shipping, the various oil price shocks since the 1970s, and changes in the goods shipped over time, changes in transportation costs are much more difficult to measure and observe.

Traditionally ad valorem measures have been used to describe transportation costs. Ad valorem, or “to value” in Latin, is the shipping price relative to the good’s value. These values usually were obtained using the “matched partner” technique. This method involves measuring the difference between reported trade flows from exporting and importing countries. Exporting countries report trade flows without including freight and insurance, while importing countries report trade flows with freight and insurance included. Assuming little or no measurement error, the difference between these values is the transportation cost for any matched pair of countries. However, Hummels and Lugovskyy (2006) demonstrate that this method is subject to many errors in measurement and reporting.

In addition to the issues raised by Hummels and Lugovskyy, several other issues are associated with the use of ad valorem values. These issues include changes in the type of goods shipped, the distance shipped, the quality of the goods, and the weight of the goods relative to the value. Over the past 40 years, the nature of international trade has changed in many ways. First, there has been a general shift in type of goods shipped from raw materials to manufactured goods. This shift has increased the value of goods and decreased the weight. Both of these shifts affect the ad valorem transportation cost measure, without providing information regarding the actual cost of shipping. The value of goods shipped also increased as the quality of imported goods increased over this time period. Finally, as transportation costs have decreased, countries have begun to shift the types of goods they ship and to whom. All of these effects work to obscure the true trend in transportation costs over the period of study (Hummels, 2007).

In addition to these shifts in the goods shipped, major changes have occurred to the method of transportation. The cost of air freight has been decreasing since the 1970s, caused by advances in jet engine and aerodynamics technology. This has caused the average weight of air freight to also

increase as goods that were previously too heavy to ship via air now became affordable. Therefore, the average weight of ocean freight increased as the relatively lighter freight moved to air freight. The increased average freight weight understates decreases in per ton transportation costs. To overcome the problems with traditional transportation cost measures, a separate transportation cost measure was generated following Hummels (2007).

Customs data were collected from the U.S. Imports of Merchandise, for the relevant time period. From this data, a fixed effects panel regression was estimated. The dependent variable was the ad valorem freight cost for commodity k shipped from exporter j at time t . The independent variables included year dummy variables and the ratio of the total weight of the commodity shipped by the value. The fixed effects (FE) are the commodity and exporting country combinations.

$$C_{k,j,t} = \alpha * \left(\frac{W_{k,j,t}}{V_{k,j,t}} \right) + \sum_{t=1973}^{2008} (\beta_t * year\ dummy) + FE_{j,k} + \mu \quad (15)$$

The weight/value ratio, $\frac{W_{k,j,t}}{V_{k,j,t}}$, captures shifts in the quality and composition of commodities over time. The fixed effects can capture variations in transportation costs due to differences in commodity weight and distance traveled.¹¹ This lets the year dummy coefficient, β_t , act as a measure of transportation costs over time (Hummels, 2007).¹² The regression included all years between 1974 and 2009. The year 2009 was the omitted dummy variable. The

¹¹ Use of a fixed effects model was confirmed using a Hausman test.

¹² This is a discrepancy in Hummels (2007): The text states the transportation cost measure above is the coefficient values for the dummy variables. Hummels' figures and Stata do-file indicate the use of the predicted values for equation 13, rather than the coefficient values. For this paper, the coefficient values are used, following the text.

regression was conducted for air freight. The year dummy coefficient values are used as a proxy for transportation cost in all future regressions.

Figure 3.3 Graph of generated transportation cost values over time.

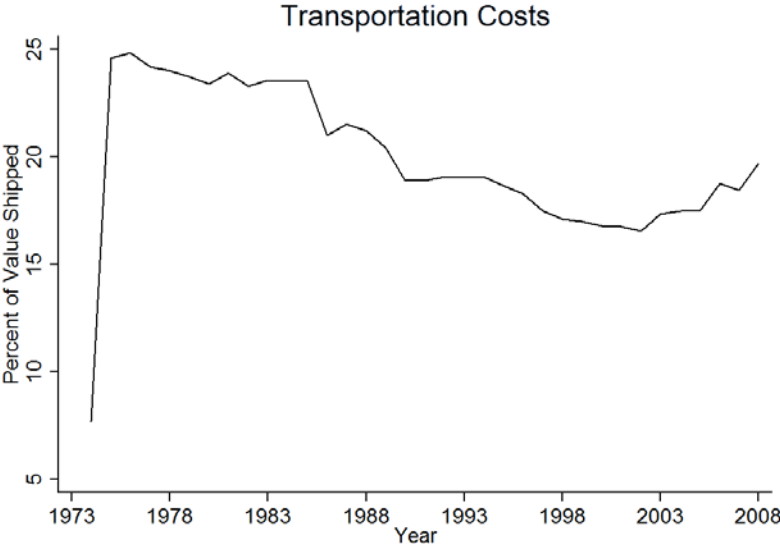


Table 3.6. Summary of Transportation Cost data

	Transportation Costs
Mean	19.910
Standard Error	3.485
Maximum	24.818
Minimum	7.700
Observations	35

From author's calculations using U.S. Census Bureau's U.S. Imports of Merchandise data.

As shown in the table and Figure 2 above, there was a significant jump in transportation costs in 1974, followed by a slower but still significant decline. From 1974 to 1975, transportation costs increased from 7.7 to 24.818, an increase of 222%. This increase could be related to OPEC's oil embargo of 1973 and 1974. From 1975 to 2002, transportation costs declined 33% from 24.818. This sustained decrease in transportation costs could have been caused by significant increases in technology (Hummels, 2007). From 2002 until 2008, transportation costs rose almost 20% from 16.523 to 19.691. This increase is most likely a result of the 2001 terrorist attack in the eastern United States. Following the attack, many new laws and policies were implemented regarding international goods traffic into the United States. The unusual increases in transportation costs in 1975 and after 2001, point to a need for greater research into the determinants of transportation costs. That, however, is , outside the scope of this chapter.

3.4 MP-ESTAR model of PPP and the BofI

The LM-type linearity tests recommended by Terasvirta (1994) show all countries except Greece rejected the null of linearity. Therefore, Greece was dropped from the sample. The delay parameter that minimized the P-values for all countries was 3 except for Finland. The delay parameter for Finland was 12.¹³ For this reason, Finland also was dropped from the sample. The series rejected the null hypothesis of homogeneity, allowing the remaining series to be combined into a panel STAR model.

For estimation, the data were separated into two groups. The first group includes all countries that retained their original currency throughout the study period, 1974 to 2008. The second group contained all countries from 1974 until 1998. This was done to include those that adopted the Euro. The separation was made to use all available data while still maintaining a balanced panel.

At this point, the first group, all countries that retained their original currency from 1974 to 2008, was separated into two subsamples for an exploratory regression and test of theory. The first 10 years and the last 10 years of the sample were regressed in a simple PSTAR model where ϕ was estimated as a simple parameter and not as a function of transaction costs. The preliminary regression was conducted using only the first group to maintain continuity.

There are three coefficient values of interest in each estimation, λ , λ^* , and ϕ . The lambdas (λ , λ^*) indicate if the series is mean reverting. ϕ describes the shape of the transition function. A negative and significant value for λ^* and $(\lambda + \lambda^*)$ demonstrate mean reverting behavior outside of the BofI. λ is allowed to be positive and significant, provided $(\lambda + \lambda^*)$ is negative and significant.

¹³ 12 was the highest reasonable delay parameter tested. It is possible that the true delay parameter for Finland is much higher.

The value of ϕ is expected to be positive and significant. Additionally, the value of ϕ should be larger in the last 10 years of the sample when transportation costs are lower.

Results from the preliminary regressions are shown below. The lambda values demonstrate the expected signs, indicating mean reversion behavior outside the BofI. It is surprising to find a significant value for λ , because $\Delta\widetilde{y}_{it}$ is expected to follow an apparent random walk. The positive values may represent a persistent force in at least one market, pushing price levels apart. After they pass beyond the BofI, the negative and significant values of λ^* indicate $\Delta\widetilde{y}_{it}$ returns to mean reverting behavior.

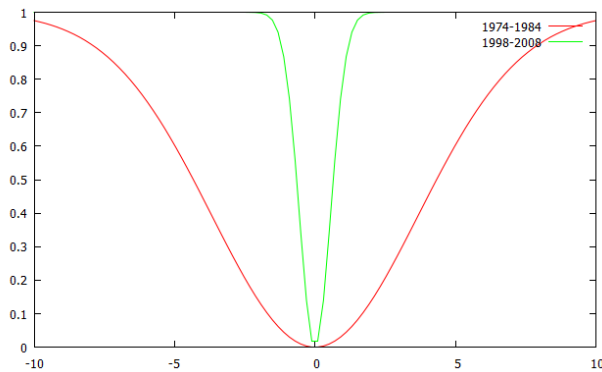
Table 3.7 Panel ESTAR regression of non-Euro countries in the first and last 10 years of the sample.

	Non-Euro countries 1974 to 1984		Non-Euro countries 1998 to 2008	
	Coefficient	Standard Error	Coefficient	Standard Error
Λ	0.0557*	0.0327	0.2481**	0.1260
λ^*	-0.0748**	0.0356	-0.2762**	0.1266
$(\lambda+\lambda^*)$	-0.0191*	0.0106	-0.0282***	0.0082
Φ	0.0370**	0.0145	1.6365**	0.7929

The first two columns of results include only the first 10 years of the sample. The second two columns include only the last 10 years of the sample. The characters*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively

The results above support the theory that the BofI narrows as transaction costs decrease. During the first subsample, the value of ϕ is 0.0370, compared to the higher ϕ value of 1.6364 in the last subsample. The value of ϕ increased almost 430% from the first 10 years to the last 10 years of the sample, while transportation costs decreased approximately 25% in the intervening years. This would indicate the BofI is sensitive to changes in transportation costs. A graph of the two transition functions provided below demonstrates the radical difference.

Figure 3.4 Transition Functions from the First and Last 10 Years of the Sample



The transition function for the first 10 years is in red. The transition function for the last 10 years is in green.

A more formal test is shown below in the regression of the Panel ESTAR model using the transition function from (14) for both the non-Euro countries from 1974 to 2008 and all countries from 1974 to 1998. As above, three coefficient values are of interest. The lambdas (λ , λ^*) still indicate if the series is mean reverting. However, instead of ϕ , δ_1 is included. δ_1 describes the shape of the transition function as transportation costs change. If δ_1 is positive and significant, then the BofI becomes larger as transportation costs increase and smaller as they decrease. This supports the theory that transportation costs are a determinant of the BofI. Results are provided below for both the non-Euro and all countries groups.

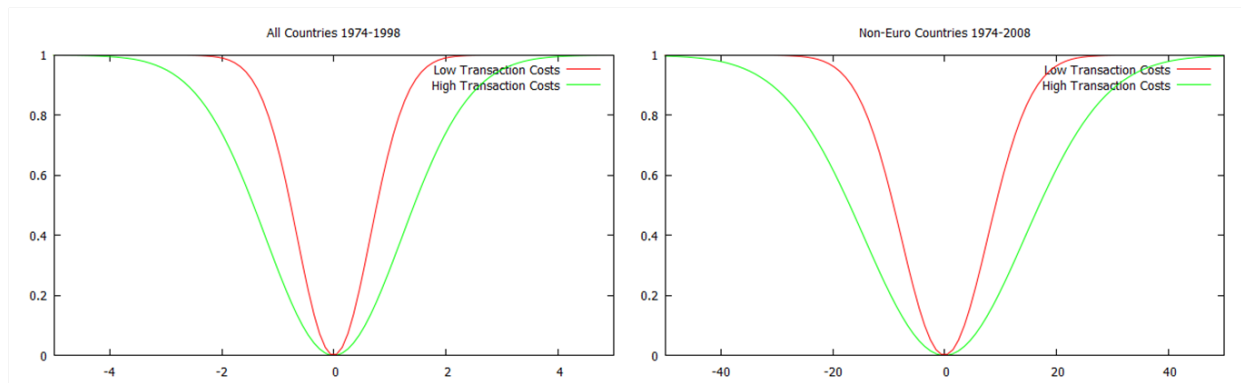
Table 3.8 Modified panel ESTAR regression results for Model 1 and Model 2

	Model 1 non-Euro countries 1974 to 2008		Model 2 all countries 1974 to 1998	
	Coefficient	Standard Error	Coefficient	Standard Error
Λ	-0.0047	0.0058	-0.0201***	0.0040
λ^*	-1.7047***	0.4610	-0.0208**	0.0093
$(\lambda+\lambda^*)$	-1.7094***	0.4574	-0.0408***	0.0083
δ_1	0.0007***	0.0002	0.0936*	0.0503

Model 1 includes non-Euro countries from 1974 to 2008. Model 2 includes all countries from 1974-1998. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

For both models, the λ coefficients are of the correct sign and significance. Both models show the presence of a BofI where mean reverting behavior occurs for large deviations from PPP but not for small deviations. The transition function provides some indication of what are large and small deviations. The transition functions graphed below show that the BofI is relatively narrow at low transportation costs (green line), though it widens considerably as transportation costs increase (red line).

Figure 3.5 Transition functions, $F(y(t-d), \tau(1,t))$ for Model 1 and Model 2



The transition functions with high transportation costs are in red. The transition functions with low transportation costs are in green.

There is a significant difference in the transition function between low and high transportation costs.

3.5 Conclusion

The failure of research to find consistent support for the PPP hypothesis can be explained by the BofI. The BofI is the range of price differentials where the cost of trade is greater than the benefit. Within the BofI, trade no longer occurs, and any trend to parity ceases. A great deal of empirical support can be found for the existence of the BofI, but no research exists into the determinants of the band (Heimonen, 2006; Michael, et al., 1997; Nakagawa, 2002; Pesaran, et al., 2009; Alan M. Taylor, 2001). This chapter modifies the traditional ESTAR model to confirm the connection between the BofI and transportation costs. The findings from the chapter provide support for one theoretical cause of the BofI. Transportation costs create the range of price differentials where the cost of trade is greater than the benefit. This paper extends the research by explaining the BofI rather than demonstrating its existence.

To construct a measure for transportation costs, this paper uses methods developed by Hummels (2007) and data from the U.S. Imports of Merchandise. This transportation cost data was used in a modified panel ESTAR model. The transition function within the ESTAR model was altered to allow the width of the transition function to vary with transportation costs. The transition function controls the movement from one autoregressive regime to another. It defines the border to the BofI. Using price level and exchange rate data for 14 countries in this modified ESTAR model, the results demonstrate that PPP exists outside the BofI and that decreases in transportation costs lead to a smaller or narrower BofI.

Additional work includes expanding the transportation cost frequency to monthly rather than to yearly data. This model also can be expanded to include other potential transaction costs. By including other transaction costs, such as tariffs, price uncertainty, and other nontariff trade barriers, it should be possible to evaluate the relative strength of various transaction costs in

determining the BofI. This in turn could lead to policy recommendations related to increasing international trade and integration.

Conclusion

Price convergence is a major assumption in many international trade models. Yet, in the real world the presence of price convergence is difficult to demonstrate empirically. Theoretically, in the absence of transaction costs, the price level in any two common-currency markets should be equal. Any differential should be eliminated through arbitrage. Price convergence studies show mixed results, and there is little consensus on the time scale of price convergence. Closely related to price convergence is the concept of purchasing power parity (PPP) which examines price convergence across currencies. PPP demonstrates mixed results across the literature as well.

Several theories have been put forward to explain the difficulty in detecting price convergence and PPP. These theories cite a variety of causes of difficulty from differing productivity growth rates (known as the Balassa-Samuelson effect) (Balassa, 1964), capital market imperfections, non-tariff trade barriers, time-varying transaction costs, general transactions costs, among many others (Hallwood & MacDonald, 2000; Alan M. Taylor, 2001). Virtual worlds (VWs) lack many of these sources of difficulty, making it easier to identify and study price convergence.

VWs represent a new tool for economic research. VWs allow millions of people from all over the world to interact, cooperate and trade on a daily basis. The economies created within VWs have become almost as complex as real world economies. This means many of the economic phenomena found in real world are also in VW economies. Because VWs are programmed, they lack some of the unknowns present in the real world (i.e. unknown transaction costs, non-tariff trade barriers, etc). This makes it easier to isolate and test certain behaviors and phenomena, such as price convergence.

The first chapter provided a brief overview of the social science literature using VWs, and a discussion of the general demographic characteristics of VW users. The non-economic research was grouped into four groups, motivation, identity, behavior, and social networks and culture, which represent the majority of VW research outside of economics. This research draws from different fields including, anthropology, communications, sociology and computer science.

While researchers have varied backgrounds, the themes of their research are consistent within each category. The research outside of economics supports the experimental economics research using VWs. The literature finds that VW users behave in a manner consistent with economic theory. VWs are a useful tool for understanding collective and individual human behavior.

The second chapter tests for price convergence using the VW, World of Warcraft (WoW). WoW is suited to price convergence research given its two country set-up with a single currency. There is a domestic goods market in each country as well as a single international goods market. This simple set up allows for calculation of high frequency price level and transaction cost data. In addition to this advantage, the developers of WoW have created 200+ parallel copies of their VW. The 12 million users world-wide are spread relatively evenly across these copies, called servers (Blizzard Entertainment, 2010). While these copies were made to reduce the computational load, it also allows for observation of shocks in multiple economies. This is much like a biologist using several Petri dishes in one study.

This study examined price convergence after a major shock to the WoW economy. The shock was caused by an expansion to WoW. This expansion introduced many new goods as well as new uses for old goods. This expansion was released for all 200+ WoW servers.

Price convergence and PPP equilibrium is assumed when prices enter the BofI (BofI). The BofI is the range of price differentials where the cost of trade is greater than the benefit. When inside the BofI there is no available profit from arbitrage, and therefore no more convergence pressure. Within the BofI price levels are determined by individual market forces and for convenience the price differential can be considered to follow a random walk.

Of the 200+ WoW servers eight were randomly chosen for this study. The price convergence tests used in this chapter were the Engle-Granger cointegration test (Engle & Granger, 1987), the exponential smooth transition autoregressive (ESTAR) test (Terasvirta, 1994), the Phillips and Sul log t convergence test (Phillips & Sul, 2007) and the BofI test (Morrison & Fontenla, 2012). The BofI test used in the first chapter was an expanded versions of the test developed in Morrison and Fontenla (2012). Price convergence was found using the Engle-Granger and BofI tests, but not when using the ESTAR and Phillip/Sul tests.

The mixed results from this chapter highlight one potential problem in testing for price convergence. The Engle-Granger and BofI tests both test for converged price series. The ESTAR and Phillips/Sul test look for converging price series. In other words the Engle-Granger and BofI test look to see if the series are at equilibrium, while the other tests look to see if the series returns to equilibrium after a shock.

Following the expansion to WoW, price levels quickly returned to the BofI, leaving few observations outside the BofI where converging behavior could be observed. The low number of observations outside the BofI could explain the failure of the ESTAR and Phillip/Sul model to find price convergence. Conversely, the large number of observations inside the BofI helped both the Engle-Granger and BofI test find price convergence. This implies that the speed of

convergence and nature of the markets need to be considered when choosing a price convergence test.

In addition to testing for price convergence, the BofI was expanded and made more robust in the second chapter. The BofI test requires the measurement of the costs and benefits of trade. The BofI test uses a one-sided t-test on the costs and benefits of trade. If the costs are consistently higher than the benefits, then the price series are within the BofI and at equilibrium. Originally developed in Morrison and Fontenla (2012), the BofI test required many simplifying assumptions be made due to low quality data. However, the high quality data from chapter two eliminated the need for simplifying assumptions. This increase in the quality and quantity of data created a more reliable price convergence test by eliminating unknowns.

The major issue preventing the use of the BofI test outside of VWs is the difficulty associated with measuring the costs of trade also called transaction costs. Without accurate measures of the transaction costs, the BofI cannot be measured until convergence has already been established. The third chapter developed a method to measure transaction costs using data from 14 countries. The chapter used a modified ESTAR model to measure the effect of transportation costs on the size of the BofI. This is a first step towards developing a practical method for accurate measurement of the BofI.

The basic ESTAR model estimates two autoregressive equations for two separate observed regimes as well as a transition function. The transition function dictates which regime is dominant. In the context of price convergence and PPP studies the two regimes are inside and outside the BofI. The autoregressive variable is the linear combination of two price levels and their nominal exchange rate.

The regime within the BofI would be expected to follow an (apparently) random walk as each price series responds to their own market forces. Outside the BofI each price series is expected to demonstrate converging behavior as price differentials are exploited by arbitrageurs.

The transition function describes the switching point between regimes, or BofI edges. This is modified by making the BofI edges a function of transportation costs. This allows the width of the BofI to vary in response to changes in transportation costs. Transportation costs were chosen as a proxy for transaction costs because they are believed to be a major component of transaction costs.

Using the price level, exchange rate, and derived transportation cost data the modified ESTAR model found that BofI does vary dramatically with transportation costs. This suggests that transportation costs are a major component of the transaction costs, and therefore a major determinant of the BofI.

Together these three chapters demonstrate the potential of VW research in economics. The first chapter introduces opportunities for future economics research, by summarizing research from other disciplines. The second and third chapters combine to show how using VWs can provide new perspective into old problems. Use of data from VWs led to the development of the BofI test, and later its expansion. Methods developed from the use of VW also provided inspiration for a new way to measure the BofI in the real world. Use of methods developed in the analysis of VWs can lead to new lines of research. They will contribute to the understanding of price convergence, PPP, and many other areas of economics.

References

- Atlas, S., & Putterman, L. (2011). Trust among the avatars: A virtual world experiment, with and without textual and visual cues. *Southern Economic Journal*, 78(1), 63-86.
- Atlas, S. A. (2008a). *Economic experiments in virtual worlds: Framing, reciprocity & trust*: ProQuest.
- Atlas, S. A. (2008b). *Inductive Metanomics: Economic Experiments in Virtual Worlds*.
- Bahmani-Oskooee, M., Kutun, A. M., & Zhou, S. (2007). Testing PPP in the non-linear STAR framework. [doi: 10.1016/j.econlet.2006.08.008]. *Economics Letters*, 94(1), 104-110.
- Bainbridge, W. (2007). The Scientific Research Potential of Virtual Worlds. *Science*, 317(5837), 472-476.
- Balassa, B. (1964). The Purchasing-Power Parity Doctrine: A Reappraisal. *The Journal of Political Economy*, 72(6), 584-596.
- Baum, C. F., Barkoulas, J. T., & Caglayan, M. (2001). Nonlinear Adjustment to Purchasing Power Parity in the Post-Bretton Woods Era. *Journal of International Money and Finance*, 20(3), 379-399.
- Bessière, K., Seay, A. F., & Kiesler, S. (2007). The ideal elf: Identity exploration in World of Warcraft. *CyberPsychology & Behavior*, 10(4), 530-535.
- Blizzard Entertainment. (2010). World of Warcraft Subscriber Base Reaches 12 Million Worldwide. Retrieved October 7th, 2010, from <http://us.blizzard.com/en-us/company/press/pressreleases.html?id=2847881>
- Bloomfield, R. (2008). Blast from the Past: Metanomics 101, Revisited. *Metanomics 2008–Business and Policy in the Metaverse*.
- Bloomfield, R., & Rennekamp, K. (2009). From the Laboratory to the Virtual World: Applications for Financial Reporting Research. *Foundations and Trends in Accounting Research, Forthcoming*.
- Bloomfield, R. J., & Cho, Y. J. (2011). Unregulated stock markets in Second Life. *Southern Economic Journal*, 78(1), 6-29.
- Castronova, E. (2001). Virtual Worlds: A First-Hand Account of Market and Society on the Cyberian Frontier. CESifo GmbH.
- Castronova, E. (2005). On the Research Value of Large Games: Natural Experiments in Norrath and Camelot. *SSRN eLibrary*.
- Castronova, E. (2008). A Test of the Law of Demand in a Virtual World: Exploring the Petri Dish Approach to Social Science. CESifo GmbH.
- Chandra, Y., & Leenders, M. A. (2012). User innovation and entrepreneurship in the virtual world: A study of Second Life residents. *Technovation*.
- Chesney, T., Chuah, S.-H., & Hoffmann, R. (2009). Virtual World Experimentation: An Exploratory Study. *Journal of Economic Behavior and Organization*, 72(1), 618-635.
- Chesney, T., Coyne, I., Logan, B., & Madden, N. (2009). Griefing in virtual worlds: causes, casualties and coping strategies. *Information Systems Journal*, 19(6), 525-548.
- Chong, Y., Jordà, Ò., & Taylor, A. M. (2012). THE HARROD–BALASSA–SAMUELSON HYPOTHESIS: REAL EXCHANGE RATES AND THEIR LONG-RUN EQUILIBRIUM*. *International Economic Review*, 53(2), 609-634.
- de Carvalho, M., & Júlio, P. (2012). Digging out the PPP hypothesis: an integrated empirical coverage. *Empirical Economics*, 42(3), 713-744.

- Debeauvais, T., Nardi, B., Schiano, D. J., Ducheneaut, N., & Yee, N. (2011). *If you build it they might stay: Retention mechanisms in World of Warcraft*. Paper presented at the Proceedings of the 6th International Conference on Foundations of Digital Games.
- Debeauvais, T., Nardi, B. A., Lopes, C. V., Yee, N., & Ducheneaut, N. (2012). *10,000 Gold for 20 Dollars: An exploratory study of World of Warcraft gold buyers*. Paper presented at the Proceedings of the International Conference on the Foundations of Digital Games.
- Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.
- Ducheneaut, N., & Moore, R. J. (2004). *The social side of gaming: a study of interaction patterns in a massively multiplayer online game*. Paper presented at the Proceedings of the 2004 ACM conference on Computer supported cooperative work.
- Ducheneaut, N., Moore, R. J., & Nickell, E. (2004). Designing for sociability in massively multiplayer games: an examination of the “third places” of SWG. *Proc. of the Other Players Conference*.
- Ducheneaut, N., & Yee, N. (2008). Collective solitude and social networks in world of warcraft. *Social networking communities and e-dating services: Concepts and implications*, 78-100.
- Ducheneaut, N., Yee, N., & Bellotti, V. (2010). The Best of Both (Virtual) Worlds: Using Ethnography and Computational Tools to Study Online Behavior. *Ethnographic Praxis in Industry Conference Proceedings*, 2010(1), 136-148.
- Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2006). Building an MMO with mass appeal a look at gameplay in world of warcraft. *Games and Culture*, 1(4), 281-317.
- Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2007). *The life and death of online gaming communities: a look at guilds in world of warcraft*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Duffy, J. (2011). Trust in second life. *Southern Economic Journal*, 78(1), 53-62.
- Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276.
- Entertainment Software Association. (2012). Essential Facts about the Computer and Video Game Industry 2007. *Press Release.[Online]*
http://www.theesa.com/facts/pdfs/ESA_EF_2012.pdf.
- Fan, C. S., & Wei, X. (2006). The Law of One Price: Evidence from the Transitional Economy of China. *Review of Economics and Statistics*, 88(4), 682-697.
- Fetscherin, M., & Lattemann, C. (2008). User acceptance of virtual worlds. *Journal of Electronic Commerce Research*, 9(3), 231-242.
- Fiedler, M. (2011). Experience and confidence in an internet-based asset market experiment. *Southern Economic Journal*, 78(1), 30-52.
- Fiedler, M., & Haruvy, E. (2009). The lab versus the virtual lab and virtual field--An experimental investigation of trust games with communication. *Journal of Economic Behavior & Organization*, 72(2), 716-724.
- Fiedler, M., Haruvy, E., & Li, S. X. (2011). Social distance in a virtual world experiment. *Games and Economic Behavior*, 72(2), 400-426.
- Fischer, C. (2009). *Price Convergence in the EMU? Evidence from Micro Data* (Discussion Paper No 06/2009). Frankfurt: Deutsche Bundesbank, Research Centre.
- Froot, K. A., & Rogoff, K. (1996). Perspectives on PPP and Long-Run Real Exchange Rates. *NBER Working Paper:4952*, 1-57.

- Fuller, W. A. (1976). *Introduction to statistical time series* (Vol. 428): Wiley-Interscience.
- Funke, K., & Koske, I. (2008). Does the Law of One Price Hold within the EU? A Panel Analysis. *International Advances in Economic Research*, 14(1), 11-24.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, 7, 7-28.
- Golub, A. (2010). Being in the World (of Warcraft): Raiding, Realism, and Knowledge Production in a Massively Multiplayer Online Game. [Article]. *Anthropological Quarterly*, 83(1), 17-45.
- Gonzalez, A., Terasvirta, T., & van Dijk, D. (2005). *Panel Smooth Transition Regression Models* (No. 165).
- Granger, C. W. J., & Terasvirta, T. (1993). *Modelling Non-Linear Economic Relationships*. Oxford: Oxford University Press.
- Greene, W. H. (2003). *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.
- Guangying, H., & Haughton, D. (2008, June 30 2008-July 2 2008). *An investigation of virtual worlds adoption: A research framework and empirical study*. Paper presented at the Service Systems and Service Management, 2008 International Conference on.
- Guo, J., & Gong, Z. (2011). Measuring virtual wealth in virtual worlds. *Information Technology and Management*, 12(2), 121-135.
- Hallwood, C. P., & MacDonald, R. (2000). *International Money and Finance*. Malden, MA: Blackwell Publishing.
- Harrison, G. W., Haruvy, E., & Rutström, E. E. (2011). Remarks on Virtual World and Virtual Reality Experiments. *Southern Economic Journal*, 78(1), 87-94.
- Haruvy, E. (2011). Challenges and Opportunities in Economics Experiments in Virtual Worlds. *Southern Economic Journal*, 78(1), 1-5.
- Hayes, A., Ulinski, A., & Hodges, L. (2010). That Avatar Is Looking at Me! Social Inhibition in Virtual Worlds. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud & A. Safonova (Eds.), *Intelligent Virtual Agents* (Vol. 6356, pp. 454-467): Springer Berlin Heidelberg.
- Heimonen, K. (2006). Nonlinear adjustment in PPP—evidence from threshold cointegration. *Empirical Economics*, 31(2), 479-495.
- Hummels, D. (2007). Transportation Costs and International Trade in the Second Era of Globalization. *Journal of Economic Perspectives*, 21(3), 131-154.
- Hummels, D., & Lugovskyy, V. (2006). Are Matched Partner Trade Statistics a Usable Measure of Transportation Costs? *Review of International Economics*, 14(1), 69-86.
- Jacobson, T., & Nessén, M. (2004). Examining world-wide purchasing power parity. *Empirical Economics*, 29(3), 463-476.
- Johnson, N. F., Xu, C., Zhao, Z., Ducheneaut, N., Yee, N., Tita, G., et al. (2009). Human group formation in online guilds and offline gangs driven by a common team dynamic. *Physical Review E*, 79(6), 066117.
- Kim, G., & Kim, Y. (2011). Purchasing Power Parity and Half Life: Another Look.
- Kobayashi, T. (2010). Bridging Social Capital in Online Communities: Heterogeneity and Social Tolerance of Online Game Players in Japan. *Human Communication Research*, 36(4), 546-569.
- Krasonikolakis, I. G., Vrechopoulos, A. P., & Pouloudi, A. (2010). *Users' personality traits in the context of virtual reality*. Paper presented at the e-Business (ICE-B), Proceedings of the 2010 International Conference on.

- Linden Labs. (2008). SL Key Metrics. *Press Release.*[Online] s3.amazonaws.com/static-secondlife-com/economy/stats_200811.xls.
- Lortie, C. L., & Guitton, M. J. (2011). Social organization in virtual settings depends on proximity to human visual aspect. *Computers in Human Behavior*, 27(3), 1258-1261.
- Lyons, K., Playford, C., Messinger, P., Niu, R., & Stroulia, E. (2009). Business Models in Emerging Online Services. In M. Nelson, M. Shaw & T. Strader (Eds.), *Value Creation in E-Business Management* (Vol. 36, pp. 44-55): Springer Berlin Heidelberg.
- MacKinnon, J. G. (2010). *Critical values for cointegration tests*: Queen's Economics Department Working Paper.
- Malaby, T. (2006). Parlaying Value Capital in and Beyond Virtual Worlds. *Games and Culture*, 1(2), 141-162.
- Mäntymäki, M., & Merikivi, J. (2010). *Investigating the drivers of the continuous use of social virtual worlds*. Paper presented at the System Sciences (HICSS), 2010 43rd Hawaii International Conference on.
- Mäntymäki, M., & Salo, J. (2010). Trust, social presence and customer loyalty in social virtual worlds. *23rd Bled eConference eTrust: Implications for the Individual, Enterprises and Society*.
- Mäntymäki, M., & Salo, J. (2011). Continuous use and Purchasing Behavior in Social Virtual Worlds. *Review Process, Computers in Human Behavior*.
- Mäntymäki, M., & Salo, J. (2011). Teenagers in social virtual worlds: Continuous use and purchasing behavior in Habbo Hotel. *Computers in Human Behavior*, 27(6), 2088-2097.
- Michael, P., Nobay, A. R., & Peel, D. A. (1997). Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation. *Journal of Political Economy*, 105(4), 862-879.
- Moore, R. J., Ducheneaut, N., & Nickell, E. (2007). Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds. *Computer Supported Cooperative Work (CSCW)*, 16(3), 265-305.
- Moore, R. J., Gathman, E., Ducheneaut, N., & Nickell, E. (2007). *Coordinating joint activity in avatar-mediated interaction*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Morrison, M., & Fontenla, M. (2012). Price convergence in an online virtual world. *Empirical Economics*, 1-12.
- Nagy, P. (2010). *Second Life, Second Choice? The effects of virtual identity on consumer behavior. A conceptual framework*. Paper presented at the Proceedings of FIKUSZ, 10th Symposium for Young Researchers.
- Nakagawa, H. (2002). Real Exchange Rates and Real Interest Differentials: Implications of Nonlinear Adjustment in Real Exchange Rates. *Journal of Monetary Economics*, 49(3), 629-649.
- Nejib, H. (2008). the Purchasing Power Parity and the Symmetry, Proportionality Conditions: Panel Cointegration Evidence from Some African Countries. *International Research Journal of Finance and Economics*, 16, 121-136.
- Nicklisch, A., & Salz, T. (2008). Reciprocity and status in a virtual field experiment. *MPI Collective Goods Preprint*(2008/37).
- Pesaran, M. H., Smith, R. P., Yamagata, T., & Hvozdnyk, L. (2009). Pairwise Tests of Purchasing Power Parity. *Econometric Reviews*, 28(6), 495-521.

- Phillips, P. C. B., & Sul, D. (2007). Transition Modeling and Econometric Convergence Tests. *Econometrica*, 75(6), 1771-1855.
- Reeves, B., Malone, T. W., & O Driscoll, T. (2008). Leadership's online labs. *Harvard Business Review*, 86(5), 58.
- Shen, J., Brdiczka, O., Ducheneaut, N., Yee, N., & Begole, B. (2012). Inferring personality of online gamers by fusing multiple-view predictions. *User Modeling, Adaptation, and Personalization*, 261-273.
- Shin, D. H. (2009). The evaluation of user experience of the virtual world in relation to extrinsic and intrinsic motivation. *Intl. Journal of Human-Computer Interaction*, 25(6), 530-553.
- Taylor, A. M. (2001). Potential Pitfalls for the Purchasing-Power-Parity Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price. *Econometrica*, 69(2), 473-498.
- Taylor, A. M. (2002). A century of purchasing-power parity. *Review of Economics and Statistics*, 84(1), 139-150.
- Terasvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. *Journal of the American Statistical Association*, 89(425), 208-218.
- Thurner, S., Szell, M., & Sinatra, R. (2012). Emergence of Good Conduct, Scaling and Zipf Laws in Human Behavioral Sequences in an Online World. [doi:10.1371/journal.pone.0029796]. *PLoS ONE*, 7(1), e29796.
- Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association*, 84(405), 231-240.
- Verhagen, T., Feldberg, F., van den Hooff, B., Meents, S., & Merikivi, J. (2012). Understanding users' motivations to engage in virtual worlds: A multipurpose model and empirical testing. *Computers in Human Behavior*, 28(2), 484-495.
- Wadley, G., & Ducheneaut, N. (2009). The 'out-of-avatar experience': object-focused collaboration in Second Life. *ECSCW 2009*, 323-342.
- Wadley, G., Gibbs, M. R., & Ducheneaut, N. (2009). *You can be too rich: Mediated communication in a virtual world*. Paper presented at the Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7.
- Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., & Nickell, E. (2006). From Tree House to Barracks The Social Life of Guilds in World of Warcraft. *Games and Culture*, 1(4), 338-361.
- Williams, D., Yee, N., & Caplan, S. (2008). Who plays, how much, and why? A behavioral player census of a virtual world. *Journal of Computer Mediated Communication*, 13(4), 993-1018.
- Wolszczak-Derlacz, J. (2008). Price convergence in the EU - an aggregate and disaggregate approach. *International Economics and Economic Policy*, 5(1), 25-47.
- Working, H. (1960). Note on the Correlation of First Differences of Averages in a Random Chain. *Econometrica*, 28(4), 916-918.
- Yee, N. (2006a). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and virtual environments*, 15(3), 309-329.
- Yee, N. (2006b). Motivations for play in online games. *CyberPsychology & Behavior*, 9(6), 772-775.

- Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009a). The Proteus Effect. *Communication Research*, 36(2), 285-312.
- Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009b). The Proteus effect implications of transformed digital self-representation on online and offline behavior. *Communication Research*, 36(2), 285-312.
- Yee, N., Ducheneaut, N., & Nelson, L. (2012). *Online gaming motivations scale: development and validation*. Paper presented at the Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems.
- Yee, N., Ducheneaut, N., Nelson, L., & Likarish, P. (2011). *Introverted elves & conscientious gnomes: the expression of personality in world of warcraft*. Paper presented at the Proceedings of the 2011 annual conference on Human factors in computing systems.
- Yee, N., Ducheneaut, N., Shiao, H. T., & Nelson, L. (2012). *Through the azerothian looking glass: mapping in-game preferences to real world demographics*. Paper presented at the Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems.
- Yee, N., Ducheneaut, N., Yao, M., & Nelson, L. (2011). *Do men heal more when in drag?: Conflicting identity cues between user and avatar*. Paper presented at the Proceedings of the 2011 annual conference on Human factors in computing systems.

Appendix A

The raw data was collected from the online service, The Armory. This service is offered by Blizzard Inc., the publishers of World of Warcraft. Among other things, The Armory provides a real-time listing of all posted auctions in all countries on all worlds. For each auction, the name of the good on sale, the quantity of the good offered for sale, the name of the seller, the current or minimum bid, and the buyout price (maximum bid) are provided. The major drawback to this service is that no historical records are posted; only the current auctions are shown.

A website, www.underminejournal.com, scans each auction house at least once an hour and provides a historical record of all auctions posted in each server. The developer of the website was employed to track nine specific goods across eight worlds, in all auction houses. I received a daily e-mail containing one file for each good in each country in each server for each hour of the day. In total, each daily email included 3,456 data files. Each file contained a listing of all auctions observed and the time of the scan. For each auction, all available information was included (name of good, quantity, seller's name, minimum bid, and buyout price).

The data collection began on December 4, 2010, three days before the release of the expansion Cataclysm on December 7, 2010. Cataclysm added several new items, new territory, and modified existing territory, as well as made several changes to the game play. This represented a large shock that would affect all worlds and countries. The magnitude of the shock would vary based on the individual characteristics of each country and world.

Sales data were extracted from the auction files by looking for auctions that disappeared from one scan to the next. This was accomplished by combining all 3,456 files. First, hourly files for each item in each country in each world were combined into daily/item/country/world files. At

this point, a variable was added to indicate the date and time of the scan; all price information was adjusted to show a per-unit price. Henceforth, listed prices are per-unit price. Next, each daily/ item/country/world file was combined into item-country/world files.

At this stage, it was necessary to extract the market price in each time period and drop all other observations. To get the market price, a variable was created for auction batches. The batches were defined based on the good, the quantity offered, the name of the seller, and the buyout price. A time variable was generated using the date and time information collected at the first condensation stage. The data then was sorted by batch and date. At this point, a variable was created counting the number of auctions in each batch in each time period. The batch variable also was altered to include the initial start time of the batch.

Then a variable was generated subtracting the number of auctions in a batch in a time period from the number of auctions in a batch in the previous time period. If this number was positive, then a new auction was added to the batch. If the number was negative, then an auction disappeared. When a new auction showed up, the batch ID for the new auction was changed, by adding the letter “n” to the beginning of the batch name. This ensured that all new auctions had a unique ID and that all batches included only those auctions posted at the same time.

In addition, this solved another problem: the fact that the same seller often would post an auction for the same item at the same bid and buyout prices several times throughout the data, i.e., once in January and once in April. The process of giving a new batch ID for each auction added to the batch solved this problem.

When an auction disappeared, it was necessary to determine if the auction had expired or if the auction had sold (i.e., the buyout price was paid). To do this, an additional variable was

generated for each batch ID. This new variable (z) counted how long the auction had been offered. Auctions can expire only after 12, 24, or 48 hours. If an auction disappeared after 12, 24, or 48 hours, then it was assumed the auction expired. If it disappeared at any other point, it was assumed the auction was sold. The sold auctions were indicated and used in generating the time period's market price.

Indicators were created to show if the auction expired and if it expired in 12, 24 or 48 hours. From these indicators, the variables ρ , and Θ were generated for final use. To create ρ , first two other variables were generated. First, ρ_1 , counted the total number of auctions. ρ_2 was created counting the total number of auctions expired. ρ was created by subtracting the ratio of ρ_2 to ρ_1 subtracted from 1.

Next, Θ s were derived first by creating several new count variables. The new variables counted how many auctions per item expired in 12, 24, and 48 hours individually. Then each new count variable was divided by ρ_2 . This generated the Θ s. The Θ s indicate the fraction of auctions per item that expired in 12, 24, or 48 hours.

At this point, all expired auctions were dropped from the data. The Θ and ρ were constants over the entire time frame and therefore were preserved with the sold auctions. The sold auctions then were sorted by date and time sold. The maximum buyout prices observed in each time period were recorded as the market price for that item in that time period.

Along the way, variables indicating the total number of items sold, total number of auctions offered, and the number of sellers per time period were generated and retained in a manner similar to that of the price data. Next, all but one observation per time period was dropped from the data. This means there was only one observation per time period in each item/country/world

file. The file was designated as a time series in Stata, at this point. Time was measured in increments of one minute, as the scans were not always exactly one hour apart.

The item files then were combined to create one file for each country in each world. Finally, the countries were combined into one file so there were a total of eight files, one for each world. Each file included both countries and all goods for the entire time period.

The next stage of the data cleaning involved aggregating the data up to hourly observations and dropping any outliers; this was done first by generating an hourly variable from the original time variable. Then 90th percentile values were generated for each price series in each country in each server. Any price greater than twice the 90th percentile was dropped. MSV values for each item were input as a constant.

Next, an hourly median price was generated for each item. This hourly median price replaced the earlier prices. Hourly median values were created for the Θ s, ρ s, and the various volume measures. All other price observations were dropped, so that there was only one observation per item per country per hour. The process was repeated at the daily level so that there was only one observation per item per country per day in each world file.

After a daily file was created, three price indices were created. The first index, *pre*, included only those items that existed before Cataclysm; the second index, *post*, included only new items introduced with Cataclysm. The final index, *index*, included all items. The indices were created as an average of all included goods weighted by the item sales volume as a fraction of total index volume. Price, Θ , ρ , and MSV values were generated for each index.

In addition to generating the indices, the variables necessary for the Phillips and Sul Log-t Convergence Test were generated for each item and index. This included total volume, cross-

sectional variance in each time period, as well as a separate constant variable indicating the first observed cross-sectional variance.

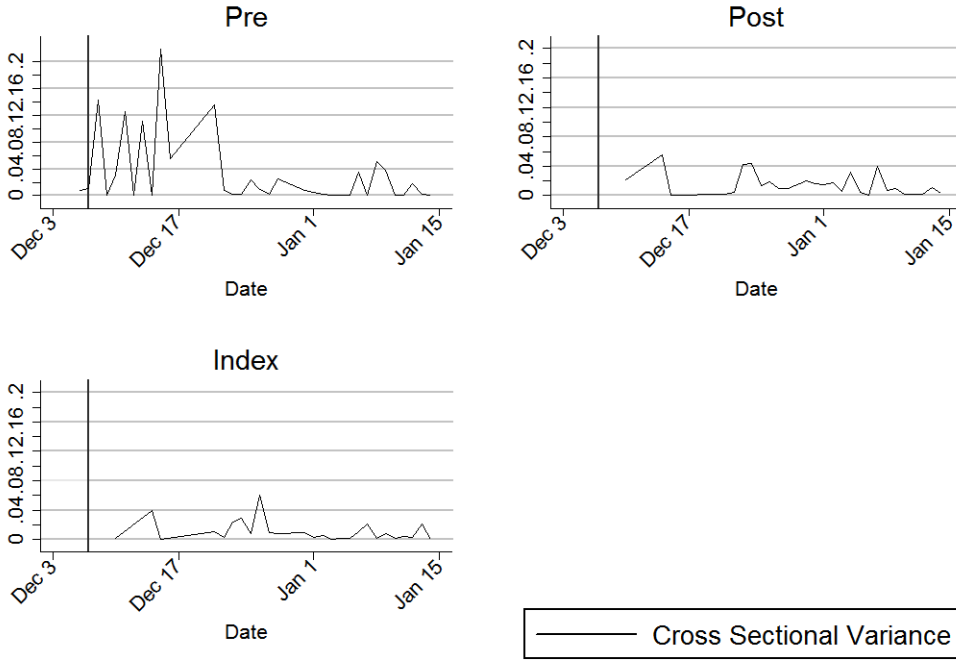
Next, the price, Θ , ρ , and MSV variables were separated into high and low categories based on the countries' goods price relative to the other country. This means if the Alliance price level for the pre index was higher than the Horde's price level for the pre index, then the Alliance price on that day was designated as the high price level for pre. In that same time period, the Horde's price level for the pre index was designated as the low price for the pre index. All Θ ρ and MSV values were designated as high or low based on the designation of the price. Next, variables for the net benefits of arbitrage were generated.

Finally, the convergence condition was generated. The conditions were formulated as the net benefit from arbitrage minus the lower price. After the conditions were generated, a simple t-test (condition = 0) was conducted on each item and index in each world.

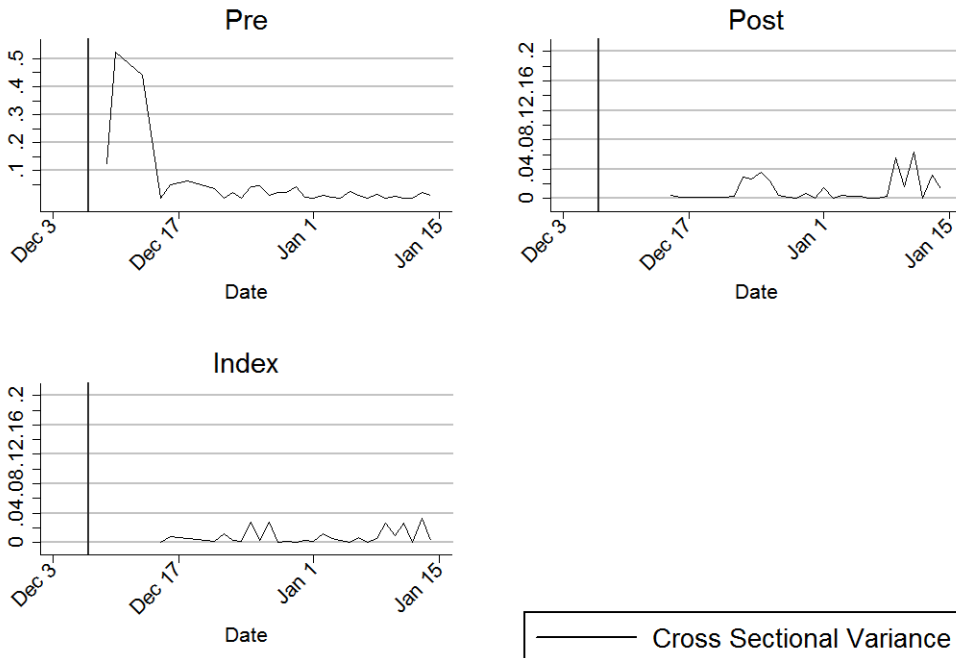
During data collection, some worlds failed to collect data due to technical difficulties. The higher than average level of users accessing The Armory during the holiday season prevented the automated data collection mechanism from completing scans of all auctions on schedule. This created some holes in the data; however, overlap of times between all worlds shows similar results despite the missing data. It is believed this issue does not significantly affect the results.

Appendix B

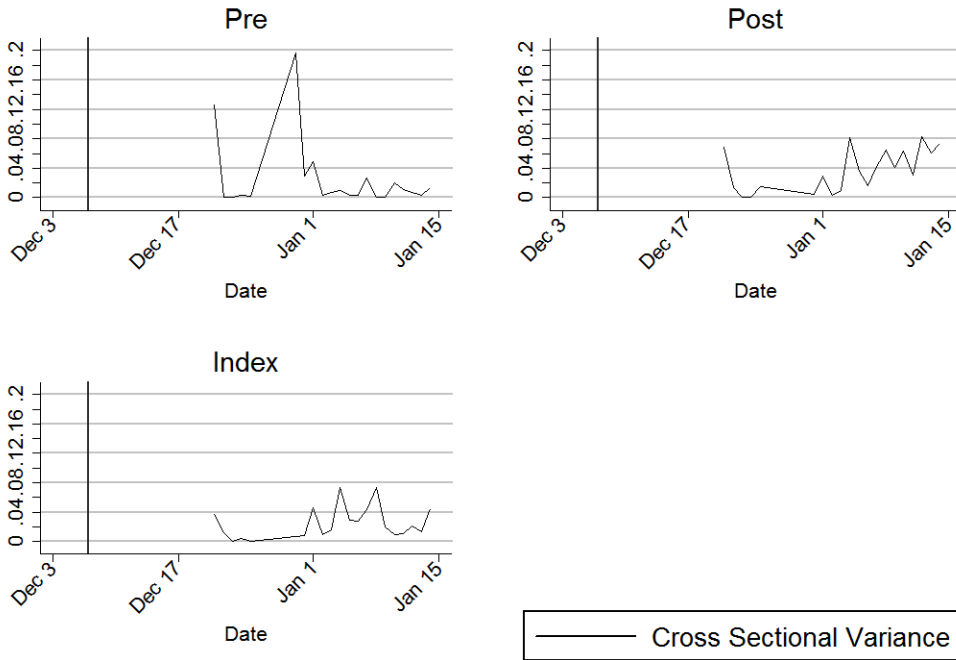
Azjol-Nerub - 1



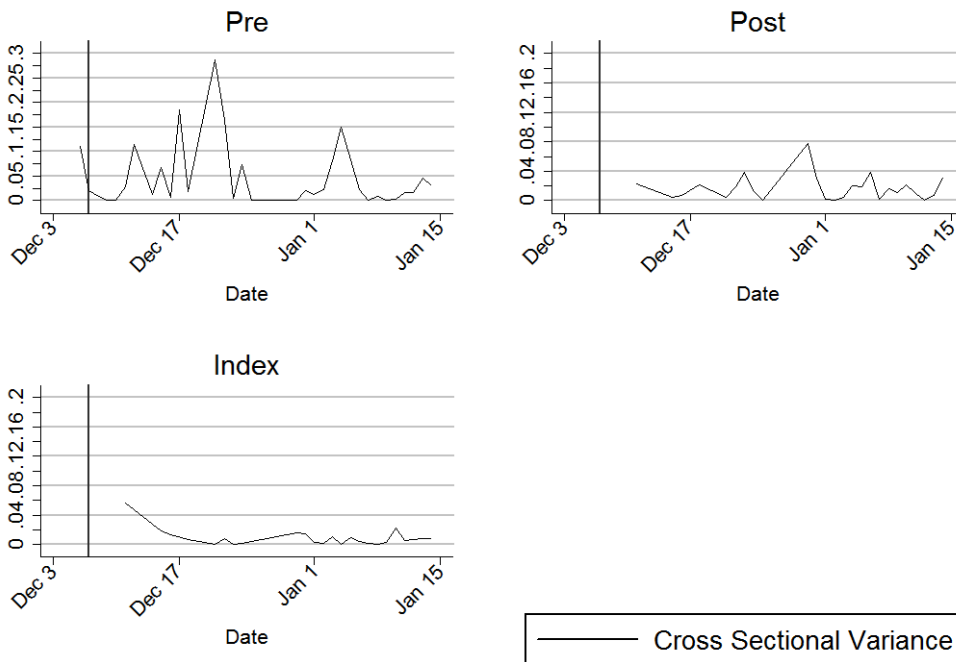
Draka - 2



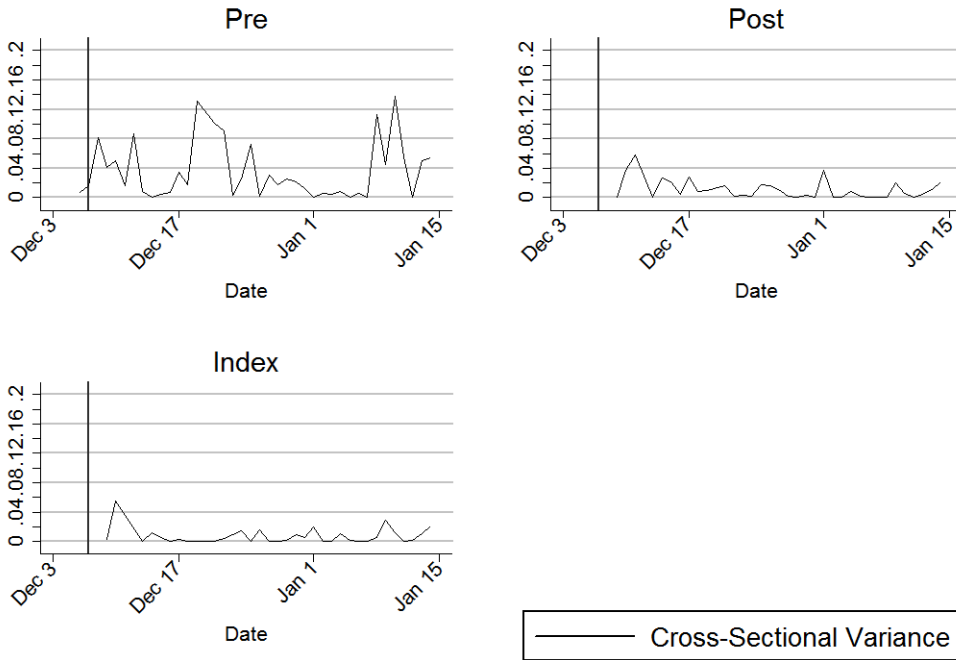
EarthenRing - 3



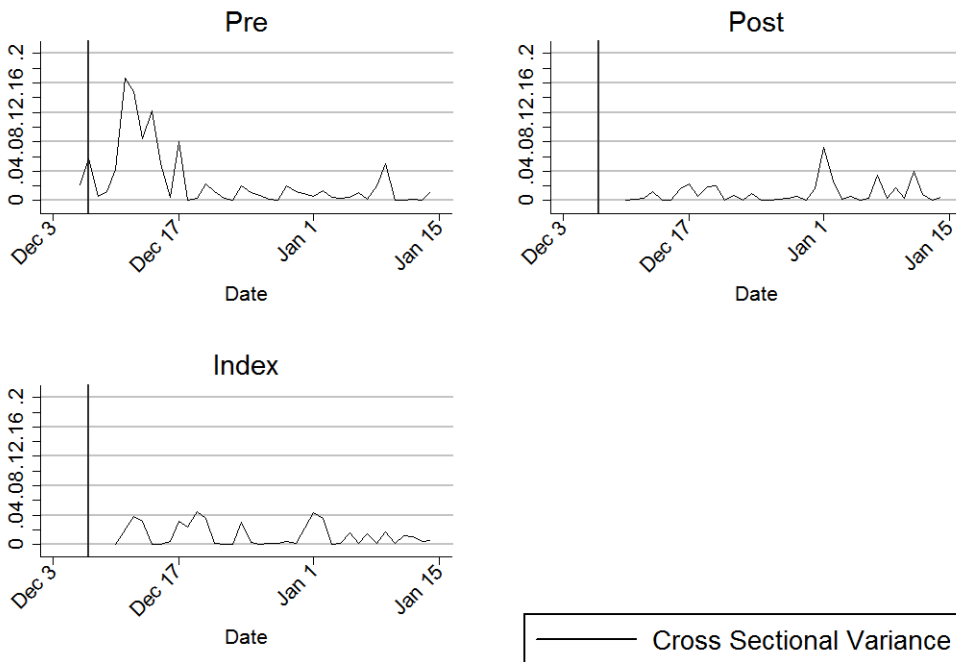
Firetree - 4



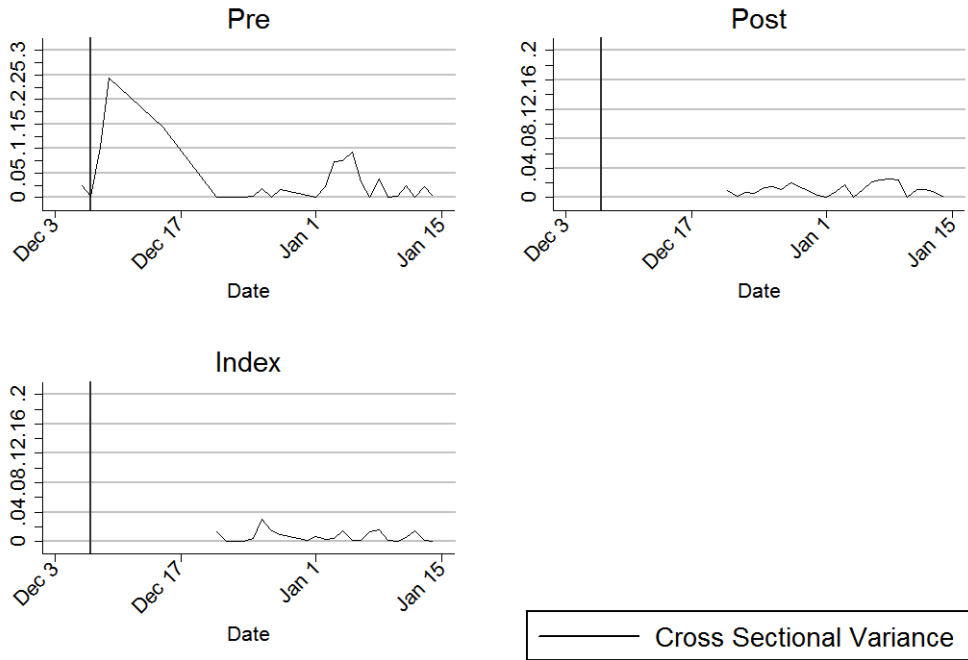
Frostmane - 5



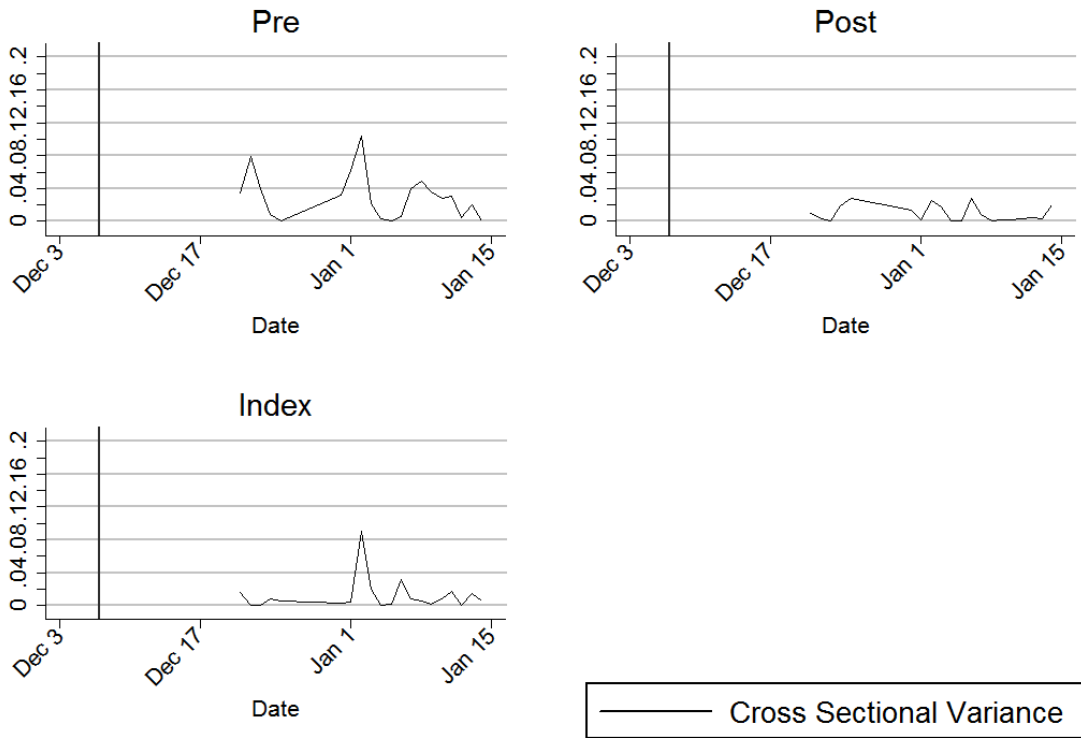
Garrosh - 6



Khaz'goroth - 7

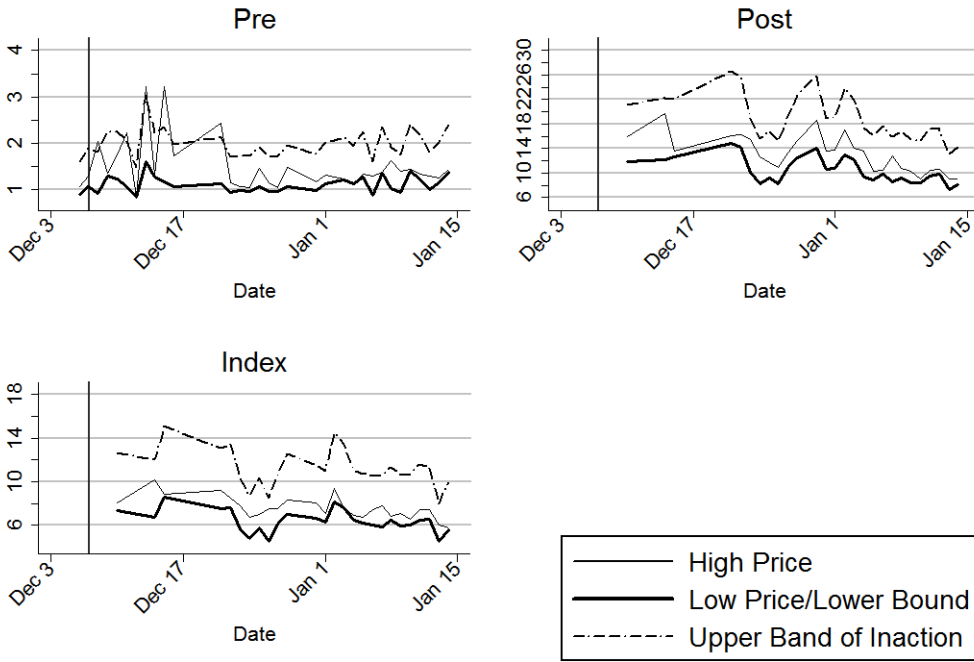


Thrall - 8

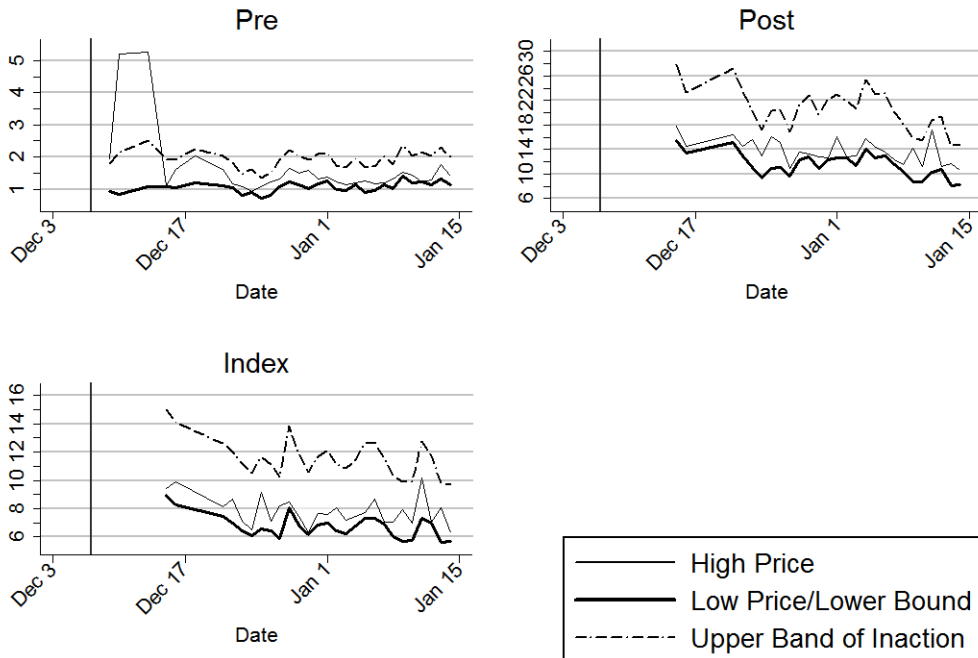


Appendix C

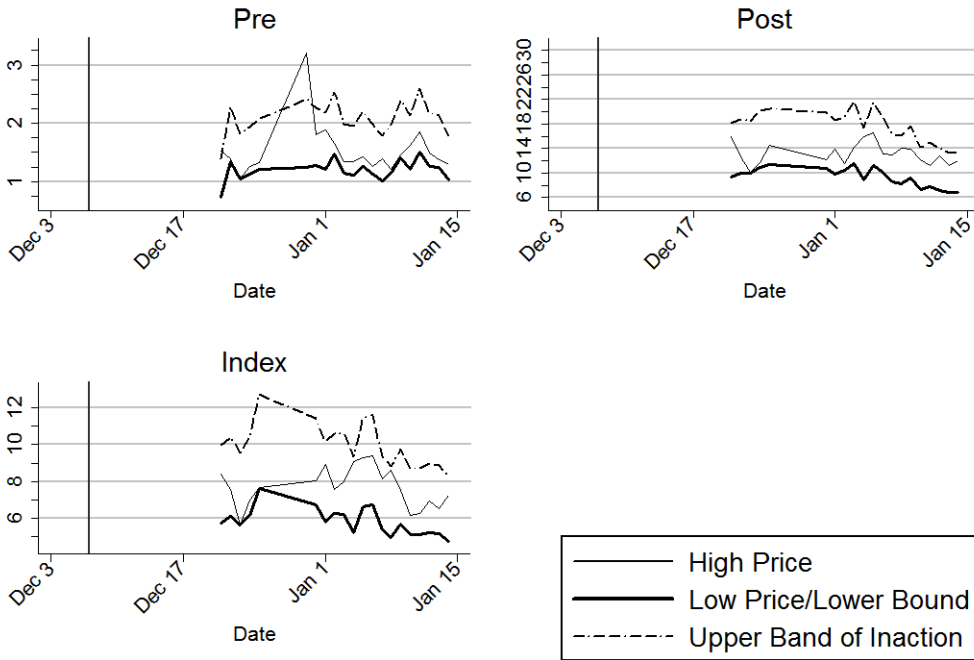
Azjol-Nerub - 1



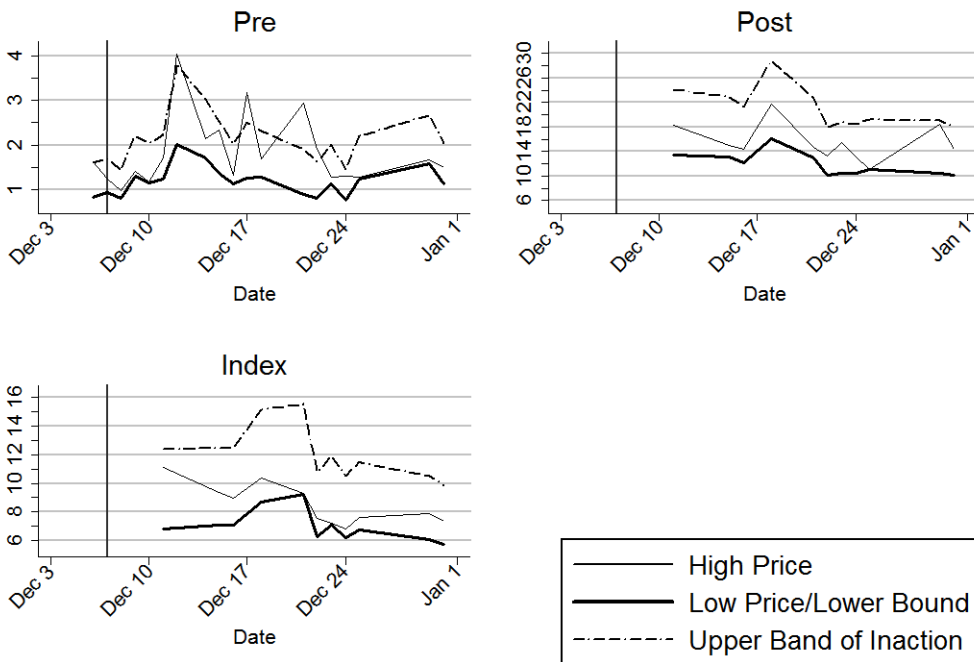
Draka - 2



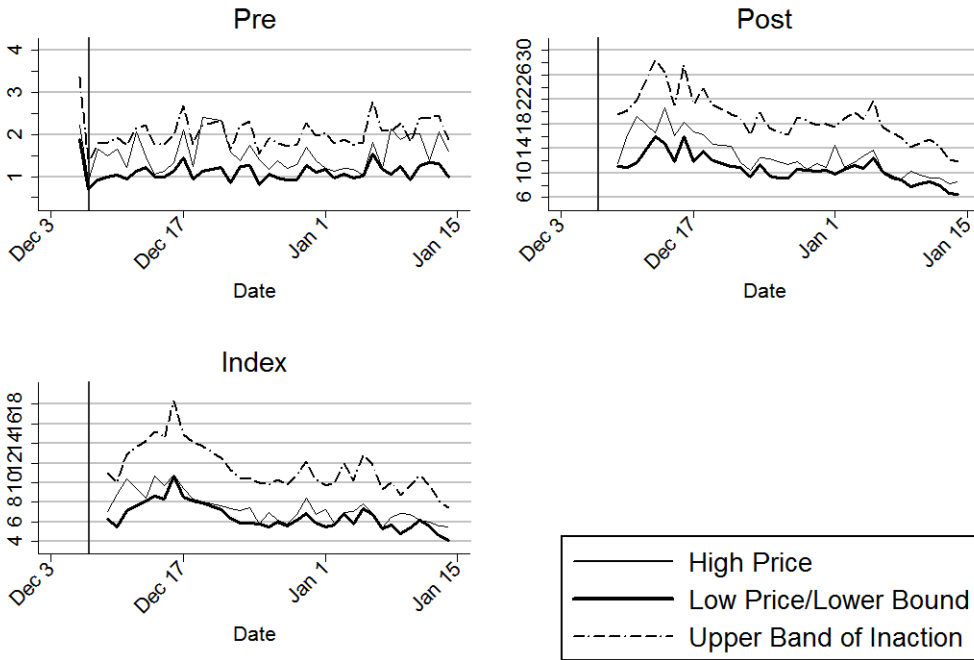
EarthenRing - 3



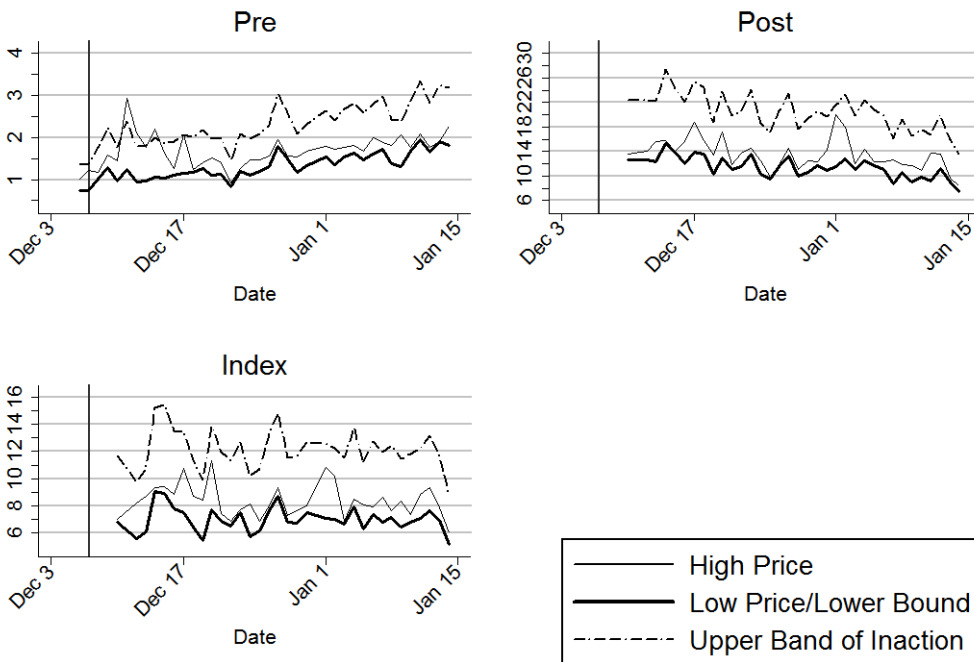
Firetree - 4



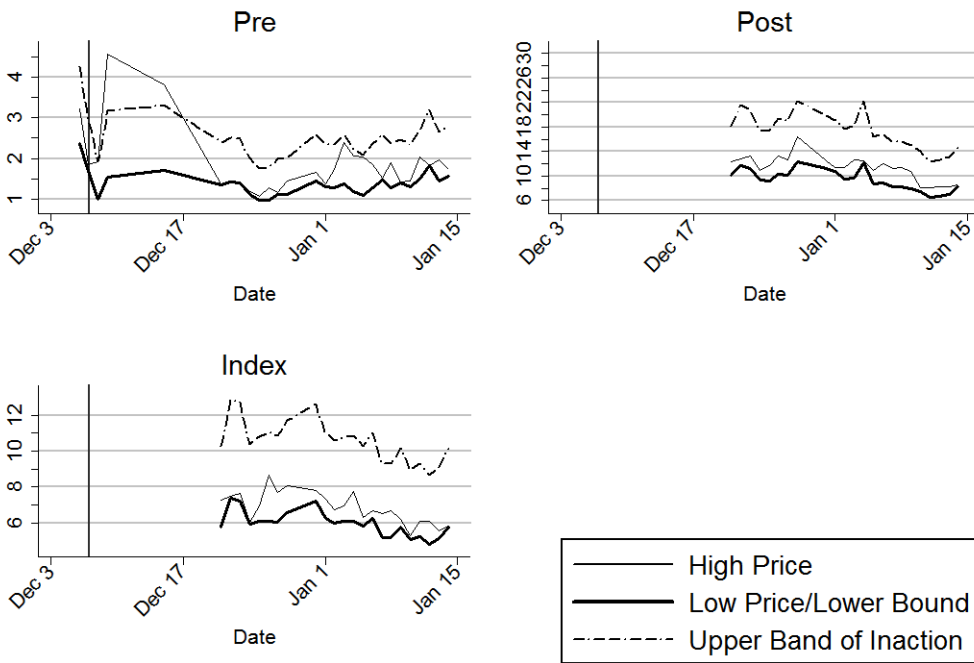
Frostmane - 5



Garrosh - 6



Khaz'goroth - 7



Thrall - 8

