June 2018

Essays in Applied Microeconomics

John Hartman
University of South Florida, jdhart306@gmail.com

Essays in Applied Microeconomics

by

John Hartman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Economics
Department of Economics
College of Arts and Sciences
University of South Florida

Major Professor: Andrei Barbos, Ph.D.
Gabriel Picone, Ph.D.
Joshua Wilde, Ph.D.
Benjamin Craig, Ph.D.

Date of Approval:
June 12, 2018

Keywords: reputational effects, credence goods, rationality, real-world outcomes

ACKNOWLEDGMENTS

I want to thank my major professor, Andrei Barbos, for all of the time and effort he has spent making this a better project. I would also like to thank the rest of the committee for their ideas and support. Specifically, Benjamin Craig has personally mentored me for the past four years, and I now consider him a good friend.

My friends and family have patiently supported me throughout this entire process. My wife, Courtney, has been forced to put up with countless late nights and weekends alone. She has been my rock to lean on. My children, Juliana and Penelope, are always there to cheer me up after a long day. My parents, Joe and Beth, have helped me with grammar and formatting, while my in-laws, Edie and Peter, have acted as a second set of parents to help make sure everything else in my life is in order. Without their help, I would not be in this position today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In the first chapter, I investigate reputational effects of the disclosure of negative information in a market affected by adverse selection. A series of recent discoveries has increased consumer concern over the presence of counterfeits in the market for fine and rare wine. For the thousands of bottles sold at auction each year, house reputation is used as a quality assurance mechanism to signal product authenticity. Using sales data from 2005-2015 for the ten largest auction houses, I study consumer reaction following two recent disclosures of an auction house having offered or sold counterfeit wine. My identification strategy to examine reputation involves a series of triple difference regressions analyzing equilibrium prices and quantities. I discover one house experienced no losses following a 2008 incident involving 107 counterfeit bottles. However, three houses associated with a 2012 incident involving thousands of bottles were found to have suffered significant reputation losses following the incident. These losses are demonstrated by a 3-8% decrease in equilibrium sales prices and a 6-9% decrease in sales quantities in the year following the disclosure.

The second chapter of my dissertation involves the transitivity of stated preferences. Revealed preference theory states that, in order for an individual's preferences to be consistent with utility maximization, they must satisfy the principle of transitivity. Any deviations from this principle result in a logically inconsistent response pattern. I develop a new framework to study the rationality of stated preferences, accounting for both the number and severity of non-transitive responses an individual makes. I implement this method using a nationally

representative survey of 3,234 respondents from the U.S. general population and discover that more than 52% of the population exhibit non-transitive preferences. In addition to measuring the number and severity of non-transitive preferences exhibited by each respondent, another aim of this manuscript is to evaluate the relationship between response transitivity and the individual outcomes of each respondent under the premise that high quality decisions are the result of greater decision-making ability. After controlling for demographic characteristics including age, education, race, gender, ethnicity, and work status, non-transitive patterns are correlated with lower incomes and poorer health.

CHAPTER ONE:

REPUTATIONAL EFFECTS OF CREDENCE GOODS: A STUDY OF FINE AND RARE WINE

1. Introduction

Information asymmetries between buyers and sellers at the time of trade cause many markets to suffer from adverse selection. Consumer uncertainty regarding product quality has the ability to depress prices, in turn lowering the overall quality of goods available in the market. To remedy this problem, markets have developed quality assurance mechanisms in the attempt to better inform consumers. Common mechanisms include product branding, word of mouth, manufacturer guarantees/warrantees, industry disclosure reports (accreditation organizations), third-party agents (*Consumer Reports*), government mandated reports (vehicle safety), and licenses to practice (bar association).[1] With the exception of government mandated reports and licenses to practice, many of these mechanisms require a carrot and stick approach where consumers reward and punish firms based upon reputation and past experiences.

The value that reputation provides is evidenced by discrepancies in price between like goods from multiple vendors. The stronger a firm's reputation in comparison to its competitors, the higher the expected price differential. Empirical research on the collectables market generally supports this assumption, while also finding:

[1] See Dranove and Jin (2010) for a full description of these mechanisms.

1) A good reputation becomes even more valuable as product uncertainty increases (Dewan & Hsu, 2004; Elfenbein, Fisman, & McManus, 2015; Lewis, 2011; Melnik & Alm, 2005)
2) A negative review has a greater impact on reputation than an additional positive review (Canals-Cerda, 2012)

Both of these stylized facts play a critical role in the design of this study. For a firm that sells multiple goods, a change in overall reputation may affect some products more than others. Items with more associated risk may see larger price fluctuations compared to those with less comparative uncertainty.

In this manuscript, I focus on auction sales of fine and rare wine to test these theories. Like other collectibles with a finite production, as rare wines are consumed, the remaining inventory increases in value. In addition, a wine's quality is generally expected to improve with age.[2] As the rarity and/or value of a wine increases, consumer concern regarding product authenticity is also expected to rise. Once a wine appreciates beyond a certain threshold (commonly considered in the $1,000 range), it becomes profitable to counterfeit. Recent advancements in printer quality and graphics design have led to improvements in the appearance of counterfeit labels to a point where, for the average consumer, they appear identical to the genuine product. The more expensive the wine, the larger the potential financial loss if a bottle is later discovered as a counterfeit. Consumers may also have less collective experience with more valuable wines, decreasing the likelihood a fake bottle would be detected. The consumer's inability to differentiate between authentic and counterfeit bottles, even after purchase, makes such wines a credence good.

[2] While all wines do eventually spoil, many collectible wines have an expected "lifespan" upwards of a century under proper storage conditions.

Utilizing the knowledge (e.g. type of glue used on the label of an authentic bottle, exact ink color, hidden bottle markings) and experience gathered over thousands of sales, auction houses have claimed the expertise to accurately assess product authenticity.

In total, wine auctions generated more than $3.3 billion in sales globally between 2006-2015. Before all sales, auction house are expected to thoroughly inspect each bottle to ensure authenticity and overall bottle condition. Through repeated sales, auction houses have developed individual reputations for product authenticity, making them a preferred method to buy and sell fine wine. Consumers expect that the likelihood of purchasing a counterfeit bottle is lower as a result of the authentication process; therefore, even a small negative review could potentially harm firm reputation.[3]

Yet, auction house inspections are not infallible. The authentication process reveals conflicting incentives: trust and reputation vs. sales. Lower standards and weaker certification measures could increase sales, particularly given the limited production and availability of the most expensive wines. However, doing so may decrease future sales and reputation if any counterfeit bottles are discovered. Each house is thus faced with the task of determining the optimum level of authentication standards to maximize expected profits. Over time, as consumers make repeated purchases, each auction house develops a reputation for their authentication practices. The better the reputation, the greater the buyer's belief that the wines offered are genuine.

Recent discoveries of international counterfeiting operations have brought the issue of product authenticity into the global spotlight. Multiple auction houses have been linked to the

[3] It has been noted that most counterfeit wine goes undetected. Because of this, one negative review could signal a much broader issue.

sale of wines produced by these operations, and it is important to understand whether firm reputation suffered as a result. In this manuscript, I study consumer behavior following two recent counterfeiting disclosures.

To access any reputation losses, I test for shifts in the demand curve for the auction houses implicated. Consumer demand for wine purchased from a specific auction house is a function of house reputation and various other factors (including the state of the global economy, customer service, and auction house experience). Through a series of difference-in-difference (DD) and triple difference (DDD) regression models estimating equilibrium prices and sales quantities, I isolate the effects of reputation on consumer demand. A decrease in either the comparative prices or sales quantity for the house(s) implicated without an increase in the other is indicative of a downward shift in the demand curve, and thus, a loss of reputation. By grouping similar wines according to previous counterfeit discoveries and recommendations of third party experts, one can test whether the value of an auction house reputation varies for different wines.

Using a rich data set covering more than 100,000 observed sales, I find that consumer responses to negative information vary. Consumers did not punish one auction house for authenticating counterfeit wine in April 2008. However, a larger 2012 discovery and arrest of an individual for counterfeiting thousands of bottles caused the three auction houses he was associated with to suffer significant reputation losses following the disclosure. The reputation losses are inferred by a 3-8% decrease in the comparative prices for wines likely to be counterfeited and a 6-9% decrease quantity of wines sold. I also observe that the value of firm reputation increases with product uncertainty, confirming much of the previous research on credence goods.

In addition to our main finding presented above, this manuscript makes several added contributions to the previous literature. To date, most studies on the effects of negative information on credence goods have utilized retail sales data (e.g., scanner data) to track changes in reputation (De Paola & Scoppa, 2013; Freedman, Kearney, & Lederman, 2012; Rao & Wang, 2017; Schlenker & Villas-Boas, 2009). A potential issue with these studies involves a lack of variation in sales price. (Breidert, Hahsler, & Reutterer, 2006). If price remains constant over time, any decrease in the demand curve for a product can only be inferred by a reduction in the equilibrium sales quantity. This matter is further complicated when researchers use competing products as a control group in the estimation strategy. Other factors, independent of reputation, could be responsible for any decrease in equilibrium quantities.[4] Potential confounders include the introduction of new competition, changes in consumer preferences regarding small differences between competing products, or adjustments to the shelf space a product occupies. The use of auction sales resolves both issues. Auctions do not suffer from the same price restrictions as retail purchases, and I use sales of identical products from firms not associated with negative information as a control.

The outline of this paper is as follows. Section 2 provides a general overview of the wine market, Section 3 discusses the literature on adverse selection, and Section 4 develops a theoretical model to test for changes in consumer behavior. Section 5 defines the estimation strategy, Section 6 describes the data, and Section 7 presents the results. Section 8 contains a brief discussion and concluding remarks.

[4] A similar argument could be made towards the use of competing auction houses. Firms could improve their customer service and increase sales. I propose a solution in the DDD models by including sales from wines unlikely to be counterfeited.

2. Reputations and Fraud in the Market for Wine

2.1 Product Quality

The quality of a wine is fundamentally determined by three variables: weather (temperature and rainfall patterns), vineyard, and winemaker/producer. For wines from known producers and vineyards, it is possible to predict the quality of a wine (and consequently price) as soon as grapes are harvested (Ashenfelter, 2008; Jones & Storchmann, 2001).

However, a recent expansion in both the number of producers, vineyards, and wine regions has made it much more difficult for consumers to ascertain the quality of wine before purchase (Wine Institute, 2015). Without full knowledge of the producer, vineyard, and vintage, consumers may choose to rely on more generic quality signals (e.g., region of production, alcohol content, or, in worst cases, label appearance or price) when making a purchase.

Issues resulting from the noise of these generic signals have given rise to consumer reliance on third-party tasting experts to aid in the decision making process (e.g., Robert Parker, Wine Spectator). These reviews often contain a short tasting description and an overall rating on a 100-point scale. Research on the effect of expert reviews on prices has been well documented, and the literature consistently shows a strong relationship between price and review score (Ali, Lecocq, & Visser, 2008; Friberg & Gronqvist, 2012; Hilger, Rafert, & Villas-Boas, 2011). As with food, individual tastes and preferences for wine vary. Yet, by comparing past personal experiences to expert reviews for those same wines, it is possible for consumers to accurately predict the quality of a wine they have never tasted.

The limited production of each wine has resulted in an active secondary market, including auction house sales and trades amongst individual consumers. The secondary market is also where a majority of the adverse selection is introduced. While consumers likely have accurate expectations about the inherent quality of each wine, they are unable to judge the authenticity of each individual bottle. To mitigate this problem, buyers use firm reputation as a signal for product authenticity.

2.2 Auction Houses and Reputations for Product Authenticity

Auctions are the preferred method for procuring old and rare bottles for many collectors, due in part to the selection and authentication process bottles must go through before a sale. It is widely regarded that consumers value the reputation of each firm, yet empirical research on the effects of auction house reputation on price remains limited. Obtaining data from sales outside auction presents a challenge. Private trades are often untraceable, and the listed retail prices for expensive wines may have room for negotiation. Thus, one is often left without a control group for comparison.

Instead of attempting to value the individual reputation of each firm, I compare sales from different auction houses over time to track changes in house reputation. A portion of the price differential for the same wine between houses can be attributed to their record of suspected prior sales of counterfeit wines. The better a firm's reputation, the higher the demand for each wine. By tracking the equilibrium prices and quantities for a wine repeatedly sold at auction by multiple firms, one can test whether auction house reputation suffers after the public discovers an auction house authenticated counterfeit wine. For the remainder of the paper, reputation is associated with consumer beliefs towards the ratio of counterfeit to genuine wines an auction house offers for sale.

2.3 History of Counterfeit Wine

Records of wine counterfeiting date back at least two thousand years to ancient Rome. Pliny the Elder wrote in the first century that even the nobility were not able to obtain wine that had not been altered in some fashion (Robinson, 1999). Possibly the most common counterfeiting method involves rebranding a cheaper wine as a more expensive one. Counterfeiters can take a cheaper wine from a neighboring vineyard or less valuable vintage, remove the labels, and attach a new forged label of a more valuable wine.

Another method to counterfeit wine involves mixing different wines. In his seminal classic on wine, "A topography of Vineyards," Andre Jullien noted that many Bordeaux merchants in the 18th century would mix expensive bottles with cheaper wines from Spain or the Rhone to increase profits (Jullien, 1816). It was common for producers to sell entire barrels of wine to merchants, allowing for the merchants to actually bottle the finished product. Some merchants were known to have added cheaper wine to increase the quantity available for sale.

Over the past 200 years, strict marketing practices have been implemented in an effort to limit such counterfeiting practices These included restricting the use of certain terms (e.g., Champagne[5]) to wines produced within a geographical boundary, quality classification systems (Bordeaux Classification of 1855[6]), grape varietal restrictions (Burgundy), and lists of permissible additives (sulfates, yeast).

Most recently, as with other markets where counterfeit goods are present (e.g., currency, pharmaceuticals), winemakers have begun using technology to prevent fraud. Advancements

---

[5] In order for a wine to legally be named Champagne, it must originate from that specific region of France.

[6] The Bordeaux Classification of 1855 ranked the best of the wines from the Medoc region of Bordeaux by placing each estate into one of five categories ranging from First growth to Fifth growth. Even though this classification is more than 150 years old, it remains relevant today. Wines given First Growth Status may sell for twenty times that of a Fifth Growth and five times that of a Second Growth. Similar ranking classifications are also provided for other regions in Bordeaux (St. Emilion), Burgundy, and Champagne.

include the use of bubble strips to note if a bottle has been opened, laser-etched bottles, individually numbered bottles and cases, and special inks on labels and foil (Lecat & Chapuis, 2017; McCoy, 2007).

However, these techniques have only become popular over the last few years. Many older wines were not produced under threat of counterfeiting. It was not until years later that they became valuable enough to counterfeit. Some wines even lack a basic production record, providing few details on how many bottles were produced, or if there were multiple versions of the same wine. Producer labels often changed from year to year, and wines sold in bulk could be bottled under multiple labels (Asimov, 2007).

2.4 Identifying Counterfeit Wine

Identifying counterfeit wine is a difficult task. One authentication method commonly used in the past involved physically tasting bottles from a lot to be sold. However, many experts have challenged this method as a true authentication technique. Not only is tasting expensive (i.e. it reduces the number of bottles available for sale), but it is highly imprecise. If an individual has been previously duped by a counterfeit, all subsequent counterfeits would have the same taste. Even when experts or consumers have tasted a genuine example of the wine, older wines have substantial bottle variation. Storing wine at different temperatures will age bottles at a different pace, and the corks used to seal wine from air deteriorate over time. A counterfeiter can also include one genuine bottle in the lot to be consumed that appears identical to the counterfeits. The label from an authentic bottle to be tasted by the expert could be replaced with a copied on that matches the counterfeits. The bottle tested by experts would be genuine while the others later offered for sale would be counterfeit.

One technique to identify counterfeits that does not require the bottle to be opened involves mass spectrometry. By measuring the molecular compounds present in each bottle, scientists can determine a wine's general age and characteristics (American Chemical Society, 2010). For wines produced after 1945, it is possible to determine wine vintage by measuring the Carbon-14 content. Atomic bombs dropped in World War II and subsequent nuclear testing during the Cold War released Carbon-14 into the atmosphere, and each living organism absorbs trace amounts of the element. The total presence of Carbon-14 in the atmosphere changes every year, and there is a commensurate shift in the amount present in each vintage of wine from similar locations. However, there is no DNA level test to prove the exact vineyard.[7] A counterfeiter could mimic a $10,000 bottle using wine from a $100 bottle of the same vintage. Given the issues with other methods, the use of authentication experts have proven to be one of the few reliable techniques to identify counterfeit bottles.

Today, authentication experts use a variety of techniques to measure a wine's authenticity. The glass is examined to study the age of the bottle. Older bottles were hand blown, and for newer wines, the shape of the bottle may be unique to individual producers. Older bottles should have a certain amount of sediment covering the inside glass. Labels are examined for misspelled words, correct font and color, and paper type. Corks must be stamped with the proper branding, and the capsule used to cover the cork must be true. Some bottles are dipped in wax while others have specific designs and labels on a metal capsule. If any one of these aspects appear incorrect, a wine may be determined as counterfeit and rejected from the consignment.

[7] Even if the individual vineyard could be determined, certain vineyards in Burgundy (including Echezeaux and Richebourg) are divided among multiple producers. There is sometimes a large difference in price depending on who produced the wine.

2.5 Counterfeiting in This Study

Both instances of wine fraud examined in this manuscript stem from the actions of one individual over the course of a decade (2003-2012). In the early 2000's, Rudy Kurniawan cornered a large portion of the market for old and rare wine by spending upwards of $1 million a month at auction (Wise, 2008). After the value of his collection had risen considerably, Kurniawan eventually began selling a portion of the wine he had previously required. Two 2006 auctions of Kurniawan's wines, titled "The Cellar I and II", were conducted by Acker Merrall & Condit (Acker) and generated more than $35 million in sales. The second of these sales still holds the record for the largest single auction in terms of total sales ($24 million) (Robinson, 2007).[8] Kurniawan also sold wine privately outside the auction market.

However, over a series of events in 2007, doubts began to arise over the authenticity of many wines sold by Rudy Kurniawan. Bottles of 1982 Le Pin, consigned by Kurniawan, were withdrawn from a Christie's auction after an examination of the corks revealed the wines to be counterfeit (Hellman, 2007). In a separate event, a small tasting of eleven old Roumier bottles purchased from "The Cellar" auctions discovered six of the eleven to be counterfeits (Steinberger, 2012). Even after these events, a select group of auction houses continued to accept and authenticate consignments from Kurniawan, often "sight unseen" (Hellman, 2008).

The first critical counterfeiting case surfaced in 2008, when 22 lots of Domaine Ponsot wines were removed from an Acker Merrall & Condit auction at the request of the winemaker. According to the winemaker Laurent Ponsot, the bottles, on consignment by Rudy Kurniawan, could not be genuine because the purported wine was never produced. The Acker catalog for the

[8] After Rudy Kurniawan's arrest in 2012, it was revealed that a significant percentage of the wines offered in "The Cellar" auctions were counterfeits.

sale (provided before the wines were removed) confirmed Ponsot's claims, yet the wines were still authenticated. A bottle dating from 1929 was offered for sale, but the catalog mentions that wine from that particular winemaker and vineyard was not made until 1934 (Kapon, 2008).

Although there is an extensive history of suspicion regarding wine fraud/counterfeiting, 2008 appears to be a turning point (McCoy, 2007; Wallace, 2009). Media reports and consumer posts on public wine forums began suggesting auction houses were failing to perform their due diligence in authenticating wines (Hellman, 2008; Wise, 2008). Some reports even questioned Acker's complicity in the sale of counterfeits (Hellman & Frank, 2009; Squires, 2008). In spite of this finding, multiple auction houses continued to discretely accept consignments from Kurniawan.

It was not until his arrest in 2012 that the scope of Kurniawan's counterfeiting operation was uncovered. (Hellman & Frank, 2009; Hernandez, 2012; Steinberger, 2012; Wine Berserkers, 2012). In the Spring of 2012, a consignment from Kurniawan to Spectrum Wine Auctions was publicly contested on a popular wine forum on the basis that many bottles were clearly counterfeit (Wine Berserkers, 2012). A 2009 lawsuit against Kurniawan for selling counterfeit bottles began to make headway, and the FBI determined that Kurniawan was not a legal resident of the U.S. (as he claimed to be) (Wallace, 2012). On March 8, 2012, Kurniawan was arrested for crimes associated with counterfeiting wine.

At the time of his arrest, a search of Kurniawan's home revealed a large counterfeiting operation complete with a corking machine, labels for expensive wines, empty bottles, and notes on mixtures of cheaper wines that mimicked more valuable ones. Mr. Kurniawan would later become the first person to be convicted of wine fraud in the U.S. (Hernandez, 2012).[9] This case received more media attention than all previous accusations of fraud (Downey, 2012; Hirsch, 2012; Lecat & Chapuis, 2017; Pfanner, 2012; Steinberger, 2012; Wallace, 2012; Wine Berserkers, 2012). While counterfeiting allegations have persisted in the years following Kurniawan's arrest, very few have been documented, and no single case has approached a similar scale.

3. Literature on Reputations Markets and Credence Goods

3.1 Adverse Selection

I use a framework similar to that adopted in several other papers on adverse selection. George Akerlof (1970) was the first to investigate the effects of adverse selection in a market for goods of uncertain quality . Akerlof's model demonstrated that asymmetric information between the buyer and the seller at the time of trade may lower the welfare of both parties. Buyer uncertainty can depress market prices, decrease product quality, and diminish the size of the market. Under certain circumstances, asymmetric information may prevent all trade from occurring. As a remedy, Akerlof suggested a market structure where quality assurance mechanisms (e.g. third-party experts, brand names, licensing organizations) are generated for the sole purpose of identifying true product quality. Given that consumers place a higher value on known goods, these institutions can increase supplier profits and consumer welfare at the same time. Buyers use the brand reputation of each mechanism as both a sign of quality and, "as a means of retaliation if the quality does not meet expectations." Specifically, Akerlof illustrates this with the example of a restaurant chain using its brand value to attract out-of-town visitors with little

[9] For those interested, a detailed description of the events is provided by Wallace (2012), Wine Berserkers (2012-), and the film, *Sour Grapes* (2016).

knowledge of the local area. Customers have confidence in the quality of products purchased from vendors with familiar positive reputations.

Darbi and Karni's (1973) work extended Akerlof's theory to include credence qualities; product qualities that cannot be observed or evaluated by normal use . In this case, the consumer will have no knowledge of whether or not the good purchased is either necessary or real (e.g. car maintenance, health care).

Due to the high costs and difficulty of obtaining reliable information about certain goods, Darbi and Karni note that individuals tend to obtain both the information and the product from the same source. For example, auction houses and used-car dealerships often act as a combined certification/sales agent. This reliance on a single firm for both may lead to a certain amount of fraud. Car mechanics could provide repairs that are not necessary, dentists might fill a cavity that does not exist, and auction houses may sell a counterfeit product.

When purchasing credence goods, consumers may be more likely to gather and share information from previous transactions to aid in the decision making process. As more information becomes available, a reputation for each firm develops. Positive experiences are expected to improve reputations, and a better reputation should coincide with the ability of a firm to obtain a higher price for a good or service. In consideration of this study, wines offered for sale from an auction house with a better reputation are expected to have a higher demand than wines from an auction house whose reputation and authentication standards are more suspect. The revelation of any new information, either positive or negative, is expected to affect future sales prices and/or quantities.

3.2 Firm Reputation

Shapiro (1983) developed a model for the reputations market . The model elicits an equilibrium, in which, once a firm establishes a reputation, it is expected to remain consistent for future time periods. A strong reputation is noted to be especially valuable when the product is rare or infrequently sold or if product quality is difficult to assess. Shapiro also proved that as information improves, the equilibrium outcome approaches that under perfect information.

Chu and Chu (1994) introduced the idea of a manufacturer renting the reputation of another firm to facilitate the sales process . A manufacturer of a high-quality good, but without an established reputation, can use the reputation of a respected dealer (e.g. auction house) as a signal to consumers. Reputable middlemen have an incentive to correctly represent the products they sell in order to maintain their individual reputation. Chu and Chu's model obtains a separating equilibrium where reputable middlemen earn positive profits while non-reputable middlemen earn no profit.[10]

However, Chu and Chu's model relies on the assumption that consumers can determine true product quality immediately after consumption. Baksi and Bose's (2007) study of third-party food labelers suggested that the separating equilibrium will not hold for credence goods. Reputable middlemen may find it profitable to misrepresent product quality if the chances of detection and prosecution are low. Baksi and Bose describe a situation where non-eco-friendly firms may deceive customers by labelling the products as eco-friendly. Consumers are willing to pay more for eco-friendly products but are unable to determine whether the product actually meets the criteria. For example, a product using genetically modified ingredients may be falsely

[10] In a separating equilibrium, information is disclosed, and consumers are able to differentiate between seller type. A seller of a genuine good or high quality good will use the reputable middlemen while the seller of the lower quality good will use the non-reputable middleman. The reputable middlemen will have a financial incentive to only accept high quality or genuine goods, and will not find it profitable to cheat the consumer.

labelled as non-GMO. Given that a non-GMO claim is nearly costless to make, there is an incentive for producers of GMO ingredients to mis-label the product. In a similar fashion, auction houses may find it profitable to certify a forged wine as authentic if consumers are unable to differentiate between genuine and counterfeit bottles.

3.3 Reputations and Credence Goods

Emons (1997) suggests a possible solution, involving the use of two separate entities for certification and sale to limit fraud in the market for credence goods. If the certifier is independent of the seller (i.e. no financial interest), there is no incentive to defraud the consumer. However, this solution is often cost prohibitive. Emons describes the market for automobile transmission repairs as an example. Both the independent certifier and mechanic (seller) would have to disassemble the transmission in order to diagnose the problem. While the consumer would never be deceived, the extra costs of an independent examination may outweigh the fraudulent mechanic's misdiagnoses. In the market for wine, a seller may choose not to have his wines authenticated by a neutral third-party expert if the costs for verification exceed the expected increase in sales price.

Dulleck and Kerschbamer (2005, 2006) revisit the issues of previous models for credence goods and propose a less restrictive model to limit fraud. However, their model still requires a set of assumptions not applicable to all credence goods. In the event that the required assumptions do not hold, the authors describe a reputations market as a second-best solution. Consumer's reliance upon reputation systems may be able to curb the amount of fraud in the market.

3.4 Empirical Research on Credence Goods

Empirical research on the effects of fraud/negative product information on consumer demand for credence goods remains somewhat limited. One possible explanation is that obtaining data presents a major challenge, as companies caught defrauding customers are usually unwilling to provide information. Recent studies have analyzed consumer response to contaminated food (De Paola & Scoppa, 2013; Schlenker & Villas-Boas, 2009), product safety recalls (Freedman et al., 2012; Garber & Adams, 1998; Prince & Rubin, 2002), accounting fraud (Toth, 2014), false advertising (Rao & Wang, 2017), and medical malpractice (Dranove, Ramanarayanan, & Watanabe, 2012). These studies have provided contrasting results, even among cases within the same industry. Consumer responses have varied from no punishment at all to a decrease in demand for the entire industry.

Dranove et al. (2012) found physicians suffered a loss in reputation after medical malpractice lawsuits. A medical malpractice lawsuit often results in a shift in patient type from those using private insurance to those on government plans. De Paola and Scoppa (2013) show that a producer caught mixing rotten ingredients in cheese experienced a loss in reputation, and both the producer and retailer (middleman) suffered financial consequences from the reputation loss. The consequences of the reputation loss were found to last more than a year after the negative information was disclosed. Rao and Wang (2017) observe that firms discovered to have made false claims regarding the healthiness of their food products suffer a significant decrease in consumer demand. The four products studied experienced a decrease in monthly revenues of between 12-67% following the exposures.

Freedman et al. (2012; toy recalls) and Prince and Rubin (2002; automobile and pharmaceutical liability) observe incidents where the firm suffers financial consequences but no loss in reputation. Toy manufacturers reported lower sales for the products directly affected by recalls but felt no impact on other items offered. For pharmaceutical and automobile manufacturers, losses are similar to the direct effect of the potential problem. No additional punishment is felt.

For some types of goods, consumers may infer that all related products are manufactured in a similar style. The association of similar products may cause a decrease in the market demand for all goods. The Freedman et al. (2012) study of toy recalls also discovered a punishment for all toys in the related industry. The sales of all toys in the product segment, including those from manufacturers whose products were not recalled, decreased by more than 30% after the recall. Schlenker and Villas-Boas (2009) discovered a similar effect following health warnings regarding contaminated meat. Industry wide, beef sales dropped by nearly 20% following the 2003 outbreak of mad cow disease.

In certain circumstances, the disclosure of negative information may not result in any punishment. Garber and Adams (1998) examined the impact of two product liability verdicts on sales in the automobile industry and found no evidence of a decrease in sales or change in stock price following a verdict. The strong loyalty many consumers have to specific vehicle manufacturers may limit the effect a product recall has on future purchases.

3.5 Consumer Inexperience

Low levels of consumer experience may reduce the importance of firm reputation. Inexperienced buyers may have limited knowledge or understanding of the market in general. They may not be aware of product fraud or the reputation systems used to combat it. An investigation of baseball card auctions by Jin and Kato (2006) found that sellers with poor reputations could increase sales prices by making non-verifiable claims about product quality . In other words, a seller could increase the sales price by simply stating the card was in outstanding (mint) condition. Individuals not familiar with the market could be more likely to accept such quality statements at face value. The authors discovered that, while sellers with a better reputation were less likely to make false claims, consumers were unwilling to pay more for the same claim made by a firm with a better reputation.

4. Auction House Reputation

It has been estimated that as much as 20% of investment grade wines sold in the market today are counterfeit (Richard, 2013). Over time, each auction house develops a reputation for authenticity based upon their prior sales record. I use a variation of Shapiro's (1983) model and Dewan and Hsu's (2004) extension to test whether auction house reputation suffers following a discovery that the house has sold or offered counterfeit wine for sale.

4.1 Buyer's Decision

For wines that are likely to be counterfeited, buyers are assumed to hold different values for authentic and counterfeit bottles. The value for a genuine bottle of wine i, during time t, is V_{Hit}, while a counterfeit bottle has value V_{Lit}, where $V_{Hit}>V_{Lit}$. Buyers believe a bottle of wine i sold by house s during time t is genuine with probability R_{ist} and counterfeit with probability (1-R_{ist}). R_{st} represents the reputation of house s at time t. House reputation may change over time as new information becomes available (e.g. lawsuits, media coverage, etc.). However, as with all reputation functions, the information used to determine each house's reputation is based upon previous sales.

It is expected that some counterfeits will never be detected. For each different wine that an auction house makes available for sale, Q_{ist} corresponds to the proportion that are counterfeit while (1-Q_{ist}) represents the proportion of authentic bottles a house authenticates during time t.[11] Authenticating and offering more counterfeit bottles for sale increases the likelihood that a counterfeiting claim will be realized. During each time period, the probability of a counterfeiting disclosure, D, is equal to:

$$P(D_{st} = 1) = f\left(\sum_{j=1}^{n}\left(M_{ist}Q_{ist} + \frac{M_{ist-1}Q_{ist-1}}{1+\rho} + \frac{M_{ist-2}Q_{ist-2}}{(1+\rho)^2} + \cdots\right)\right) \quad (4.1)$$

where f is an increasing function and $0 \leq P(D_{st}) \leq 1$. D_{st} is a dummy variable equal to one if, during the current time period, the auction house was discovered to have previously authenticated counterfeit wine. M_{ist} represents the quantity of each wine i, offered for sale by house s, during time t, and j denotes the number of unique wines sold by the auction house. As time passes after a counterfeit bottle has been sold or offered for sale, the probability that it will be discovered as counterfeit decreases for each subsequent time period. In Equation 4.1, this is accounted for by the inclusion of a discount factor, ρ.

The individual values for R_{ist} depend upon the specific characteristics of each wine, B_{it} (namely price and vintage), and previous disclosures against each house. For wines not valuable enough to counterfeit, R_{ist} is expected to be equal for all firms during all time periods (i.e. R_{ist}=1). This assumption plays a critical role in the estimation strategy described in Section 5.

Consumers also discount previous discoveries of counterfeit wines as time passes (denoted by ϱ in Equation 4.2), so the effect of a disclosure years before carries less weight than a discovery during the previous time period. The reputation function for each wine is as follows:

[11] Due to differing levels of knowledge, it is possible that some auction houses consciously accept counterfeits while others unknowingly make errors. Both cases are treated equally in this study.

$$R_{ist} = g(\, B_{it}, D_{st}, \frac{D_{st-1}}{1+\varrho}, \frac{D_{st-2}}{(1+\varrho)^2}, \ldots) \tag{4.2}$$

For wines likely to be counterfeited, each additional disclosure is expected to decrease the value of R_{ist}. Using Equation 4.2, we can determine the expected value a consumer has for each wine:

$$V_{ist} = R_{ist}V_{Hit} + (1 - R_{ist})V_{Lit} \tag{4.3}$$

Values for authentic and counterfeit bottles, V_{Hit} and V_{Lit}, are independent of auction house reputation and equal for all firms.

4.2 Auction House Prices

To estimate the sales price for each bottle, Y_{ist}, we must also consider various other factors beyond V_{ist}. These include transaction costs and the quantity of bottles sold by each house. As more bottles of the same wine are sold by the house during each time period, it is expected that the equilibrium price will decrease (i.e. $Y'_{ist} < 0$). Transactions costs (e.g. shipping fees, time and effort required as part of the bidding process, and customer service) also play a significant role in the sales price. Houses with lower transaction costs may have higher final sales prices than houses with better reputations. In Section 5, I propose an estimation method that controls for outside factors, allowing us to determine if firm reputation suffers following the disclosure of negative information.

5. Estimation Strategy and Methodology

For each auction house, any change in the comparative equilibrium prices and/or quantities is the result of a shift in house supply or demand (or both). As our goal is to elicit if the auction houses suffered a reputation loss following the disclosures, we seek to uncover a downward shift in the demand curve. A downward shift in the demand curve (and thus reputation) can be identified by a decrease in the comparative equilibrium prices or sales quantities without an increase in the other. Through a series of difference-in-difference and triple difference regression models, I am able to test whether consumer demand shifted following the discovery that an auction house has authenticated counterfeit bottles for sale.

5.1 Empirical Model for Auction House Prices

Difference-in-difference (DD) and triple difference (DDD) approaches are commonly used in natural experiments to study the outcomes of a certain event or treatment and have become popular due to their ability to capture important variables omitted from the data (Angrist & Pischke, 2009). A description of each approach is provided in the following sections.

5.1.1 Difference-in-Difference Price Model

Following Angrist and Pischke (2009) and Greene (2012), the difference-in-difference approach first defines treatment and the time it occurred. For this project, time is a binary variable denoting if a sale (of bottle *i*) occurred before (*Time*=0) or after (*Time*=1) an auction house (*s*) was discovered to have authenticated counterfeit bottles and offered them for sale.

$$Y_{is0} \text{ if } Time = 0$$
$$Y_{is1} \text{ if } Time = 1$$

If we assume demand and supply for the entire market remained constant, the effect of the discovery of an auction house authenticating counterfeit wines could be measured as

$$Y_{is} = Y_{is0} + (Y_{is1} - Y_{is0})Time$$

where the difference between Y_{i1} and Y_{i0} is equal to the size of the effect.

However, as displayed in Figure 1, the assumption that demand[12] remains constant over time is highly suspect. As with nearly all collectable items, the demand for investment grade wines varies over time due to factors unrelated to discoveries of counterfeit bottles, including the state of the global economy and changes in consumer tastes. To control for these shifts in demand, one can include a sample group of control observations separate from the treatment group.

In this study, treatment (House) is a binary variable differentiating firms discovered to have accepted/authenticated counterfeit bottles (*House*=1) from firms not implicated (*House*=0).

$$Y_{i0t} \text{ if } House = 0$$
$$Y_{i1t} \text{ if } House = 1$$

Under this model, the size of the effect is measured as

$$(Y_{i11} - Y_{i10}) - (Y_{i01} - Y_{i00})$$

By separating the effects of two different binary variables (Time and House), the equation above provides the basic difference in difference model.[13]

When aggregating multiple wines and observations, a regression model will likely improve the estimates of the treatment effect. In the regression model, I control for the fixed effects of each individual wine, auction house, sales date, sales location, bottle size, and lot size. The regression difference-in-difference equation is provided below:

$$\log(Y_{ist}) = \alpha_0 + \beta_1 X_{ist} + \beta_2 \delta_i + \beta_3 \gamma_s + \beta_4 \lambda_t + \beta_5 House_s + \beta_6 Time_t \qquad (5.1)$$
$$+ \beta_7 House_s \bullet Time_t + \varepsilon_{ist}$$

[12] Or supply, for that matter.
[13] A table showing an example calculation is provided in the appendix.

Y_{ist} is price of bottle (*i*) from house (*s*) in time (*t*). Sales price is logged to allow for a direct comparison between wines of different values (i.e. hundreds. vs. thousands of dollars)[14]. X denotes the continent of sale, number of bottles in the lot, bottle size, and a dummy for whether or not the designated lot consisted of an entire case of wine. δ represents individual bottle fixed effects, λ is a time variable with quarterly fixed effects, and γ represents the fixed effects for each auction house. The auction house fixed effects control for differences in the transaction costs of each auction house. It is also important to note that the fixed effects for each auction house and time period absorb the effects of the binary variables, *House* and *Time*. The simplified model is shown below:

$$\log(Y_{ist}) = \alpha_0 + \beta_1 X_{ist} + \beta_2 \delta_i + \beta_3 \gamma_s + \beta_4 \lambda_t + \beta_5 House_s \bullet Time_t + \varepsilon_{ist} \qquad (5.2)$$

The coefficient of interest, β_5, determines whether auction houses receive lower sales prices after it is discovered that they sold or attempted to sell counterfeit wine, where:

$$\beta_5 = [(\overline{y}|S = 1, T = 1) - (\overline{y}|S = 1, T = 0)] - [(\overline{y}|S = 0, T = 1) - (\overline{y}|S = 0, T = 0)]$$

$$= \Delta(\overline{y}|treatment) - \Delta(\overline{y}|control)$$

One potential criticism of the difference-in-difference model is that it fails to control for outside factors that could also influence price (e.g. the amount of pre-auction advertising, available information about the consigner, auction attendance). An auction including rarer wines may garner more attention from consumers, increasing the price of all wines offered at the sale. Lowering one's certification standard is one method to increase the inventory of rare wines available for sale.

[14] A Modified Park Test suggested that a generalized linear model using a gamma distribution may be preferable to the log-linear model presented in Section 5.2. The results of this model are reported in the appendix. Given that there are only minor differences between the estimates of the two models, I choose to direct most attention to the more basic OLS model.

5.1.2 Triple Difference Price Model

Following the theoretical framework of Gruber (1994) and Chetty et al. (2009), I include a control group of less expensive bottles to create a triple difference regression to mitigate the potential effects of the factors mentioned above. Wines below a certain value are thought to be less attractive to counterfeiters, implying differences in reputation between auction houses should not affect prices.

The DDD model includes a third binary variable, fake, indicating whether or not the wine is likely to be counterfeited.

$$Y_{0st} \text{ if } Fake = 0$$
$$Y_{1st} \text{ if } Fake = 1$$

Using the same fixed effects as the difference-in-difference regression model, price is calculated as:

$$\log(Y_{ist}) = \alpha_0 + \beta_1 X_{ist} + \beta_2 \delta_i + \beta_3 \gamma_s + \beta_4 \lambda_t + \beta_5 Fake_i + \beta_6 House_s + \beta_7 Time_t + \beta_8 Fake_i \bullet House_s + \beta_9 Fake_i \bullet Time_t + \beta_{10} House_s \bullet Time_t + \beta_{11} Fake_i \bullet House_s \bullet Time_t + \varepsilon_{ist} \quad (5.3)$$

under the triple differences-in-differences model. In a similar manner to Equation 5.1, the bottle, house, and time fixed effects negate the vales for *Fake*, *House*, and *Time*. The triple difference equation thus becomes:

$$\log(Y_{ist}) = \alpha_0 + \beta_1 X_{ist} + \beta_2 \delta_i + \beta_3 \gamma_s + \beta_4 \lambda_t + \beta_5 Fake_i \bullet House_s + \beta_6 Fake_i \bullet Time_t + \beta_7 House_s \bullet Time_t + \beta_8 Fake_i \bullet House_s \bullet Time_t + \varepsilon_{ist} \quad (5.4)$$

For this model, β_8 is the main variable of interest used to examine changes in equilibrium prices after a counterfeiting scandal:

$$\beta_8 = [(\overline{y}|I = 1, S = 1, T = 1) - (\overline{y}|I = 1,\ S = 1, T = 0)]$$
$$- [(\overline{y}|I = 1, S = 0, T = 1) - (\overline{y}|I = 1,\ S = 0, T = 0)]$$
$$- [(\overline{y}|I = 0, S = 1, T = 1) - (\overline{y}|I = 0,\ S = 1, T = 0)]$$
$$- [(\overline{y}|I = 0, S = 0, T = 1) - (\overline{y}|I = 0,\ S = 0, T = 0)]$$
$$= \Delta(\overline{y}|treatment) - \Delta(\overline{y}|control)$$

This value is similar to β_5 from Equation 5.2.

5.2 Empirical Model for Auction House Quantity Supplied

The quantity (*M*) of each wine an auction house sells during each time period is determined by a combination of individual house characteristics (reputation, commission fees, customer service) and the quantity of all bottles offered in the market (determined by market supply and demand). Without further information, any change in the number of bottles auctioned by a single house after a disclosure of negative information could simply be the result in the change in market forces affecting all firms. As with the model for prices described in the previous section, the inclusion of sales from auction houses not accused (*House*=0) is expected to control for any changes that affect the entire market.

5.2.1 Difference-in-Difference Quantity Model

By restricting the model to only include sales from wines likely to be counterfeited, the quantity of wine *i* sold by house *s* during time *t* can be estimated as:

$$M_{ist} = \alpha_0 + \beta_1\delta_i + \beta_2\gamma_s + \beta_3\lambda_t + \beta_4 House_s + \beta_5 Time_t + \beta_6 House_s \bullet Time_t + \varepsilon_{ist} \quad (5.5)$$

where $\delta, \gamma,$ and λ again represent the fixed effects of each bottle, auction house, and time period.[15] As with the models for prices specified in Section 5.2, the fixed effects for each house and time period capture the values for *House* and *Time.* A simplified version of equation 5.5 is specified below:

$$M_{ist} = \alpha_0 + \beta_1\delta_i + \beta_2\gamma_s + \beta_3\lambda_t + \beta_4 House_s \bullet Time_t + \varepsilon_{ist} \tag{5.6}$$

In this equation, β_4 is our coefficient of interest used to determine changes in the quantity of bottles sold following the discovery that an auction house has authenticated counterfeit wine.

5.2.2 Triple Difference Quantity Model

The addition of a set of wines unlikely to be counterfeited (*Fake*=0) may help control for potential changes in house specific characteristics that don't involve reputation. An auction house that improves its seller experience (e.g. lower commission fees, minimum price guarantees, payment schedule) may attract more sellers. For each wine, the quantity of bottles (*M*) sold in each time period is estimated as:

$$\begin{aligned} M_{ist} = \alpha_0 + \beta_1\delta_i + \beta_2\gamma_s + \beta_3\lambda_t + \beta_4 Fake_i + \beta_5 House_s + \beta_6 Time_t \\ + \beta_7 Fake_i \bullet House_s + \beta_8 Fake_i \bullet Time_t + \beta_9 House_s \bullet Time_t \\ + \beta_{10} Fake_i \bullet House_s \bullet Time_t + \varepsilon_{ist} \end{aligned} \tag{5.7}$$

Following the previous models, the values for $\delta, \gamma,$ and λ capture the effects of *Fake, House, and Time*; therefore, Equation 5.7 can be simplified to:

$$\begin{aligned} M_{ist} = \alpha_0 + \beta_1\delta_i + \beta_2\gamma_s + \beta_3\lambda_t + \beta_4 Fake_i \bullet House_s + \beta_5 Fake_i \bullet Time_t \\ + \beta_6 House_s \bullet Time_t + \beta_7 Fake_i \bullet House_s \bullet Time_t + \varepsilon_{ist} \end{aligned} \tag{5.8}$$

[15] One minor change from the model specified in Equation 5.1 is that each time period now represents a half-year as opposed to a quarter.

The values for β_7 in the DDD model determine whether auction houses have a decrease in the quantity of rare and expensive wine supplied following the disclosure of negative information. This information can be used in combination with the results from Equation 5.4 to test whether or not auction houses see a shift in the demand curve as a result of a counterfeiting disclosure. A reduction in comparative prices coupled with a decrease (or unchanging) quantity of bottles supplied is indicative of a downward shift in the demand curve.

6. Data

The estimation strategy I employ requires a longitudinal dataset containing sales information before and after the treatment date. In total, I examine the sales data from 2005-2015 for the ten largest auction houses: Acker Merrall & Condit (Acker), Bonhams & Butterfields (Bonhams), Christies, Hart Davis Hart (HDH), Heritage, K & L, Morrell & Company (Morrell), Sothebys, Spectrum, and Zachys. In addition to price, other variables in the dataset contain information regarding the producer, vineyard, vintage, auction date, sales location, bottle size, and lot size for each bottle sold.

6.1 Selection of Wines

To minimize potential region and producer specific effects, I restrict attention to sales from wines produced in Bordeaux or Burgundy. Wines produced from these regions are often prized for both their quality and aging potential (Robinson, 1999). Given their demand, wines from Bordeaux and Burgundy are most targeted by counterfeiters (Kapon, 2006). Wines from other regions are also counterfeited[16], but the list of potential marks is limited.

[16] Wines from other regions targeted by counterfeiters include Screaming Eagle from California's Napa Valley and Penfolds Grange from Australia.

A further restriction limits attention to wines from highly rated vintages with at least 50 sales observations[17] from 2003-2015 in the effort to capture pre and post-intervention sales prices. A highly rated vintage was given a rating of at least an 85/100 by tasting experts (Leve, 2016; Parker, 2015). These wines often take longer to mature, appear more frequently at auction, and have a higher likelihood of being purchased as an investment. For example, a bottle of 1981 Lafite (a vintage rated 84 by expert Jeff Leve) can be found for under $400 while a case of the 1982 vintage (with a rating of 96) recently sold for $55,000 ($4,600 a bottle) (Leve, 2016). In total, the vintages selected range from 1929-2002 and include 116,094 observed sales of 286 unique wines.[18] These wine account for a significant percentage of all auction house sales. In 2011, the sales of the included wines represented more than a third of the total sales revenue for all auction houses combined.

6.2 Description of Independent Variables

Each of the 286 wines is categorized by a unique combination of producer, vineyard, and vintage. These individual wine fixed effects are used in the regression models to control for differences in value. In addition to capturing all producer, vineyard, and vintage effects, the 286 fixed effects dummies also capture region specific effects.

The model also controls for auction date, location, and bottle size, and lot size. Dates of sale are grouped by quarter to control for changes in price not associated with counterfeiting accusations (i.e. market supply and demand). The value of collectible wine tends to follow the

[17] Sales observations are measured in lots. Lots usually contain between 1-12 bottles of the same wine, and multiple lots of the same wine can be offered at a single auction.

[18] Due to the sales criterion and required designation of whether or not a wine is likely to be counterfeited, not all vintages could be included for each producer. For instance, Margaux is only included for 11 of the 20 Bordeaux vintages. A listing of all wines and vintages is included in the appendix.

global economy. Prices generally rise during economic expansions and wane during contractions.

Discrepancies between import fees, taxes, shipping, or other non-bottle specific costs may result in price differences between locations. Thus, I include dummy variables for each location (Asia, Europe, Internet, and North America).

Size dummies control for the volume of wine per bottle. 92% (N=106,255) of all sales observed were for 750ml bottles while the other 8% (N=9,809) were for larger format bottles (1.5L and larger). Apart from the difference in volume, larger format wines are expected to age at a slower pace and are also often shown as centerpieces to a collection. These bottles generally sell for a premium beyond that expected by volume alone (i.e. the larger the bottle, the higher the price/ml).

The number of bottles in each lot separated into two variables. The first variable controls for the number of bottles in the lot. Generally, larger lots sell for a higher price per bottle than lots with fewer bottles. It is expected that wines from the same lot will have a similar storage history and taste more homogenous than wines from separate lots.

The second lot size variable is a binary variable designating whether the sale was part of a full case (or more) of twelve 750ml bottles, six 1.5L bottles, or two 3L bottles. Cases of wine often come in decorative wooden crates providing aesthetic value for consumers wanting to display their collection. A full case may also signal that the wine was originally purchased as an investment, increasing the chances that the bottles have been properly stored. As such, there is generally a significant increase in price for wines sold by the case.

6.2.1 Determining Wines Likely to be Counterfeited

The triple difference regression framework classifies each wine into one of two groups depending upon whether or not the wine has been or is likely to have been counterfeited. Counterfeiters may target certain wines based upon their value, perceived risk of discovery, or effort required to generate the counterfeit. Wines were labeled as likely to be counterfeited based upon lists provided by third-party authentication experts (City National Bank, 2015; Downey, 2012; Gray, 2014; Haughney, 2015; Robinson, 2014). In general, price appears to be the major factor for determining whether or not a wine is considered a potential for counterfeiting. William Edgerton, a third-party expert, has noted that, "any wine over $1,000 is fair game [for counterfeiters] " (City National Bank, 2015). [19]

It is likely that the $1,000 price point is not a firm threshold for counterfeiters. In the effort to create a rigid cut-point differentiating wines likely to be counterfeited from those unlikely to be counterfeited, all wines with a maximum annual average price between $601-$999 were excluded from the triple difference regressions.

Wines deemed unlikely to be counterfeited have a maximum annual average of $600 for all years between 2006 and 2012.[20] Given that these wines are unlikely to be faked, auction house reputation should only have a minimal effect (if any) on the price. Appendix 4 provides a list and designation (by counterfeit group) of the wines included in the regression model.

[19] Wines considered likely to be counterfeited from the Bordeaux region include various vintages of Lafite, Latour, Mouton, Margaux, Haut Brion, Petrus, Cheval Blanc, Palmer, Lafleur, Latour a Pomerol, Trotanoy, and La Mission Haut Brion. From Burgundy, producers include Domaine De La Romanee Conti (DRC), Ponsot, Dujac, Henri Jayer, and De Vogue.

[20] The control group includes less valuable vintages from the group above as well as wines from other respected producers (including Leoville Las Cases, Montrose, Ducru-Beaucaillou, Calon Segur, Leoville Barton, Leoville Poyferre, Lynch Bages, Talbot, La Lagune, Cos D'estournal, Palmer, and Beychevell from Bordeaux and Jadot, Leroy, Louis Latour, Bonneau, Mommessin, and Faively from Burgundy).

6.3 Description of Dependent Variables

The regression models used in this analysis focus on the effects of a counterfeiting disclosure on an auction house's equilibrium sales prices and quantities. The price of each bottle is first calculated by adding any buyer's commission fees to the hammer price of each lot. This number is then divided by the number of bottles in the lot to find the per bottle price.[21] In the regressions using price as the dependent variable, each observation represents the sale of a single lot.

For the regressions with quantity as the dependent variable, each observation represents the number of bottles of wine *i* that were sold by one auction house during a single time period. Unlike price, the values for quantity do take lot size into consideration.

For each wine included in the study, I observe all sales from the selected auction houses. Provided two minor assumptions, it is important to note that the data on quantity does not suffer from any censoring issues. The first assumption states that there are still authentic bottles of each wine sold in the market. I also assume that each auction house is willing to sell authentic bottles of each wine selected. By making these assumptions, we have no missing observations. If an auction house doesn't sell a wine during the specified time period, the value for M_{ist} is 0. Without these assumptions, we would not be able to differentiate between auction houses with no sales and those not in the market.[22]

[21] Given a buyer's commission fee of 20%, the price per bottle from a 12 bottle lot with a hammer price of $5,000 would be $500. ($5,000*1.2)/12

[22] For instance, these assumptions prevent the study of the counterfeit Ponsot wines from the 2008 counterfeiting discovery. Because the authentic wines were never produced, some houses may have been unwilling to sell any bottles. I try to avoid this issue by limiting the selection to only include wines with at least 50 sales observations.

6.4 Disclosure Dates and Selection of Houses Implicated

2008 Study

On April 23, 2008, Acker Merrall & Condit was sued by Bill Koch for selling counterfeit bottles consigned by Rudy Kurniawan in 2005-2006. Two days later, only minutes before an Acker auction, bottles consigned from Kurniawan were withdrawn from sale after they were declared counterfeit (Hellman, 2008). The identified winemaker for those bottles was at the auction and publicly declared he never produced those specific wines. For this case, Acker is the treatment house (House=1) accused of selling/authenticating counterfeit wine and Bonhams & Butterfields, Hart Davis Hart, Morrell & Company, Sotheby's, and Zachys are the control houses not implicated (*House*=0).[23] The 2008 study includes sales from April 1, 2007-March 31, 2009. Because no auction house sales were held between April 23-April 25, 2008, April 23, 2008 is used as the treatment date for the DD and DDD regressions.

2012 Study

The second disclosure studied in this manuscript involves the arrest of Rudy Kurniawan in March 2012 for counterfeiting wine. To test whether or not auction house reputation suffered after this scandal, I treat Acker, Christies, and Spectrum as the treatment group of auction houses accused of selling counterfeit bottles. Acker publicly accepted consignments of Kurniawan's wines from 2006-2008 and continued to privately accept consignments until at least 2011, Christie's accepted consignments from 2006-2012, and Spectrum held the February 2012 auction that helped lead to Kurniawan's arrest (Hellman, 2017; Hernandez, 2012; Wine Berserkers, 2012). Bonhams & Butterfields, Hart Davis Hart, Heritage Auctions, K&L Wines, Morrell &

[23] Christies is not included in this set of regressions because of a similar, but less publicized, withdrawal of counterfeits in 2007.

Company, Sotheby's, and Zachys are the control houses not implicated (*House*=0). The time dummy in the model denotes whether the sale occurred before or after March 8, 2012 (the date Rudy Kurniawan was arrested). Sales observations range from April 1, 2011-March 31, 2013.

7 Results

7.1 Equilibrium Prices Following the 2008 Disclosure

The base 2008 Difference-in-Difference regression model covers 8,785 sales of 178 unique wines, an average of 49.2 sales per wine. As indicated in Column 1 of Table 3, the results appear contrary to expectations. Equilibrium prices for bottles auctioned by Acker, Merrill and Condit, the wine house implicated in the 2008 fraud allegations, rose by more than 8% (*Time* and *House)* compared to those from auction houses not accused in the year following the disclosure.

A possible factor for the base model findings may be due to Acker's timely expansion into the Asian market after Hong Kong's February, 2008 removal of beer and wine duties (See Table 4). Asian auction prices during the model time period were generally higher than their European, North American, and internet counterparts, possibly a result of their relative market infancy and relatively limited susceptibility to reputational changes. Given that both the 2008 and 2012 fraud discoveries were primarily based on North American auctions, it is possible that consumers in other locations were less aware of those charges and thereby less prone to make changes in the prices they were willing to pay. For this reason, Column 2 in Table 3 only reflects sales conducted from North America. Yet, restricting the model to the North American market (Column 2) found Acker's prices rose nearly 9% compared to other houses in the year following the counterfeiting scandals.

Columns 3 and 4 in Table 3 limit the regression model to sales of cases/large-format bottles and very expensive ($2,000+) bottles, respectively. Twelve-bottle cases, large-bottle wines and expensive wines are often purchased for investment rather than consumption. As such, many of these lots are accompanied by original purchase receipts and a full-storage history to address concerns that the bottles may be either counterfeit or have been poorly stored. Investors may be more concerned with auction house reputation if wines would require re-authentication for future resale. The coefficients for β_5 (*Time* and *House*) in both Columns 3 and 4 are smaller than those from the Base DD and North American Columns, but remain both positive and significant. Acker's prices for cases and large formats rose by 5% after April 2008 while the prices for very expensive bottles rose by more than 7%. To study the full equilibrium effects of the disclosure, we must also look at the equilibrium quantities of bottles sold for these specifications. The findings are presented in Section 7.2.

Another rationale for the increase in equilibrium prices involves unobservable factors unique to each individual auction. Acker was often praised for providing customers with highly detailed catalogs, pre-auction wine tastings, and gourmet dinners that may have enhanced consumer experience (Kapon, 2007; McInerney, 2008). Other unobservable characteristics include shipping fees, bidder excitement, auction timing (e.g., currency fluctuations), and personal relationships with the auctioneer. Including wines unlikely to be counterfeited should help control for the unobservable characteristics of each auction.

7.1.1 Triple Difference Estimates (DDD)

The 2008 triple difference regression (Table 5) includes an increased sample size, bringing the totals to 286 wines and 14,189 observed sales. Column 1 of Table 5 displays the estimates for the full DDD model. These figures show positive, yet insignificant, results for β_8 (*Time, Fake,* and *House*). The coefficient for β_5 (*Time* and *House*) appears to capture many of the unobserved effects of the DD model, noting that prices for all wines sold by Acker rose compared to those sold by other auction houses.

As with the DD model, this finding can at least partially be attributed to Acker's position as a leading house in the Asian market. Including the data from all sales, from 2007-2009, auctions in Asia generated sales prices more than 20% higher than auctions elsewhere. Column 2 limits the model to only include North American sales. While the DDD coefficient is smaller than that found in the DD model for North American sales, it remains significantly positive. The response is also more than three times larger than that of the base model. In North America, Acker's price for wines likely to be counterfeited rose by more than 4% in the year following the April 2008 events compared to other auction houses.

Column 3 limits the analysis to sales of wines from Burgundy. Burgundy represents a much smaller geographical region than Bordeaux and only accounts for 13% of all sales in the dataset.   However, due to its limited production and increasing popularity among collectors, many of the most expensive wines in the world originate from this region. One particular example, Domaine de la Romanee-Conti's Romanee Conti, was also the favorite of Rudy Kurniawan (Hernandez, 2012).[24] As the wines pulled from auction were also Burgundy, it is possible prices of Burgundy sold by Acker would decrease more than the prices of other wines after April 2008. The results show the opposite to be true. In the DDD regression (Time, Fake, and House) including wines unlikely to be counterfeited, prices of Burgundy auctioned by Acker rose by nearly 12% relative to those sold by the control houses.[25]

---

[24] Due to his affinity for DRC, Rudy earned the moniker Dr. Conti. It was even included as his alias during trial in 2012.

[25] It should be noted that this model produces results with the poorest fit among the six DDD equations. The average number of observations per wine (22.5) is significantly lower than that of the next lowest model (Equation 2 averages 38.4).

Column 4 reflects sales of older wines produced before 1986.[26] Experts have noted a higher ratio of counterfeit-to-authentic bottles for older vintages, leading to speculation that consumer response to fraud may be stronger for this subgroup (Bell, 2014; City National Bank, 2015; Hirsch, 2012). This estimation does result in a negative, albeit insignificant, coefficient for the DDD term in support of this theory.

All models in this analysis rely on the assumption that changes in sales prices are consistent for all regions (i.e. no arbitrage). Even for auctions held only a few weeks apart, product availability and fluctuation in exchange rates may cause uneven shifts in prices the base DDD model cannot capture. The base model uses continent and quarterly fixed effects, and, if prices in one part of the world changes at a different pace than others, the model estimates may be confounded. Through the use of semiannual continent fixed effects, Column 5 tests whether price follows similar trends worldwide. The strong resemblance between these estimates and those for Column 1 supports this assumption, suggesting the price differential between sales locations remained consistent over time.

Column 6 limits the regression to the sale of bottles produced after 1994.[27] This helps provide a comparison to the results of Column 4 to test whether reputation is more important for older wines. The coefficients for the DDD term are again insignificant, but the reduction in significance for many of the auction house fixed effects does hint that consumers may be more

[26] Laurent Ponsot has claimed that as much as 80% of pre-1980 Burgundy on the market is counterfeit. Restricting the model to pre-1980's vintages produces similar results to Column 4, but with much larger standard errors. Both 1982 Bordeaux and 1985 Burgundy are considered fully mature and have been praised as outstanding vintages by critics. Given the high prices bottles from these vintages command, they are also commonly counterfeited.

[27] I chose this cutoff as bottles from 1990 (the last great vintage before 1995) were finally reaching the prime drinking window while the newer vintages were still likely to improve.

concerned with house reputation for older wines. A study of the equilibrium quantities is expected to provide additional evidence in regards to this theory.

7.2 Equilibrium Quantities

In order to determine whether auction houses suffer a loss in reputation following a counterfeiting disclosure, we must also study the quantity of wine sold by each house. To account for the differences in frequency in which certain wines appear at auction, a logarithmic transformation was performed on the dependent variables for the OLS regressions, log(1+M_{ist}).

For these regressions, it is important to remember that quantity is a count variable with a lower bound of 0. The use of linear regression models can provide good estimates for the average partial effects, but these estimates could also prove problematic, given that the estimates can produce negative values for quantity (Wooldridge, 2010). As such, each of the equations was also run using a negative binomial regression model (NegBin II). Along with Poisson regression models, negative binomial models are often used when the outcome of interest is a count variable. These models use an exponential functional form and have the added advantage of being able to directly estimate the effects of each model, E(y|x) without the logarithmic transformation used with the OLS model, E[log(1+y)|x].

Negative binomial models are often preferred to Poisson models when the data suffers from overdispersion (i.e. the conditional variance is greater than the mean). Specifically, the NegBin II model assumes that the amount of overdispersion is expected to increase with the mean. For the regression estimating the quantity of bottles supplied, it is expected that the NegBin II model specification will provide more accurate and interpretable estimates than an OLS or Poisson model.

7.2.1 Equilibrium Quantities Following the 2008 Disclosure

Table 8 displays the difference-in-difference estimates for changes in the quantity of bottles Acker sold following the 2008 counterfeiting study. Given the results for the prices models, the findings for quantity sold are not surprising. While statistically insignificant, the results from the base NegBin II DD model in Column 2 suggest Acker was able to increase the quantity of wines sold following the discovery. When combined with the results from Column 1 of Table 3, these estimates indicate an increase in the demand curve for wines offered by Acker that are valuable enough to counterfeit.

Column 3 displays estimates from a model restricting observations to wines produced before 1986. While the estimates for the change in equilibrium quantities of older wines are still positive, they are only a seventh of the size of the base model. These findings provide further evidence about the relationship between product uncertainty and firm reputation. Of all the model specifications performed on the prices following the 2008 event (Tables 3 and 5), the only estimate with a negative coefficient for the treatment effect dealt with older wines.

Columns 4 of Table 8 focuses on the sales of cases and large format bottles The negative coefficient for the treatment variable suggests that, like consumers of older wines, buyers of full cases and large format bottles may be more sensitive to the disclosure of negative information. The increase in equilibrium prices found in Table 3 may have been a result of a decrease in the supply curve of these wines, and not from an outward shift in demand curve (as found in the base model).

The results from a model restricting the selected wines to only include sales from expensive $2,000+ wines (Column 5) in largely reflects the base model, noting an increase in Acker’s demand curve for these wines in the year following the April 2008 discovery.

7.2.2 Triple Difference Estimates

Estimates from models including sales from wines unlikely to be counterfeited are displayed in Table 9. Interpreting the treatment effect for the DDD models on quantity supplied is not as straightforward as with the DD models for supply or DDD models for prices. It is likely that consignments of wine unlikely to be counterfeited are at least partially dependent upon house reputation.

When individuals decide to sell wine at auction, they generally consign all bottles (to be sold during that time period) through a single auction house due to negotiation costs and consignment fees (larger consignments often have lower fees). Frequently, these consignments contain wines of varying price points. Some wines may be valuable enough to counterfeit while others are not. Ceteris paribus, a decrease in auction house reputation is likely to affect the quantity of all bottles supplied. Any reputation loss lowers the expected price sellers receive for wines likely to be counterfeited, decreasing the total benefit for the entire collection, and thus, the likelihood that an auction house is offered the consignment. For this reason, it is not clearly apparent how to interpret treatment effects with the DDD models. If we believe that the supply of wines unlikely to be counterfeited is not independent from the supply of wines likely to be counterfeited, it may be best to interpret the treatment effect by adding the coefficient for β_6 (*Time* and *House*) to β_7 (*Time*, *Fake*, and *House*).

Using this method to calculate the treatment effect, we see a strong resemblance between the DD and DDD models. The base model (Column 2) shows a 9% increase in the quantity of bottles supplied by Acker following the disclosure. When combined with the 1% increase in equilibrium prices found with the base DDD prices model (Column 1 of Table 5), it is evident that Acker did not suffer a loss in reputation following the 2008 disclosure.

A model restricting the sales to North America (Column 3) does shows a 24% reduction in the quantity of bottles sold by Acker in the year following the disclosure. Yet, it is expected that the real cause for this result was due to Acker's choice to focus on expanding sales in Asia. The base DDD price model found that auctions in Asia generated sales prices more than 20% higher than auctions in other locations, and the results from this regression and Table 4 simply show Acker chose to focus on the Asian marketplace more than nearly all other auction houses.

7.3 Equilibrium Prices Following the 2012 Disclosure

The 2012 difference-in-difference base equation (Column 1 of Table 6) covers a total of 15,471 sales observations of 178 unique wines. The most important finding is the significantly negative coefficient for the treatment effect, *Time* and *House*, noting that relative prices for Acker, Christies, and Spectrum dropped by more than 4% in the year following Mr. Kurniawan's arrest. Unlike the 2008 case, the evidence obtained from the formal arrest of Rudy Kurniawan may have been too much for consumers to overlook or ignore. The 2008 study involved significantly fewer counterfeit bottles and was based on allegations of fraud by the purported winemaker rather than criminal charges from law enforcement (Hellman, 2015; Wallace, 2012).

Although, as in 2008, bottles sold in Asia sold for a higher price than those auctioned elsewhere, the price differential is much smaller than those from the 2008 case. For example, sales prices from North American auctions were 18% lower relative to Asia for the 2008 study (Table 3), but only 9% for the 2012 study (Table 6). The results displayed in Table 4 suggest a possible explanation. As more auction houses focused on the Asian market, the increase in the quantity of bottles supplied likely contributed to the lower price premiums.

As with the 2008 DD regression, Column 2 of Table 5 restricts sale price data to the North American market. The results, however, are very different. Whereas prices in the North

American sustained the highest relative gains in the 2008 study, they show the highest relative losses in the 2012 study of more than 7%. Column 3 also indicates a significant drop of nearly 7% for the prices of cases and large formats sold by the houses accused of accepting counterfeits.

In addition to sales of cases and larger volume bottles, buyers of the most valuable wines may behave in a different manner due to higher prices and increased rarity. Column 4 explores this possibility to see if buyers of the most expensive bottles (having an average annual sales price of at least $2,000) were more sensitive to the counterfeiting allegations. The coefficient for this Column is the only insignificant finding for the treatment effect in the DD regressions. The relative price decrease for these very expensive wines was less than 1.5 %.

The DD model does not control for outside factors affecting the prices of all wines, and it is possible that the houses implicated could have reduced advertising expenses or suffered a loss in auction turnout as a result of the negative information. As with the 2008 case, the full DDD model will address this consideration.

7.3.1 Triple Difference Estimates (DDD)

The base triple difference regression (Table 7) covers 25,078 observations, representing the sale of 170,711 total bottles and 286 different wines. This regression, like the DD models, indicates a decrease in equilibrium prices for the wine houses associated with the 2012 disclosure. The inclusion of an additional control group for wines unlikely to be counterfeited produced a smaller relative loss (3%) than the base DD regression (4.4 %), but the result remains significant.

Column 2 restricts sales to North America, and finds that prices from the treatment group for all wines (*Time* and *House*) showed a significant relative loss of 6+ %. The coefficient for wines likely to be counterfeited (*Time, Fake,* and *House*) indicates an additional relative loss of only 0.1% compared to the treatment group of wines unlikely to be counterfeited. Some buyers may have chosen to boycott all sales from the auction houses implicated.

Column 3 tested if there was a stronger reaction for older bottles of wine, using the same model as Column 4 from Table 5. As noted previously, many of the wines Mr. Kurniawan consigned were from older vintages. A lack of production records and anti-counterfeiting measures make many older wines an easy target for counterfeiters. In this model, I find a treatment effect of more than 8% following the disclosure. These results are fairly consistent between Asia and North America, with both regions showing a similar relative drop in equilibrium prices for wines sold by Acker, Christies, and Spectrum.

Another explanation for the price drop of Acker, Christies, and Spectrum is through a change in average observable characteristics during the time period. Once Mr. Kurniawan was arrested, the average bottle condition may have suffered for the houses if no other counterfeiters were able to supply bottles of the same observable quality. Instead of punishing the auction houses implicated, the price drop may be due to a change in the average observable condition for each wine.

To test whether changes in observable bottle condition caused the drop in prices, Column 4 uses sales of cases or large format bottles while Column 5 contains wines produced after 1994. The tighter dispersion of auction house and geographic fixed effects as well as the improved model fit suggest bottle condition is more uniform for cases, large formats, and newer wines. The significant relative drop in prices of more than 6% for case and large format sales largely supports the base model, implying that a change in bottle condition was not the cause for the price decrease.

Column 5 limits the regression to bottles produced after 1994 and uses the same setup as that restriction from the 2008 study. This regression model also allows us to test for differences in purchasing behavior for wines from newer vintages when these results are compared to those from Column 3. Many producers began taking significant steps to prevent counterfeiting in the latter stages of the 20th Century (McCoy, 2007). These techniques began with specific bottle etchings or label identifiers and have continued to become more sophisticated with each passing year.

As the results show, sales prices for newer bottles were not affected to the same degree as sales of older wines. Knowledge of anti-counterfeiting measures may have given consumers increased confidence in the authenticity of younger wines, regardless of seller reputation. Newer wines were offered during the Cellar I, Cellar II, and subsequent Kurniawan auctions, but most of the contested bottles were from older vintages (Hellman & Frank, 2009; Kapon, 2006). While the 2008 comparison between older and younger wines yielded insignificant results, both studies provide evidence that reputation is more valuable when buyer uncertainty increases.

The retail sector may also be a cause for the diminished DDD coefficients found with younger wines. Relative Differences in auction house fixed effects are smaller for newer vintages, likely due to both similarities in observable characteristics and availability of these wines in the retail sector. The increased availability means that the price differences between retail and auction are likely smaller for newer vintages, leaving less room for reputational losses in the model.

Other regressions (including those using semiannual region controls) were also tested, and the results were near identical to their base model counterparts.

7.4 Equilibrium Quantities Following the 2012 Disclosure

Results from the effects of the 2012 disclosure on equilibrium sales quantities are displayed in Tables 10 and 11. While insignificant, the estimates for the base DD NegBin II model in Column 2 show that Acker, Christies, and Spectrum sold 6% fewer bottles of wines likely to be counterfeited in the year following the arrest of Rudy Kurniawan. Column 3 shows a decrease of nearly 9% in the quantity of older bottles sold, and Column 5 shows a highly significant reduction of more than 32% in the quantity of very expensive $2,000+ bottles sold.

The only regression showing a positive change in quantity dealt with the sales of full cases and large format bottles (Column 4). A time lag between the date bottles are consigned and when they are sold may mitigate the size and significance of our findings from Columns 2-5. The process of cataloging, shipping, authenticating, and advertising a consignment often takes months. Seller's midway through the auction process when Kurniawan was arrested could have decided to continue with a planned sale due to the additional costs associated with starting the process over again using another firm. These costs are likely higher for shipments of full cases and large format wines, due to the additional shipping costs associated with each lot.

To test whether a potential lag could be affecting the estimates for the treatment effect, I compared sales from March 2011-March 2012 with those from September 2012 (six months after Kurniawan's arrest) to September 2013. The results, displayed in Column 4 of Table 9, show a significant decrease in the quantity of wines sold by Acker, Christies, and Spectrum. A model using sales from September 2011-September 2012 as the control time period also yielded similar results. In general, the DD results show a decrease in the quantity of all wines sold by the three auction houses implicated. This is evidenced by a negative coefficient for *Time* and *House* across nearly all models.

7.4.1 Triple Difference Estimates Following the 2012 Disclosure

Table 11 provides the estimates using a DDD model and shows. These estimates are again largely supportive of the DD regressions displayed in Table 10. The base NegBin II model (Column 2) shows a 6% decrease in equilibrium sales quantities for the houses implicated while the model using lagged sales notes a decrease of nearly 20% from September 2012-September 2013 (Column 4).

8. Discussion and Concluding Remarks

8.1 Observations from the 2008 Study

The results for the 2008 analysis suggest Acker did not suffer a loss in reputation after it was discovered the auction house authenticated 22 lots of counterfeit wines. These findings closely mirror the media coverage describing the events, which generally exonerated Acker for its role in the sale of counterfeits (Hellman, 2008; Hellman & Frank, 2009; McInerney, 2008; Wise, 2008). This lack of retribution may have provided encouragement for other auction houses (Christies and Spectrum) to accept bottles from Mr. Kurniawan (Wine Berserkers, 2012). [28]

There are a few possible explanations for Acker's ability to maintain its reputation after this incident. Individuals may have felt this to be an isolated incident, and thus below the threshold needed to reconsider house reputation.

Another explanation involves the personal reputation of Rudy Kurniawan. After the April 2008 auction, Acker publicly cut ties with Rudy Kurniawan and John Kapon (Acker's CEO) offered an apology on a popular wine forum (Squires, 2008). Other individuals also came to

[28] Mr. Kurniawan largely disappeared from public events after this event and began using middlemen in order to hide his consignments from the public. These middlemen include Darmawan Saputra, Antonio Castanos, Marc Lazar and Richard Brierley. However, it has been discovered that the auction houses often knew Kurniawan was the source behind bottles consigned by these individuals.

Kapon's aid, providing personal testimonials on his behalf. While the stories were not widely published until years later, it appears that some collectors may have become weary of wines consigned by Rudy Kurniawan before the April 2008 auction (Barzelay, 2012). Because Rudy Kurniawan consigned so much wine through Acker, it is possible that consumers associated Kurniawan's personal reputation with Acker's. Unless the consignor was specifically named before the auction, some consumers may have assumed that Kurniawan was the source for all rare wines offered by Acker. This association could have depressed Acker's prices prior to the 2008 auction, in turn artificially increasing the results for the post-intervention change in prices.

8.1.1 Acker's Guarantee as Insurance

For many years, Acker has offered a money back guarantee for all wines purchased at auction. Previous research on baseball cards suggests that having a product guarantee may have cushioned Acker from a drop in prices following the 2008 discovery of counterfeit wine (Haley & Van Scyoc, 2010; Jin & Kato, 2006). Auctions with buyers insurance were able to generate a significantly higher final sales price. A study on used tractors also hinted that warranties could downplay the importance of reputations, although many of the results were of little or no statistical significance (Roberts, 2011). Acker's guarantee acts as a form of insurance for customers against counterfeit bottles, making the buyer feel more assured about the authenticity of each bottle offered.

Acker's guarantee was also likely a major reason why Rudy Kurniawan was able to continue consigning wines after the April 2008 discovery (Hellman & Frank, 2009). As was later discovered, Acker previously accepted a series of returns from The Cellar auctions due to suspicions that the bottles were counterfeit. (Hellman, 2012; Hernandez, 2012; Wallace, 2012). In exchange for returning the counterfeit bottles, individuals may have been asked to sign nondisclosure agreements to prevent the information from becoming publicly available. This concealment may have limited the damage to Acker's reputation while also allowing Kurniawan to continue consigning wine.

8.2 Observations from the 2012 Study

The arrest of Rudy Kurniawan and negative publicity surrounding the Spring 2012 Spectrum auction provide a stark contrast to the 2008 scandal. Unlike the 2008 case, the FBI investigation provided concrete evidence against Mr. Kurniawan and the auction houses with which he was associated. The regression estimates show Acker, Christies, and Spectrum suffered a loss in reputation as a result of the scandal. The DDD estimates for prices indicate a relative price decrease between 3-8% from March 2012-March 2013 for the auction houses implicated. During the same time period, equilibrium sales quantities decreased by between 6-9%. A model looking at sales quantities from September 2012-September 2013 found a significantly larger effect of more than 20%. Combined, these results indicate that Acker, Christies, and Spectrum suffered a downward shift in the demand curve caused by a loss in reputation as a result of the counterfeiting discovery.

8.3 General Findings and Study Implications

The cases studied in this manuscript track consumer behavior following two recent disclosures that auction houses had authenticated and offered counterfeit wine for sale. While the risk of fraud detection may be higher for auction houses selling a higher proportion of counterfeit wines, the discovery of a counterfeit bottle is a discrete event. It is difficult to predict when the next counterfeiting allegation will be made, but recent evidence suggests the punishment received by the auction houses after the 2012 case was not strong enough for all firms to significantly improve their authentication standards. Three new counterfeiting claims were brought against international auction houses from May 2015-April 2017; one of which involved counterfeit Burgundy authenticated by Acker Merrall & Condit (Gray, 2015; Lichfield, 2016; Millar, 2017; Wang, 2017; Wine Berserkers, 2012).

Meanwhile, the number of remaining authentic bottles of each collectible wine decreases every year as some are consumed. Coupled with a growing worldwide demand for wine, the market has experienced in a steep increase in prices for older wines. In 2003, a bottle of 1945 DRC Romanee-Conti could be purchased for less than $3,000. At the height of the market in 2011, a bottle of that same wine sold for $124,000. While the appreciation has been most spectacular for the rarest bottles, nearly all collectable wines have become more valuable (Figure 1).

An increase in the sales prices attracts more counterfeiters, and likely also increases consumer skepticism. It has been estimated that as much as 20% of the fine wine currently bought and sold worldwide is counterfeit (Krebiehl, 2017; Richard, 2013; Taylor, 2013). This ratio may only increase in the future as more counterfeit bottles are produced and authentic bottles are consumed. Results from the DDD model on prices (Table 7) note that the punishment received by auction houses was strongest for the wines most appealing to counterfeiters (older vintages and larger quantities). The negative coefficients for β_6 (Time and Fake) also note that consumers may have felt growing distrust for all auction houses. Increasing concerns regarding the authentication standards of all auction houses has led to the establishment of third-party firms that solely function as an authentication service.

8.3.1 Increasing Use of Third-Party Experts

In May 2016, Sotheby's auctioned 20,000 bottles from the collection of Bill Koch. In addition to the authentication offered by Sotheby's, these wines were also inspected by third-party experts (Bloomberg Markets, 2014). Nearly all of the 2,730 lots sold above their pre-auction estimates as Sotheby's CEO noted, "[buyers] were prepared to spend more money to acquire wines of impeccable quality (Meltzer, 2016)." Many of the rarest wines from the collection sold for more than double the expected price. Results from this auction (and others where the wines have been vetted by a neutral party) can be used in future studies to estimate the adverse selection discount consumers place on wines sold at auction.

If neutral party experts are shown to instill more consumer confidence than the auction houses themselves, we may eventually find a scenario where auction houses solely function as a connecting agent between sellers and consumers (à la ebay). The separation of sales and authentication has become widely accepted in other collectibles markets with adverse selection; namely coins, autographs, and sports cards.

8.4 Study Limitations and Future Projects

Due to its reliance on repeat sales, the model used in this study is unable to assess the effects of counterfeiting scandals on reputation for the rarest and most expensive wines. As the rarest bottles most often receive the most pre-auction publicity, it is possible that reputation is even more valuable for these wines. The results for the limited number of older wines in this data set suggest that the three auction houses associated with Rudy Kurniawan may have suffered a larger decline in prices and quantities sold (compared to the results) for the rarest wines after his arrest.

8.4.1 Future Projects

The results from the 2012 study suggest that all auction houses suffered a decrease in demand as a result of the counterfeiting discovery. While most of the decrease was likely due to outside forces, some of the decrease may have been caused by a decrease in the reputation of all firms. Two approaches have been identified to provide further evidence in future studies of counterfeiting. The first involves tracking private and retail sales over the same time frame to study if the frequency and price of private trades also change after each case of fraud. The second involves a similar comparison between wines authenticated by auction houses and those from third-party experts.

The growing use of third-party experts also affords a variety of other new study topics. One involves a deeper investigation on the underlying incentives experts face. The use of a third-party expert does not necessarily prevent fraud. If the individual buyer is not paying for the authentication, there may be incentives for an expert to authenticate counterfeit wine in order to increase future business (Hubbard, 1998). If the seller (or auction house) pays for the third-party authentication, we are left with the same potential issues consumers already face with the auction house authentication practices. Nonetheless, recent sales from collections authenticated by third-party experts obtained substantially higher prices than expected (Abernethy, 2017; Meltzer, 2016). Will this trend continue over time?

To date, most of the research conducted on the theory of credence goods has focused on markets where the expert faces differential treatment costs (Bonroy, Lernarie, & Tropeano, 2013; U. Dulleck & Kerschbamer, 2006; Emons, 1997). An expert performing car maintenance, even if it isn’t necessary, faces higher costs than if no maintenance was performed at all. Unlike the expert who provides a diagnosis and treatment, an authentication expert only performs a diagnosis (i.e. a certifying agent).

If certifiers are required to be truthful, it is possible to eliminate fraud in the market (Stahl & Strausz, 2017). Yet, requiring the certifier to be truthful may not always be applicable to real-world situations (Dranove & Jin, 2010). Some unique factors in the authentication of collectibles (i.e. technological advances, likelihood of product resale, possibility that the consumer utilizes multiple third-party experts) may allow for the derivation of certain equilibriums where third-party experts always provide the correct diagnosis, even if they are not required to.

However, there could be incentives for the expert to provide false information about other parts of the market. An expert may overstate the true amount of fraud present in the market to increase the price differential between wines authenticated by a neutral expert and the auction house and bottles only authenticated by the auction house. Expert authenticators, such as Carfax, often employ scare tactics to induce consumers to have a product authenticated (Carfax, 2016). I plan to test if, under certain conditions, experts will purposefully overstate the number of counterfeit goods publicly offered for sale. These general market claims are separate from the individual bottle authentications, meaning that an expert caught overstating the number of counterfeit goods in the market may suffer no loss in individual reputation (i.e. consumers still believe the expert always provides the correct diagnosis for each individual bottle).

8.5 Conclusion

This manuscript analyzes consumer behavior following a discovery that an expert has authenticated counterfeit goods. The results of this study show that, in the market for fine and rare wine, consumer reliance upon firm reputation can be an effective tool to punish firms for authenticating counterfeit bottles. While not all disclosures result in a reputation loss, I find that three auction houses suffered significant reputation losses following the discovery that each house had authenticated a significant amount of counterfeit wine. This is shown through a series of difference-in-difference and triple difference regressions that control for other factors affecting consumer demand.

This study confirms results from multiple previous studies on credence goods, showing that reputation value increases with product uncertainty. It is hopeful that the theoretical model and estimation strategy used in this study will also prove useful for future studies on firm reputation in the market for credence goods.

References

Abernethy, D. (2017). Sotheby's : "The Don Stott Cellar" Part III Brings $5.2 Million at Sotheby's NY, Achieving 180% of High Estimate [Press release]. Retrieved from http://www.4-traders.com/SOTHEBYS-11842/news/Sothebys-The-Don-Stott-Cellar-Part-III-Brings-5-2-Million-at-Sotheby-s-NY-Achieving-180-of-Hi-24467415/

Ackerlof, G. (1970). The market for lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), 488-500.

Ali, H. H., Lecocq, S., & Visser, M. (2008). The impact of gurus: Parker grades and En primeur wine prices. *Economic Journal, 118*(529), F158-F173. doi:10.1111/j.1468-0297.2008.02147.x

Angrist, J. D., & Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton: Princeton Univ Press.

Ashenfelter, O. (2008). Predicting the quality and prices of Bordeaux wine. *The Economic Journal, 118*(529), F174-F184.

Asimov, E. (2007, March 14, 2007). Sniffing Out Fraud in the Bottle. *New York Times*. Retrieved from http://www.nytimes.com/2007/03/14/dining/14pour.html?_r=0

Baksi, S., & Bose, P. (2007). Credence goods, efficient labelling policies, and regulatory enforcement. *Environmental & Resource Economics, 37*(2), 411-430. doi:10.1007/s10640-006-9032-0

Barzelay, D. E. (2012). The Rise and Fall of a Wine Counterfeiter. Retrieved from https://oldvinenotes.com/2012/06/18/the-rise-and-fall-of-a-wine-counterfeiter/

Bell, K. K. (2014). Vintage Fraud: How to Spot a Counterfeit Wine. Retrieved from http://www.forbes.com/sites/katiebell/2014/06/02/vintage-fraud-how-to-spot-a-counterfeit-wine/#228bfbd7111d

Bonroy, O., Lemarie, S., & Tropeano, J. P. (2013). Credence goods, experts and risk aversion. *Economics Letters, 120*(3), 464-467. doi:10.1016/j.econlet.2013.05.033

Bloomberg Markets. (2014). How Billionaire Bill Koch Fights Fake Wine. *Bloomberg*.

Breidert, C., Hahsler, M., & Reutterer, T. (2006). A review of methods for measuring willingness-to-pay. *Innovative Marketing*.

Canals-Cerda, J. J. (2012). The value of a good reputation online: an application to art auctions. *Journal of Cultural Economics, 36*(1), 67-85. doi:10.1007/s10824-011-9156-0

Carfax. (2016). Americans Driving more than a Quarter-Million Flood Damaged Cars [Press release]. Retrieved from https://www.carfax.com/press/releases/americans-driving-more-than-a-quarter-million-flood-damaged-cars

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review, 99*(4), 1145-1177. doi:10.1257/aer.99.4.1145

Chu, W. J., & Chu, W. (1994). SIGNALING QUALITY BY SELLING THROUGH A REPUTABLE RETAILER - AN EXAMPLE OF RENTING THE REPUTATION OF ANOTHER AGENT. *Marketing Science, 13*(2), 177-189. doi:10.1287/mksc.13.2.177

City National Bank (2015). Fraud in Your Wine Cellar? Three Tips to Spot Counterfeit Wine. Retrieved from http://newsroom.cnb.com/fraud-in-your-wine-cellar

Darby, M. R., & Karni, E. (1973). FREE COMPETITION AND OPTIMAL AMOUNT OF FRAUD. *Journal of Law & Economics, 16*(1), 67-88. doi:10.1086/466756

De Paola, M., & Scoppa, V. (2013). Consumers' Reactions to Negative Information on Product Quality: Evidence from Scanner Data. *Review of Industrial Organization, 42*(3), 235-280. doi:10.1007/s11151-012-9365-5

Dewan, S., & Hsu, V. (2004). Adverse selection in electronic markets: Evidence from online stamp auctions. *Journal of Industrial Economics, 52*(4), 497-516. doi:10.1111/j.0022-1821.2004.00237.x

Downey, M. (2012). Winesearcher Columnist Maureen Downey on Wine Fraud. Retrieved from http://www.chaiconsulting.com/2012/08/09/winesearcher-feature-columnist-maureen-downey-on-wine-fraud/

Dranove, D., & Jin, G. Z. (2010). Quality Disclosure and Certification: Theory and Practice. *Journal of Economic Literature, 48*(4), 935-963. doi:10.1257/jel.48.4.935

Dranove, D., Ramanarayanan, S., & Watanabe, Y. (2012). Delivering Bad News: Market Responses to Negligence. *Journal of Law & Economics, 55*(1), 1-25. doi:10.1086/661227

Dulleck, U., & Kerschbamer, R. (2006). On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature, 44*(1), 5-42. doi:10.1257/002205106776162717

Dulleck, U., & Kerschbamer, R. (2007). Price discrimination via the choice of distribution channels.

Elfenbein, D. W., Fisman, R., & McManus, B. (2015). Market Structure, Reputation, and the Value of Quality Certification. *American Economic Journal: Microeconomics, 7*(4), 83-108. doi:http://www.aeaweb.org/aej-micro/

Emons, W. (1997). Credence goods and fraudulent experts. *Rand Journal of Economics*, *28*(1), 107-119. doi:10.2307/2555942

Freedman, S., Kearney, M., & Lederman, M. (2012). PRODUCT RECALLS, IMPERFECT INFORMATION, AND SPILLOVER EFFECTS: LESSONS FROM THE CONSUMER RESPONSE TO THE 2007 TOY RECALLS. *Review of Economics and Statistics, 94*(2), 499-516. doi:10.1162/REST_a_00162

Friberg, R., & Gronqvist, E. (2012). Do Expert Reviews Affect the Demand for Wine? *American Economic Journal-Applied Economics, 4*(1), 193-211. doi:10.1257/app.4.1.193

Garber, S., & Adams, J. (1998). Product and stock market responses to automotive product liability verdicts. *Brookings Papers on Economic Activity*, 1-53.

Gray, W. B. (2014). The 16 Most Counterfeited Wines. Retrieved from http://www.wine-searcher.com/m/2014/08/the-16-most-counterfeited-wines

Gray, W. B. (2015). Questions Raised by Hong Kong Wine Auction. Retrieved from http://www.wine-searcher.com/m/2015/12/questions-raised-by-hong-kong-wine-auction

Greene, W. H. (2012). *Econometric Analysis* (Seventh Edition ed.). Boston: Pearson Education Limited.

Gruber, J. (1994). THE INCIDENCE OF MANDATED MATERNITY BENEFITS. *American Economic Review, 84*(3), 622-641.

Haley, M. R., & Van Scyoc, L. (2010). Adverse selection, seller reputation and buyer insurance in online auctions for 1960s-era collectible baseball cards. *Applied Economics Letters, 17*(14), 1341-1345. doi:10.1080/13504850902984303

Haughney, C. (2015). Selling the Cellar. Retrieved from http://www.foodrepublic.com/2015/11/05/auction-houses-complex-relationship-with-wine/

Hellman, P. (2007). Countering Counterfeit Wines. *New York Sun*. Retrieved from http://www.nysun.com/food-drink/countering-counterfeit-wines/65159/

Hellman, P. (2008). Domaine Ponsot Proprieter Halts Sale of Fake Bottles. *Wine Spectator*. Retrieved from http://www.winespectator.com/webfeature/show/id/Domaine-Ponsot-Proprietor-Halts-Sale-of-Fake-Bottles_4131

Hellman, P. (2012). Justice Department Rejects Kurniawan's Illegal Search Accusations. Retrieved from http://www.winespectator.com/webfeature/show/id/47580

Hellman, P. (2015). Rudy Kurniawan's Last Auction. *Wine Spectator*.

Hellman, P. (2017). *In Vino Duplicitas: the rise and fall of a wine forger extraordinaire*. New York: Experiment.

Hellman, P., & Frank, M. (2009, December 15, 2009). The Crusade Against Counterfeits. *Wine Spectator,* 46-66.

United States of America v. Rudy Kurniawan (S.D.N.Y. 2012).

Hilger, J., Rafert, G., & Villas-Boas, S. (2011). EXPERT OPINION AND THE DEMAND FOR EXPERIENCE GOODS: AN EXPERIMENTAL APPROACH IN THE RETAIL WINE MARKET. *Review of Economics and Statistics, 93*(4), 1289-1296. doi:10.1162/REST_a_00117

Hirsch, J. (2012). The Wine Detective: Maureen Downey Reveals Forged Bottles. Retrieved from http://archives.sfweekly.com/sanfrancisco/the-wine-detective-maureen-downey-reveals-forged-bottles/Content?oid=2186203

Hubbard, T. N. (1998). An empirical examination of moral hazard in the vehicle inspection market. *Rand Journal of Economics, 29*(2), 406-426. doi:10.2307/2555895

Jin, G. Z., & Kato, A. (2006). Price, quality, and reputation: evidence from an online field experiment. *Rand Journal of Economics, 37*(4), 983-1004.

Jones, G. V., & Storchmann, K. H. (2001). Wine market prices and investment under uncertainty: an econometric model for Bordeaux Crus Classes. *Agricultural Economics, 26*(2), 115-133. doi:10.1016/s0169-5150(00)00102-x

Jullien, A. (1816). *Topographie de tous les vignobles connus*. Paris.

Kapon, J. (2006). The Cellar I and II Auction Catalog.

Kapon, J. (2007). Results & Tastings Notes from Evening Auction. *Vintage Tastings by John Kapon*. Retrieved from http://www.vintagetastings.com/all/2007-5-11_00110.html

Kapon, J. (2008). A Selection from the Cellar of Robert A. Rosania.

Krebiehl, A. (2017). Fighting Fraud: Knowledge is Power. *The World of Fine Wine*.

Lecat, B., & Chapuis, J. B. C. (2017). Fraud and counterfeit wines in France: an overview and perspectives. *British Food Journal, 119*(1), 84-104. doi:10.1108/bfj-09-2016-0398

Leve, J. Bordeaux Vintage Chart 1959 to Today, Vintage Rankings Characteristics. Retrieved from http://www.thewinecellarinsider.com/wine-topics/bordeaux-wine-buying-guide-tasting-notes-ratings/bordeaux-wine-vintage-chart/

Lewis, G. (2011). Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors. *American Economic Review, 101*(4), 1535-1546. doi:http://www.aeaweb.org/aer/

Lichfield, J. (2016, May 22, 2016). Auction of world's most expensive wine thrown into chaos after claims they're fake. *Independent*. Retrieved from http://www.independent.co.uk/news/world/europe/wine-fraud-most-expensive-bottle-auction-burgundy-romanee-conti-domain-authenticity-geneva-a7042336.html

McCoy, E. (2007, March 8, 2007). Wine industry fights back against counterfeiters. *New York Times*. Retrieved from http://www.nytimes.com/2007/03/08/business/worldbusiness/08iht-wine.4849816.html?_r=0

McInerney, J. (2008). Billionaire Winos. *Men's Vogue, September 2008*.

Melnik, M. I., & Alm, J. (2005). Seller reputation information signals, and prices for heterogeneous coins on eBay. *Southern Economic Journal, 72*(2), 305-328.

Meltzer, P. D. (2016, May 24, 2016). Sotheby's Scores Big with Bill Koch's $21.9 Million Wine Auction. *Wine Spectator*.

Millar, R. (2017). Russian Wine Fraudster Jailed in France. Retrieved from https://www.thedrinksbusiness.com/2017/05/russian-wine-fraudster-jailed-in-france/

Parker, R. (2015). The Wine Advocate Vintage Guide. Retrieved from http://www.erobertparker.com/newsearch/vintagechart1.aspx/VintageChart.aspx

Pfanner, E. (2012, June 19, 2012). Fraud Charges Threaten Burgundy's Vaunted Reputation. *New York Times*, p. 83. Retrieved from http://www.nytimes.com/2012/06/19/business/global/fraud-charges-threaten-burgundys-vaunted-reputation.html?_r=0

Prince, D. W., & Rubin, P. H. (2002). The effects of product liability litigation on the value of firms. *American Law and Economics Review, 4*(1), 44-87.

Rao, A., & Wang, E. (2017). Demand for "Healthy" Products: False Claims and FTC Regulation. *Journal of Marketing Research, 54*(6), 968-989. doi:10.1509/jmr.15.0398

Richard, D. (2013, November 1, 2013). Vins de Bordeaux: le trafic de fausses etiquettes explose sur internet. *Sud Oest*. Retrieved from http://www.sudouest.fr/2013/11/01/etiquettes-trompeuses-1216931-2780.php

Roberts, J. W. (2011). Can Warranties Substitute for Reputations? *American Economic Journal-Microeconomics, 3*(3), 69-85. doi:10.1257/mic.3.3.69

Robinson, J. (1999). *Oxford companion to wine*: Oxford University Press.

Robinson, J. (2007). That crazy, single-owner Acker sale. Retrieved from http://www.jancisrobinson.com/articles/that-crazy-single-owner-acker-sale?layout=pdf

Robinson, J. (2014). A nose for crime: counterfeit Wines. Retrieved from http://www.ft.com/cms/s/2/6e8e66d8-8829-11e3-8afa-00144feab7de.html

Schlenker, W., & Villas-Boas, S. B. (2009). Consumer and Market Responses to Mad Cow Disease. *American Journal of Agricultural Economics, 91*(4), 1140-1152. doi:10.1111/j.1467-8276.2009.01315.x

Shapiro, C. (1983). PREMIUMS FOR HIGH-QUALITY PRODUCTS AS RETURNS TO REPUTATIONS. *Quarterly Journal of Economics, 98*(4), 659-679. doi:10.2307/1881782

Society, A. C. (2010). Detecting fake wine vintages: It's an (atomic) blast. Retrieved from https://www.acs.org/content/acs/en/pressroom/newsreleases/2010/march/detecting-fake-wine-vintages-its-an-atomic-blast.html

Squires, M. (2008). WS article on fake Ponsot lot.. Retrieved from http://www.erobertparker.com/bboard/

Stahl, K., & Strausz, R. (2017). Certification and Market Transparency. *Review of Economic Studies, 84*(4), 1842-1868. doi:10.1093/restud/rdw064

Steinberger, M. (2012, July 2012). A Vintage Crime. *Vanity Fair*.

Taylor, R. (2013). Who are They Fooling? (A Lot of Folks). *Wine Spectator*.

Toth, A. (2014). REPUTATION EFFECTS IN THE MARKET OF CERTIFIERS: EVIDENCE FROM THE AUDIT INDUSTRY. *Economic Inquiry, 52*(2), 505-517. doi:10.1111/ecin.12063

Wallace, B. (2009). *The Billionaire's Vinegar: The Mystery of the World's Most Expensive Bottle of Wine*: Broadway Books.

Wallace, B. (2012, May 13, 2012). Chateau Sucker. *New York*.

Wang, N. (2017). Acker Faces Fake WIne Accusation over DRC Lots.   Retrieved from https://www.thedrinksbusiness.com/2017/04/acker-faces-fake-wine-accusation-over-drc-lots/

Wise, J. (2008). The New Business of Old Wine.   Retrieved from http://digital.modernluxury.com/display_article.php?id=88469

Wine Berserkers. (2012). RUDY KURNIAWAN & GLOBAL WINE AUCTION FRAUD THREAD Retrieved from http://www.wineberserkers.com/forum/viewtopic.php?f=1&t=61172

Wine Institute. (2015). Number of California Wineries.   Retrieved from http://www.wineinstitute.org/resources/statistics/article124

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*: MIT press.

Table 1.1: Bottle Characteristics- Producer Sales Volume per Year

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Auction House											
Acker Merrall Condit	1,355	1,397	1,798	2,158	2,001	2,893	2,966	2,512	1,669	1,457	1,484
Bonhams & Butterfields	252	429	516	602	438	492	778	493	452	415	352
Christies	2,499	2,494	1,928	2,033	1,800	1,962	2,409	2,249	1,867	1,499	1,054
Hart Davis Hart	582	653	1,057	1,408	931	1,616	1,863	1,028	1,247	1,123	1,020
Heritage Auctions							520	302	228	325	221
K & L Wines							842	851	857	939	899
Morrell & Company	335	327	361	260	265	264	404	258	208	146	42
Sothebys	1,638	1,684	2,085	1,880	1,480	2,587	2,870	2,281	1,745	1,914	991
Spectrum Wine Auctions					178	1,200	1,418	759	396	482	393
Zachys	1,566	1,708	1,671	2,040	1,988	1,963	2,815	2,297	1,379	1,350	1,251
Producer											
Lafite	900	884	1,084	1,404	1,085	1,954	2,983	1,813	1,163	1,047	784
Latour	794	780	836	952	831	1,273	1,522	1,225	829	827	634
Mouton	956	1011	1,166	1,220	1,127	1,725	2,134	1,667	1,123	1,121	773
Margaux	668	741	832	961	867	1,220	1,612	1,184	944	862	630
Haut Brion	565	553	611	715	631	856	1,116	854	738	661	408
DRC La Tache	216	318	342	355	365	400	437	465	313	348	258
Ponsot Clos de la Roche	81	41	43	45	58	41	35	60	47	62	31
DRC Romanee Conti	125	190	204	198	172	234	304	232	214	169	133
DRC Romanee St. Vivant	71	98	103	109	113	109	150	160	137	143	104
DRC Richebourg	129	130	144	134	134	165	200	188	157	137	108
Petrus	361	433	427	386	350	506	576	433	379	311	287
Cheval Blanc	419	470	454	438	447	583	716	573	442	426	307
Le Pin	18	44	41	48	25	36	40	39	35	35	36
Dujac Clos de la Roche	19	30	30	20	24	33	20	27	36	32	32
Henri Jayer Echezeaux	15	21	15	17	5	15	21	23	18	11	16
DRC Echezeaux	44	76	76	79	82	86	120	107	83	114	74
DRC Grands-Echezeaux	62	73	91	103	107	106	128	136	101	118	72
Ducru-Beaucaillou	261	208	262	267	183	286	417	313	307	263	231
Calon Segur	90	78	132	148	90	136	143	131	117	142	146
Leoville Barton	192	199	165	168	133	198	299	204	234	192	187
Leoville Poyferre	110	102	93	125	87	127	159	114	114	122	109
Lynch Bages	336	444	408	454	350	502	821	550	507	472	422
Talbot	102	117	84	108	73	86	122	111	91	102	88
La Lagune	64	35	38	35	28	41	52	31	42	45	47
Beychevelle	85	137	55	94	83	96	120	109	93	126	88
Les Forts de Latour	27	24	27	61	47	62	117	65	85	91	60
Cos d'Estournel	252	237	286	308	256	343	453	396	309	344	315
Mommessin Clos de Tart	26	25	40	40	29	53	51	58	71	60	38
De Vogue Bonnes Mares	36	40	34	53	49	43	48	56	43	42	44
Louis Latour Corton-Charlemagne	21	38	39	27	19	18	19	24	13	35	27
Bonneau Corton-Charlemagne	19	21	20	16	25	22	27	35	17	20	19
Lafleur	62	78	83	89	78	92	89	77	82	48	59
Latour a Pomerol	18	18	14	18	21	20	22	24	17	11	16
Trotanoy	31	33	40	42	27	43	39	43	36	31	21
La Mission Haut Brion	245	246	311	257	277	348	429	407	281	224	224
Leoville Las Cases	335	270	337	428	336	563	640	503	345	429	361
Palmer	189	161	142	152	135	178	226	200	153	127	145
Montrose	184	178	195	196	229	249	345	242	198	223	222
Henri Jayer Richebourg	6	14	5	4	8	16	17	18	9	4	7
Henri Jayer Cros Parantoux	42	37	52	34	25	60	58	63	62	26	67
De Vogue Musigny	51	59	55	73	70	53	58	70	63	47	77
Continent											
Asia				794	2,058	3,958	6,335	3,790	2,244	2,039	1,431
Europe	2,243	2,688	2,430	2,484	1,734	1,819	2,266	2,004	1,697	1,418	877
Internet				445	622	1,144	1,856	2,362	2,263	2,464	2,053
North America	5,984	6,004	6,986	6,658	4,667	6,056	6,428	4,874	3,844	3,729	3,346
Mean Price											
Bottle Unlikely To be Faked	$181	$214	$273	$250	$232	$293	$323	$282	$284	$278	$274
Bottle Likely to be Counterfeited	$1,388	$2,018	$2,550	$2,373	$2,241	$3,074	$2,981	$2,511	$2,656	$2,613	$2,769
Total	8,227	8,692	9,416	10,381	9,081	12,977	16,885	13,030	10,048	9,650	7,707

Table 1.2: Auction House Sales by Year (in Millions of U.S. Dollars)

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Acker	16.6	17.0	20.9	60.3	59.9	59.8	44.2	98.5	111	83.3	63.8	61.8	69.9
Zachys	15.3	26.0	33.8	34.7	52.4	47.7	50.7	56.5	79.0	70.2	53.0	45.2	55.5
Sothebys	24.0	21.0	29.1	37.4	49.3	44.6	41.8	88.3	85.5	64.4	57.8	65.3	60.4
Christies	31.0	36.0	42.0	58.6	58.1	55.7	42.4	71.5	85.3	83.8	68.0	53.0	57.2
Hart Davis Hart			9.5	13.8	26.9	32.3	24.0	39.2	37.4	26.3	36.1	42.8	41.5
Bonhams				7.5	7.3	9.83	4.87	7.19	16.5	16.1	18.2	13.3	17.9
Spectrum							3.5	15.7	24.2	11.9	10.1	5.6	7.0
Morrell & Co					6.1	4.1	1.87	5.8	3.9	2.9	1.7	2.4	
Heritage									11.9	7.1	7.1	12.3	7.2
World Total	96	109	166	241	301	276	233	408	478	389	337	352	346

Source: Wine Spectator

Table 1.3: 2008 Difference-in-Difference Price Estimates

	(1) Base DD	(2) North America	(3) Case and Large Format	(4) Very Expensive
Acker	Base	Base	Base	Base
	-	-	-	-
Bonhams & Butterfields	-0.133***	-0.153***	-0.121***	-0.149***
	(0.0185)	(0.0201)	(0.0232)	(0.0311)
Hart Davis Hart	-0.0494***	-0.0510***	-0.0552***	-0.0837***
	(0.0130)	(0.0126)	(0.0202)	(0.0197)
Morrell & Company	-0.174***	-0.172***	-0.164***	-0.207***
	(0.0209)	(0.0206)	(0.0267)	(0.0517)
Sothebys	0.00378	0.00204	-0.0229	0.00566
	(0.0168)	(0.0172)	(0.0223)	(0.0384)
Zachys	-0.0663***	-0.0700***	-0.0913***	-0.0993***
	(0.0109)	(0.0116)	(0.0171)	(0.0171)
Time and House (DD)	**0.0810*** **	**0.0849*** **	**0.0504** **	**0.0770*** **
	(0.0135)	**(0.0154)**	**(0.0195)**	**(0.0263)**
Asia	Base	Base	Base	Base
	-	-	-	-
Europe	-0.138***		-0.101***	-0.183***
	(0.0223)		(0.0257)	(0.0481)
Internet	-0.227***		-0.336***	-0.266***
	(0.0211)		(0.0826)	(0.0324)
North America	-0.181***		-0.187***	-0.172***
	(0.0146)		(0.0172)	(0.0163)
Lot Size	0.00610***	0.00565***		0.0149***
	(0.00155)	(0.00141)		(0.00250)
Case Dummy	0.0371***	0.0472***		-0.0176
	(0.0131)	(0.0124)		(0.0154)
Observations	8,785	7,017	3,669	3,001
R-squared	0.816	0.830	0.921	0.758
Number of Unique Wines	178	177	133	100

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

Table 1.4: Expansion of Sales into Asia

Percentage of Lots Sold in Asia								
House	2008	2009	2010	2011	2012	2013	2014	2015
Acker Merrall & Condit	18.54%	30.68%	49.12%	49.22%	37.06%	43.08%	40.77%	40.97%
Bonhams & Butterfields	16.28%	21.69%	13.62%	17.74%	16.23%	15.04%	5.30%	9.09%
Christies	3.59%	8.94%	23.29%	37.73%	29.44%	22.60%	24.28%	19.92%
Sothebys		27.50%	50.06%	45.09%	37.62%	35.30%	38.40%	26.54%
Spectrum Wine Auctions				73.70%	56.52%			
Zachys	10.93%	39.29%	36.58%	52.90%	36.13%	30.38%	24.00%	25.42%
Percentage of Total Bottles Sold in Asia								
House	2008	2009	2010	2011	2012	2013	2014	2015
Acker Merrall & Condit	25.35%	40.90%	61.80%	65.21%	55.36%	58.39%	55.32%	50.82%
Bonhams & Butterfields	24.28%	25.91%	14.18%	20.18%	16.72%	12.17%	6.28%	7.31%
Christies	3.38%	14.58%	29.43%	41.96%	33.62%	29.44%	30.47%	22.76%
Sothebys		26.45%	49.32%	47.86%	37.41%	39.14%	42.67%	29.20%
Spectrum Wine Auctions				86.55%	72.33%			
Zachys	14.07%	43.22%	42.46%	59.29%	42.61%	37.90%	30.89%	31.83%
Percentage of Expensive ($2,000+/Bottle) Lots Sold in Asia								
House	2008	2009	2010	2011	2012	2013	2014	2015
Acker Merrall & Condit	38.56%	52.43%	86.18%	79.84%	65.82%	71.68%	67.84%	65.90%
Bonhams & Butterfields	24.43%	22.94%	44.67%	24.41%	18.60%	15.38%	4.31%	9.76%
Christies	12.01%	26.03%	31.78%	53.40%	34.91%	50.66%	44.77%	26.00%
Sothebys		47.09%	54.96%	60.05%	49.40%	44.03%	46.92%	42.89%
Spectrum Wine Auctions				97.39%	58.06%			
Zachys	16.30%	48.60%	56.08%	76.38%	53.29%	55.87%	45.00%	55.27%

Note: HDH, Heritage, K & L, and Morrell & Company have not auctioned wine in Asia.
The values reported represent percentages based upon sales included in the data set.

Table 1.5: 2008 Triple Difference Price Estimates

	(1) Base DDD	(2) North America	(3) Burgundy	(4) Old Wines	(5) Semiannual Continent	(6) New Wines
Acker	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-
Bonhams & Butterfields	-0.111***	-0.122***	-0.196***	-0.186***	-0.0953***	-0.0704**
	(0.0189)	(0.0197)	(0.0403)	(0.0390)	(0.0188)	(0.0310)
Hart Davis Hart	-0.0238	-0.0129	-0.151***	-0.0694*	-0.0275	-0.00585
	(0.0180)	(0.0182)	(0.0417)	(0.0394)	(0.0178)	(0.0274)
Morrell & Company	-0.138***	-0.124***	-0.203***	-0.198***	-0.141***	-0.0897**
	(0.0178)	(0.0177)	(0.0492)	(0.0355)	(0.0174)	(0.0349)
Sothebys	0.0321	0.0444**	-0.0684	-0.00454	0.0175	0.0250
	(0.0201)	(0.0207)	(0.0485)	(0.0377)	(0.0205)	(0.0384)
Zachys	-0.0491***	-0.0386**	-0.178***	-0.0704*	-0.0685***	-0.0632***
	(0.0163)	(0.0165)	(0.0367)	(0.0365)	(0.0162)	(0.0237)
Time and House	0.0580***	0.0431**	0.0279	0.0978***	0.0444***	0.0568***
	(0.0145)	(0.0175)	(0.0474)	(0.0340)	(0.0152)	(0.0199)
Time and Fake	-0.00431	-0.00494	-0.0262	0.00125	-0.0101	0.00330
	(0.00895)	(0.0101)	(0.0203)	(0.0168)	(0.00931)	(0.0129)
Fake and House	0.00887	0.0300	-0.151***	0.0201	0.00477	-0.0323
	(0.0193)	(0.0191)	(0.0375)	(0.0416)	(0.0191)	(0.0305)
Time, Fake, and House(DDD)	**0.0127**	**0.0436***	**0.117****	**-0.0321**	**0.0156**	**0.0224**
	(0.0199)	**(0.0234)**	**(0.0529)**	**(0.0407)**	**(0.0213)**	**(0.0313)**
Asia	Base		Base	Base		Base
	-		-	-		-
Europe	-0.221***		-0.0929*	-0.265***		-0.212***
	(0.0206)		(0.0560)	(0.0365)		(0.0381)
Internet	-0.240***		-0.234***	-0.299***		-0.242***
	(0.0151)		(0.0349)	(0.0289)		(0.0227)
North America	-0.207***		-0.171***	-0.217***		-0.203***
	(0.0118)		(0.0201)	(0.0254)		(0.0157)
Lot Size	0.00191	0.00148	0.00573**	0.00206	0.00147	0.00500**
	(0.00137)	(0.00130)	(0.00235)	(0.00234)	(0.00135)	(0.00210)
Case Dummy	0.0434***	0.0569***	0.0136	0.0848***	0.0562***	-0.0190
	(0.00958)	(0.00936)	(0.0214)	(0.0159)	(0.00960)	(0.0160)
Observations	14,189	10,837	2,390	5,555	14,189	4,679
R-squared	0.787	0.808	0.501	0.719	0.774	0.840
Number of Unique Wines	286	282	106	141	286	86

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

Table 1.6: 2012 Difference-in-Difference Price Estimates

	(1) Base DD	(2) North America	(3) Case and Large Format	(4) Very Expensive
Acker	Base	Base	Base	Base
	-	-	-	-
Bonhams & Butterfields	-0.140***	-0.200***	-0.0636***	-0.101**
	(0.0214)	(0.0180)	(0.0220)	(0.0449)
Christies	0.0237*	-0.0484***	0.0459**	0.0542**
	(0.0127)	(0.0110)	(0.0190)	(0.0261)
Hart Davis Hart	-0.0721***	-0.103***	-0.0783***	-0.0411**
	(0.0116)	(0.0117)	(0.0136)	(0.0183)
Heritage Auctions	-0.124***	-0.148***	-0.115***	-0.0867***
	(0.0125)	(0.0119)	(0.0224)	(0.0222)
K & L Wines	0.0163		-0.0700*	0.0736
	(0.0192)		(0.0362)	(0.0580)
Morrell & Company	-0.120***	-0.112***	-0.0475*	-0.132***
	(0.0219)	(0.0254)	(0.0243)	(0.0394)
Sothebys	-0.0254**	-0.0768***	-0.00605	-0.0147
	(0.0122)	(0.0162)	(0.0136)	(0.0183)
Spectrum Wine Auctions	-0.144***		-0.0922***	-0.143***
	(0.0175)		(0.0245)	(0.0295)
Zachys	-0.0726***	-0.0933***	-0.0744***	-0.0746***
	(0.0113)	(0.0107)	(0.0147)	(0.0192)
Time and House (DD)	**-0.0441*** **	**-0.0710*** **	**-0.0682*** **	**-0.0145**
	(0.0118)	**(0.0139)**	**(0.0151)**	**(0.0209)**
Asia	Base		Base	Base
	-		-	-
Europe	-0.116***		-0.0928***	-0.163***
	(0.0147)		(0.0148)	(0.0302)
Internet	-0.183***		-0.130***	-0.271***
	(0.0193)		(0.0280)	(0.0363)
North America	-0.0949***		-0.0982***	-0.117***
	(0.00825)		(0.0122)	(0.0129)
Lot Size	-0.000230	0.000456		0.00740**
	(0.00122)	(0.00120)		(0.00361)
Case Dummy	0.0670***	0.0583***		0.0288
	(0.0119)	(0.00932)		(0.0257)
Observations	15,471	6,090	6,657	4,908
R-squared	0.775	0.778	0.892	0.669
Number of Unique Wines	178	177	143	100

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

Table 1.7: 2012 Triple Difference Price Estimates

	(1) Base DDD	(2) North America	(3) Old Wines	(4) Case and Large Format	(5) New Wines
Acker	Base	Base	Base	Base	Base
	-	-	-	-	-
Bonhams & Butterfields	-0.116***	-0.174***	-0.183***	-0.0539***	-0.0572***
	(0.0147)	(0.0146)	(0.0289)	(0.0175)	(0.0201)
Christies	0.0174*	-0.0317***	0.0193	0.0322**	0.0377***
	(0.00950)	(0.00957)	(0.0203)	(0.0136)	(0.0115)
Hart Davis Hart	-0.0229**	-0.0568***	-0.0541**	-0.00379	0.00948
	(0.0113)	(0.0147)	(0.0231)	(0.0141)	(0.0154)
Heritage Auctions	-0.104***	-0.129***	-0.126***	-0.0866***	-0.119***
	(0.0122)	(0.0157)	(0.0241)	(0.0193)	(0.0154)
K & L Wines	0.0364***		0.0153	-0.0152	0.0367*
	(0.0119)		(0.0225)	(0.0280)	(0.0189)
Morrell & Company	-0.113***	-0.120***	-0.197***	-0.0645**	-0.0428**
	(0.0173)	(0.0230)	(0.0301)	(0.0252)	(0.0179)
Sothebys	0.0103	-0.0500***	-0.0288	0.0365**	0.0265
	(0.0108)	(0.0151)	(0.0228)	(0.0143)	(0.0160)
Spectrum Wine Auctions	-0.127***		-0.205***	-0.0991***	-0.0593***
	(0.0117)		(0.0194)	(0.0189)	(0.0162)
Zachys	-0.0426***	-0.0638***	-0.0830***	-0.0289**	-0.0289*
	(0.00960)	(0.0138)	(0.0176)	(0.0127)	(0.0148)
Time and House	-0.0167	-0.0662***	0.0192	-0.0145	-0.0249
	(0.0114)	(0.0163)	(0.0254)	(0.0147)	(0.0160)
Time and Fake	-0.0334***	-0.0330***	0.00174	-0.0392***	-0.0438**
	(0.0120)	(0.0121)	(0.0188)	(0.0127)	(0.0197)
Fake and House	0.0338***	0.0265*	0.0284	0.0573***	0.0188
	(0.0112)	(0.0152)	(0.0223)	(0.0161)	(0.0186)
Time, Fake, and House (DDD)	**-0.0303***	**-0.00112**	**-0.0843****	**-0.0638*****	**-0.0123**
	(0.0179)	**(0.0211)**	**(0.0383)**	**(0.0237)**	**(0.0264)**
Asia	Base		Base	Base	Base
	-		-	-	-
Europe	-0.110***		-0.185***	-0.0899***	-0.0711***
	(0.0113)		(0.0272)	(0.0116)	(0.00959)
Internet	-0.164***		-0.239***	-0.126***	-0.108***
	(0.0114)		(0.0229)	(0.0201)	(0.0117)
North America	-0.100***		-0.142***	-0.0969***	-0.0843***
	(0.00710)		(0.0151)	(0.00908)	(0.00951)
Lot Size	-0.00256***	-0.00154	-0.00538***		-0.00105
	(0.000732)	(0.000946)	(0.00162)		(0.000871)
Case Dummy	0.0690***	0.0641***	0.119***		0.0377***
	(0.00684)	(0.00668)	(0.0149)		(0.00664)
Observations	25,078	9,686	9,031	11,697	8,688
R-squared	0.740	0.729	0.634	0.859	0.860
Number of Unique Wines	286	284	141	249	86

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

Table 1.8: 2008 Difference-in-Difference Estimates for Quantity of Bottles Sold

	(1) OLS DD	(2) Negative Binomial	(3) Old Wines	(4) Case and Large Format	(5) Very Expensive
Acker	Base	Base	Base	Base	Base
	-	-	-	-	-
Bonhams & Butterfields	-1.080***	-1.982***	-1.688***	-2.136***	-2.483***
	(0.0732)	(0.138)	(0.199)	(0.209)	(0.230)
Hart Davis Hart	-0.322***	-0.427***	-0.547***	0.130	-0.667***
	(0.0665)	(0.0957)	(0.146)	(0.167)	(0.165)
Morell & Company	-1.227***	-2.395***	-2.069***	-2.053***	-3.135***
	(0.0749)	(0.183)	(0.221)	(0.271)	(0.374)
Sothebys	-0.0548	-0.00868	0.0183	0.344*	-0.189
	(0.0798)	(0.104)	(0.139)	(0.198)	(0.176)
Zachys	0.174**	0.0925	0.238*	0.167	-0.0834
	(0.0672)	(0.0930)	(0.130)	(0.191)	(0.148)
Time and House (DD)	**0.206*** **	**0.136**	**0.0181**	**-0.0641**	**0.103**
	(0.0682)	**(0.0998)**	**(0.148)**	**(0.210)**	**(0.182)**
One Year-Six Months Prior	Base	Base	Base	Base	Base
	-	-	-	-	-
Six Months Prior-Disclosure	0.390***	0.563***	0.496***	0.549***	0.598***
	(0.0337)	(0.0604)	(0.0834)	(0.113)	(0.0915)
Disclosure-Six Months Post	-0.0137	-0.0296	-0.252**	0.0807	-0.293**
	(0.0372)	(0.0763)	(0.110)	(0.131)	(0.145)
Six Months -One Year Post	0.0859**	0.114	-0.00447	0.249*	0.104
	(0.0406)	(0.0833)	(0.124)	(0.134)	(0.151)
Observations	3,888	3,888	2,304	3,216	1,608
R-squared	0.302				
Number of Unique Wines	162	162	96	134	67

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

Table 1.9: 2008 Triple Difference Estimates for Quantity of Bottles Sold

	(1) OLS DDD	(2) Negative Binomial	(3) North America
Acker	Base	Base	Base
	-	-	-
Bonhams & Butterfields	-0.747***	-1.241***	-1.779***
	(0.0998)	(0.133)	(0.126)
Hart Davis Hart	-0.0887	-0.152	-0.0803
	(0.0975)	(0.120)	(0.111)
Morell & Company	-0.963***	-1.761***	-1.644***
	(0.0974)	(0.150)	(0.138)
Sothebys	0.312***	0.394***	-0.238**
	(0.111)	(0.123)	(0.106)
Zachys	0.435***	0.349***	0.322***
	(0.0972)	(0.111)	(0.104)
Time and House	0.283**	0.0710	-0.122
	(0.111)	(0.138)	(0.153)
Time and Fake	-0.0604**	-0.0890*	-0.0765
	(0.0299)	(0.0468)	(0.0542)
Fake and House	0.292**	0.339**	0.163
	(0.113)	(0.132)	(0.120)
Time, Fake, and House (DDD)	**-0.0766**	**0.0205**	**-0.150**
	(0.131)	**(0.169)**	**(0.189)**
One Year-Six Months Prior	Base	Base	Base
	-	-	-
Six Months Prior-Disclosure	0.308***	0.394***	0.391***
	(0.0319)	(0.0509)	(0.0551)
Disclosure-Six Months Post	0.0391	0.0430	-0.0896
	(0.0383)	(0.0644)	(0.0711)
Six Months -One Year Post	0.0726*	0.0845	-0.0888
	(0.0391)	(0.0648)	(0.0740)
Observations	6,696	6,696	6,648
R-squared	0.231		
Number of Unique Wines	279	279	277

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

Table 1.10: 2012 Difference-In-Difference Estimates for Quantity of Bottles Sold

	(1) **OLS**	(2) **Negative Binomial**	(3) **Old Wines**	(4) **Case and Large Format**	(5) **Very Expensive**	(6) **Time Lag**
Acker	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-
Bonhams & Butterfields	-1.227***	-1.973***	-1.744***	-2.112***	-2.073***	-1.797***
	(0.0736)	(0.154)	(0.179)	(0.256)	(0.277)	(0.131)
Christies	-0.139**	-0.0807	0.102	-0.122	-0.172	0.321***
	(0.0567)	(0.0774)	(0.109)	(0.129)	(0.114)	(0.0761)
Hart Davis Hart	-0.805***	-1.096***	-1.613***	-1.150***	-1.946***	-0.911***
	(0.0658)	(0.109)	(0.187)	(0.159)	(0.196)	(0.104)
Heritage Auctions	-1.269***	-2.128***	-2.063***	-2.293***	-2.521***	-2.125***
	(0.0704)	(0.128)	(0.203)	(0.206)	(0.257)	(0.121)
K & L Wines	-1.414***	-2.881***	-2.727***	-4.525***	-3.682***	-2.800***
	(0.0765)	(0.127)	(0.186)	(0.296)	(0.222)	(0.121)
Morell & Company	-1.553***	-3.179***	-2.953***	-3.278***	-3.794***	-3.187***
	(0.0834)	(0.141)	(0.184)	(0.197)	(0.244)	(0.141)
Sothebys	-0.253***	-0.197**	-0.209	-4.15e-05	-0.736***	-0.0637
	(0.0648)	(0.0942)	(0.139)	(0.148)	(0.144)	(0.0867)
Spectrum Wine Auctions	-0.840***	-1.116***	-0.862***	-1.584***	-1.286***	-1.197***
	(0.0667)	(0.0906)	(0.122)	(0.163)	(0.167)	(0.101)
Zachys	0.0385	-0.0141	0.0374	-0.128	-0.642***	-0.0482
	(0.0663)	(0.0868)	(0.113)	(0.143)	(0.141)	(0.0833)
Time and House (DD)	**-0.0457**	**-0.0671**	**-0.0927**	**0.0932**	**-0.387*** **	**-0.230** **
	(0.0442)	**(0.0843)**	**(0.129)**	**(0.145)**	**(0.143)**	**(0.0940)**
One Year-Six Months Prior	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-
Six Months Prior- Disclosure Date	0.0739**	0.110*	0.116	-0.0130	0.311***	0.124**
	(0.0285)	(0.0604)	(0.0875)	(0.103)	(0.106)	(0.0603)
Disclosure Date- Six Months Post	-0.160***	-0.346***	-0.421***	-0.586***	-0.219	
	(0.0303)	(0.0727)	(0.122)	(0.122)	(0.144)	
Six Months – One Year Post	0.0130	-0.0116	0.0456	-0.171	0.142	0.0293
	(0.0335)	(0.0750)	(0.112)	(0.128)	(0.132)	(0.0783)
One Year Post- 18 Months Post						-0.490***
						(0.0740)
Observations	6,480	6,480	3,840	5,520	2,680	6,480
R-squared	0.309					
Number of Unique Wines	162	162	96	138	67	162

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

Table 1.11: 2012 Triple Difference Estimates for Quantity of Bottles Sold

	(1) **OLS**	(2) **Negative Binomial**	(3) **North America**	(4) **Time Lag**
Acker	Base	Base	Base	Base
	-	-	-	-
Bonhams & Butterfields	-1.244***	-1.573***	-1.286***	-1.433***
	(0.0728)	(0.119)	(0.130)	(0.109)
Christies	-0.0514	-0.00692	-0.334***	0.213***
	(0.0519)	(0.0619)	(0.0842)	(0.0659)
Hart Davis Hart	-0.746***	-0.782***	0.615***	-0.766***
	(0.0639)	(0.0855)	(0.103)	(0.0861)
Heritage Auctions	-1.465***	-2.133***	-0.679***	-2.071***
	(0.0662)	(0.101)	(0.124)	(0.0983)
K & L Wines	-1.422***	-2.399***	-14.18***	-2.289***
	(0.0653)	(0.102)	(1.289)	(0.100)
Morell & Company	-1.680***	-2.807***	-1.918***	-2.777***
	(0.0691)	(0.117)	(0.151)	(0.119)
Sothebys	-0.280***	-0.198**	-0.124	-0.137*
	(0.0657)	(0.0846)	(0.113)	(0.0832)
Spectrum Wine Auctions	-0.839***	-1.024***	-14.83***	-1.125***
	(0.0540)	(0.0731)	(1.706)	(0.0841)
Zachys	0.00699	-0.0338	0.740***	-0.0958
	(0.0631)	(0.0782)	(0.104)	(0.0806)
Time and House	-0.0521	-0.0885	0.0680	-0.301**
	(0.0684)	(0.111)	(0.158)	(0.130)
Time and Fake	0.0732***	0.0505	0.101*	0.0171
	(0.0247)	(0.0473)	(0.0586)	(0.0432)
Fake and House	-0.0794	-0.0571	0.153	-0.0232
	(0.0667)	(0.0849)	(0.131)	(0.0876)
Time, Fake, and House (DDD)	**0.00644**	**0.0242**	**-0.0753**	**0.105**
	(0.0814)	**(0.138)**	**(0.212)**	**(0.158)**
One Year-Six Months Prior	Base	Base	Base	Base
	-	-	-	-
Six Months Prior-Disclosure Date	0.128***	0.188***	0.316***	0.190***
	(0.0257)	(0.0440)	(0.0565)	(0.0437)
Disclosure Date--Six Months Post	-0.224***	-0.348***	-0.383***	
	(0.0309)	(0.0586)	(0.0754)	
Six Months -One Year Post	-0.0150	-0.0275	0.0524	0.0167
	(0.0302)	(0.0559)	(0.0789)	(0.0569)
One Year Post- 18 Months Post				-0.433***
				(0.0525)
Observations	11,160	11,160	11,120	11,160
R-squared	0.283			
Number of Unique Wines	279	279	278	279

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

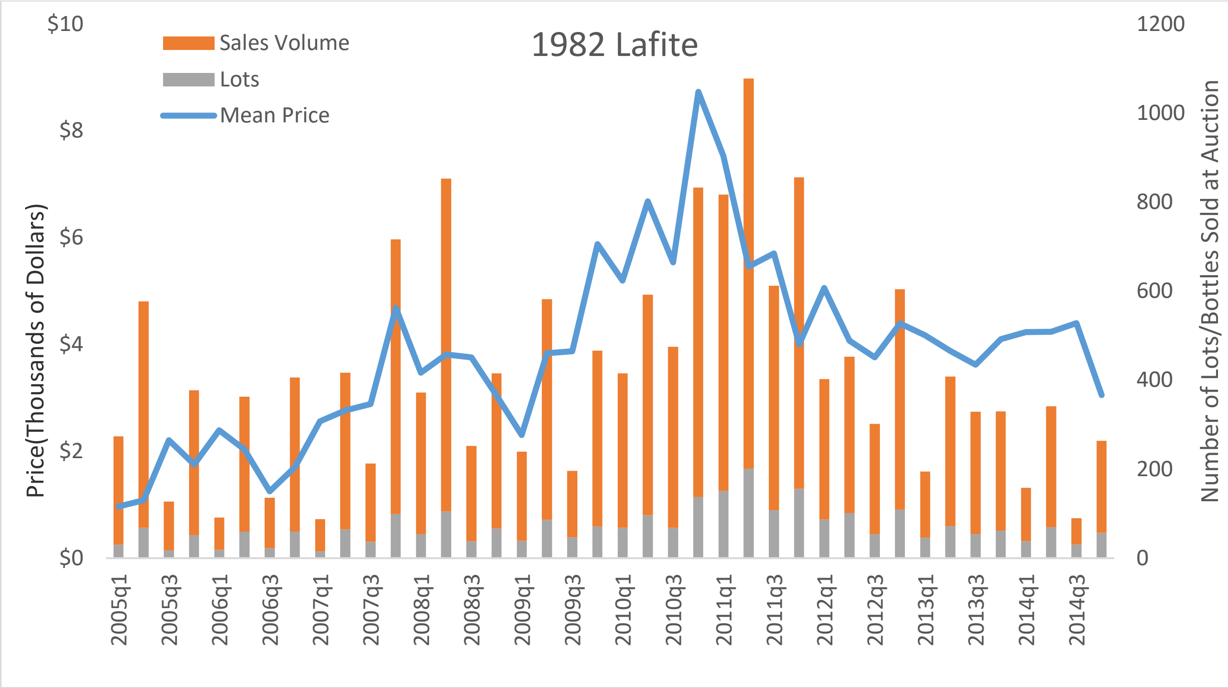


Figure 1.1: Sales Patterns

CHAPTER TWO:

TRANSITIVITY OF STATED PREFERENCES AND INDIVIDUAL CHARACTERISITICS

1. Introduction

One of the most common assumptions used in economics states that, given a set of feasible options, individuals will always choose the one that maximizes their well-being. Revealed preference theory allows for a variety of rationality tests regarding this assumption when multiple choices from the same individual can be observed. To satisfy the rationality tests, preferences must remain complete and transitive.

In this chapter, I study the rationality of stated preferences (referred to in the following as preferences) by examining whether they satisfy the principle of transitivity. The principle of transitivity states if A is revealed preferred to B and B is revealed preferred to C, then it should not be the case that C is revealed preferred to A. Any response pattern failing to meet this criteria represents a failure of transitivity and is inconsistent with utility maximizing behavior.

A number of studies have utilized transitivity tests to study rationality, and it has been widely observed that individuals do not always satisfy transitivity (Lancsar & Louviere, 2006). Yet, in the field of health specifically, most studies have only evaluated transitivity on a binary scale (i.e., investigating only whether an individual's preferences are transitive), ignoring both the frequency and severity of any transitivity violations.[29] A few studies have combined the frequency and severity of transitivity violations into a single measure, but no study has separated the two into individual measures of transitivity (Devlin, Hansen, Kind, & Williams, 2003; Lamers, Stalmeier, Krabbe, & Busschbach, 2006). This is likely due to the joint identification problem of quantity and severity. Using only a single measure of transitivity, it is impossible to differentiate respondents suffering multiple, less severe, failures of transitivity from those exhibiting fewer, but more severe, failures. There may be important differences between these individuals that go unnoticed when the quantity and severity of transitivity failures are not evaluated individually.

I solve the identification problem by including two transitivity measures (Count and Order), one of which is nested inside the other, to evaluate both the quantity and severity of failures of transitivity displayed by each individual. The Order Measure of transitivity calculates the number of transitivity failures, while the Count Measure of transitivity combines both the number and severity of transitivity failures into a single value. The average severity of transitivity failures can then be determined by dividing values for the Count Measure by the Order Measure.

I employ these measures using responses from the Timing, Duration, and Lifespan (TDL) Study, an online, cross-sectional, survey of adults from the U.S. general population. In addition to various questions regarding each individual's demographic and socioeconomic characteristics, respondents were asked to complete a series of questions regarding their preferences for health. These questions asked individuals to state their preferences over trade-offs between a longer

[29] The severity of a transitivity violation has been measured in a variety of ways. Some studies measure severity by calculating the required adjustment to the budget constraints for choices to remain rational. A failure of transitivity requiring a larger adjustment is determined as more severe.

lifespan with health problems and a shorter lifespan with no health problems. By varying the intensity of health problems and lifespan over a series of 30 questions, we are able to determine whether each respondent remained transitive in their preferences.

In addition to measuring the number and severity of non-transitive preferences exhibited by each respondent, another aim of this manuscript is to evaluate the relationship between response transitivity and the individual characteristics of each respondent under the premise that high quality decisions are the result of greater decision-making ability. Specifically, I test the hypothesis that fewer failures of transitivity are associated with positive real-world outcomes.

The two real-world outcomes I study in this paper include income and health. An individual's income is the result of a lifetime of choices. The better decisions an individual makes, the more likely they are to earn higher incomes. A similar rational may be used when evaluating the relationship between failures of transitivity and health. While some aspects of health are predetermined by outside factors, including genetics and geographic environment, many others represent a choice (e.g., diet, activity level, annual check-ups). Even after controlling for education and demographic characteristics, individuals with more non-transitive responses may also be more likely to make poor health-related decisions.

I find that failures of transitivity are common for the individuals in my sample; more than 52% of respondents suffer at least one failure of transitivity during the experiment. In addition, greater numbers of transitivity failures are associated with lower incomes and poorer health. Specifically, a standard deviation decrease in the number of transitivity failures is associated with a $3,100 to $3,400 increase in household income. Conditional on an individual suffering from at least one failure of transitivity, more severe failures are also associated with lower incomes. These findings largely align with previous preference research on the topic, suggesting that individuals with a greater decision-making ability enjoy better real-world outcomes (Brown, Kapteyn, Luttmer, & Mitchell, 2017; Choi, Kariv, Muller, & Silverman, 2014; Echenique, Lee, & Shum, 2011).

The remainder of the chapter is organized as follows. Section 2 contains a review of the literature analyzing stated preferences, while Section 3 outlines the transitivity measures used in this manuscript. The estimation strategy and methodology are provided in Section 4, and Section 5 describes the data. The results are presented in Section 6. Section 7 discusses the major findings, implications, limitations, and future work. Conclusions are summarized in Section 8.

2. A Review of the Literature

2.1 Rationality and Real-World Outcomes

Recently, a number of papers have examined the relationships between the rationality of preferences and real-world outcomes. Saelensminde (2002) studied individual preferences for public transport and found that nearly two-thirds of the respondents reported at least one failure of transitivity. Individuals with lower education levels were found to be more likely to respond in a non-transitive manner.

Other research has focused on the transitivity of financial decisions. Echenique, Lee, and Shum (2011) measured the severity of violations to the general axiom of revealed preferences (GARP).[30] The severity of each violation is equivalent to the amount of money an individual could save by making rational decisions. The authors used scanner-level data to track grocery store purchases and found, on average, households "waste" about 6% of their monthly food

[30] GARP extends the weak axiom of revealed preferences (WARP) by allowing for ties. GARP states that, if "A" is revealed preferred to "B" when both are available, then "B" must not be strictly revealed preferred to "A."

expenditures as a result of irrational purchasing patterns. Older, less educated, and poorer households were found to make more severe violations.

Studies by Choi, Kariv, Muller, and Silverman (2014) and Brown, Kapteyn, Luttmer, and Mitchell (2017) evaluated the relationships between transitivity and real-world outcomes in experimental settings. The results from each were largely supportive of the findings presented by Echenique, Lee, and Shum. Brown et al. used an annuity valuation experiment of social security benefits. In addition to age and education, the authors also found associations between irrational responses and individuals with weaker numerical abilities.

Choi et al. took the additional step of investigating the causal relationship between rationality and wealth in an analysis of risk preferences. Controlling for income, education, work status, and other individual characteristics, stronger adherence to GARP was found to correspond to greater household wealth. The intuitive nature of these results helps provide a basis for the present study.

2.2. Transitivity in the context of Preferences for Health

Multiple studies have also evaluated the transitivity of health preferences (Bleichrodt & Prades, 2009; Engel, Bansback, Bryan, Doyle-Waters, & Whitehurst, 2016; Johnson & Mathews, 2001; Lamers et al., 2006; McIntosh & Ryan, 2002; Ryan, Watson, & Entwistle, 2009; Schwappach & Strasmann, 2006; Yang, van Busschbach, Timman, Janssen, & Luo, 2017).[31] Some of these used a similar set of questions to those asked in the TDL Study we employ, whereas others investigated the transitivity of preferences for medications, surgeries, screening tests, and physician visits. In nearly all of the studies, failures of transitivity were common occurrences. A study by Devlin,

[31] In health, failures of transitivity are often described as a "logical inconsistency."

Hansen, Kind, and Williams (2003) focusing on individual preferences for health found that nearly 80% of respondents reported at least one failure over a series of 13 choice tasks. Non-transitive preferences have been shown to be correlated with age, education, and socioeconomic status (Andrade, Noronha, Kind, Reis, & de Carvalho, 2016; Badia, Roset, & Herdman, 1999; Craig & Ramachandran, 2006; Devlin et al., 2003; Dolan & Kind, 1996).

Apart from transitivity, other commonly used rationality tests to evaluate preferences for health involve monotonicity and preference reversals. The monotonicity assumption of normal goods states that individuals should never choose a weakly dominated option from a given set of choices. Preference reversals involve a situation in which an individual first states that A is preferred to B and then later states that B is preferred to A. Three of the most relevant studies to examine the rationality of stated preferences for health are described below.

San Miguel, Ryan, and Amaya-Amaya (2005) examined the rationality of preferences for physician visits using tests for monotonicity and preference reversals. The authors found that tests for preference reversals captured more failures of rationality than those involving monotonicity. Younger and older individuals and those with less education were also discovered to be more likely to exhibit irrational preferences.

A study by Ozdemir, Mohamed, Johnson, and Hauber (2010) used four rationality tests (monotonicity, preference reversals, and two involving transitivity) to analyze preferences for medical treatments. The results showed that, of the four tests, respondents were most likely to fail the less-restrictive transitivity measure.[32] Specifically, 32% of respondents failed the transitivity measure, 25% had at least one preference reversal, and 18% failed the monotonicity test. This finding suggests that individuals are more likely to break transitivity than the other

[32] This measure is very similar to the one used in this study (described in Section 3.1).

axioms of rationality. Results from the four tests were then combined into a single binary variable (i.e., to determine whether respondents failed at least one of the tests) to study the relationship between irrational preferences and individual characteristics. Male respondents, individuals with fewer years of education, and those with lower incomes were more likely to fail at least one of the rationality tests.

Al Sayah, Johnson, Ohinmaa, Xie, and Bansback (2017) assessed the rationality of health preferences using a combination of preference reversals and monotonicity. Respondents were deemed inconsistent (i.e., they failed the rationality test) if they gave a weakly dominated option an equivalent or higher value than a dominant one. Using a lenient rationality test that only captured severe errors, the authors still found that more than 11% of respondents exhibited irrational preferences. Individuals with inadequate health literacy (assessed using the three-question Brief Health Literacy Screen), lower incomes, and more than 59 years of age were significantly more likely to fail the rationality test.[33]

3. Examining the Transitivity of Preferences

3.1 Experimental Design

Economic valuation studies are often performed to help determine preferred options when budgets are limited. One of the fastest growing and increasingly popular valuation methods involves the use of discrete choice experiments (DCEs) (de Bekker-Grob, Ryan, & Gerard, 2012). Frequently, these questions present hypothetical situations mimicking choices individuals have previously faced or are expected to face in the future. Focusing on health specifically,

[33] Consistency was also found to be correlated to chronic health problems. However, the only significant result compared individuals with at most one chronic condition (from a list of 15) with those who had at least three chronic conditions. The inclusion of two other health related questions in the regression model may have also confounded the results.

DCEs often require respondents to compare the opportunity costs of potential health interventions. A cancer treatment drug may extend lifespan but lower expected quality of life, and a surgery to relieve back pain could decrease mobility. By varying the potential outcomes over a choice set, it is possible to measure individual and societal preferences for health.

The TDL Study asked individuals to complete a set of 30 paired comparison choice tasks using health states described by the EQ-5D-5L, one of the most commonly used instruments to measure health outcomes, along with a lifespan attribute (Euroqol, 2018; Herdman et al., 2011).[34] The EQ-5D-5L has five items (domains) regarding an individual's problems with mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[35] Responses to each domain have five incremental levels, and each increase is indicative of a greater severity of problems (no problems, slight, moderate, severe, and extreme problems). Figure 1 provides a sample version of the instrument. Since 2008, the EQ-5D has been the preferred measure of health-related quality for all health technology appraisals conducted by the National Institute for Health and Care Excellence (NICE) in the United Kingdom (NICE, 2017).

An example question may ask respondents to state a preferred option between the following choices:

A. 30 days with the following health problems: severe problems with mobility, severe problems with self-care, severe problems performing usual activities, moderate pain or discomfort, and severe anxiety or depression. Following these 30 days, the individual will experience 25 days with no health problems

B. 7 days with no health problems

[34] An example version of the questionnaire is included in the appendix.
[35] Specifically, the EQ-5D refers to mobility as an individual's ability to walk about, self-care as the level of difficulty an individual has washing or dressing oneself, and usual activities includes work, study, and family/leisure activities. There are no further descriptions for pain/discomfort or anxiety/depression.

where, for each question, the pairs present a tradeoff between lifespan and health. Option A is composed of a longer lifespan while option B has a shorter lifespan with no health problems (*TB*). Option A can be decomposed into two separate parts: time spent with health problems (*HS)*, and time with no health problems (*TA*). Each option begins "today" and culminates in death (i.e., the respondent dies in 55 days in option A and in 7 days in option B).[36]

Individuals were also randomly assigned to one of six time groups at the beginning of the survey corresponding to the duration of time in health state *HS* (1 day, 7 days, 30 days, 9 weeks, 12 months, or 10 years) for all of the 30 questions. For each individual, *TA* and *TB* also followed the same time frame (days, weeks, months, or years) as *HS*.

Within each of the six time durations, questions varied by EQ-5D-5L health state *HS*, duration of time with no problems following the time with health problems *TA*, and the duration of time with no health problems as the alternate choice *TB*. The pair selection was randomized by individual, and no two respondents were exposed to the same 30 questions. In total, the sample included 1,965 unique paired-comparison questions. A similar design has been used in previous health valuation experiments (Craig et al., 2014; de Bekker-Grob et al., 2012; Oppe, Devlin, van Hout, Krabbe, & de Charro, 2014).

Before completing the paired comparison choice tasks, each respondent was also asked to complete a series of questions regarding their demographic and socioeconomic characteristics, as well as current health. Health was assessed using the EQ-5D-5L questionnaire (Figure 1), and the

[36] Including death at the end of each option is required to determine the logical order of each health state. Without death, it is impossible to state that 10 years with slight health problems has a greater value than 1 year with severe health problems. Individuals with no health problems may see the 10 years as a negative attribute, while those currently experiencing moderate/severe health problems are likely to view the longer time with only slight problems as a positive attribute.

socioeconomic questions provide details regarding an individual's household income, work status, and level of education.

3.2 Defining a Failure of Transitivity

Assuming that preferences remain stable over the course of the experiment, all failures of transitivity meet the following two conditions. First, question 2 must contain a health state that weakly dominates the health state provided in question 1 (i.e. one health state always has a lower or equivalent level of problems than the other across every domain; $HS_2 \geq HS_1$). Second, the individual must prefer a shorter lifespan with no health problems (option B) in question 2 and a longer lifespan with health problems (option A) for question 1. In addition, one of the following three specifications must also be met:

1) $TA_2 \geq TA_1$ and $TB_1 \geq TB_2$
2) $TA_1 \leq TB_1$ or $TA_2 \geq TA_1$, and $TB_2 = 0$
3) $TB_1 - TA_1 \geq TB_2$.

A visual example for a failure of transitivity is provided in Figure 2.

3.3 Measuring Each Failure of Transitivity

I use two methods to examine each failure of transitivity. Among four options, suppose that option A objectively dominates options B and C (e.g., A has fewer health problems than either B or C across all domains). Assume that over the course of three choice tasks, an individual reveals that options B and C are both preferred to option D, but option D is preferred to option A. Should this be counted as two failures of transitivity, given that there are two sets of responses failing the transitivity test? Or, because both failures originate from the same response (option D revealed preferred to A), is this better represented as a single failure of transitivity? In this study, I choose to utilize both methods for the analysis. The Count Measure classifies this response pattern as two failures of transitivity, whereas the Order Measure counts this as a single failure.[37] A visual representation of both measures is provided in Figure 3.

The Order Measure of Transitivity, denoted $tran_o$, is nested within the Count Measure, $tran_c$ (i.e., the Count Measure includes all failures observed using the Order Measure). By definition, $tran_c \geq tran_o$. If an individual suffers only a single failure of transitivity, both measures will have a value of one. However, if the same question is involved in multiple failures, it will only be counted once with the Order Measure. Conceptually, $tran_o$ represents the fewest number of responses that would need to be changed in order for all responses to behave in a transitive manner.

The Count Measure of Transitivity, $tran_c$, represents the total number of response pairs that generate a failure of transitivity. With the Count Measure, a more irrational response is likely to be connected to multiple failures of transitivity. The Count Measure combines both the frequency and severity of transitivity failures into a single value. By incorporating the values of the Order Measure with the Count Measure, we are able to individually examine both the frequency and severity of transitivity failures suffered by each respondent.

3.4 Determining the Severity of Each Failure of Transitivity

In a financial experiment in which each item in a bundle has a given monetary value (e.g., grocery items, investments), multiple mechanisms have been developed to evaluate the severity of each failure of transitivity (Afriat, 1972; Echenique et al., 2011; R. Varian, 1994). The severity of each failure is often calculated by examining the fraction of money that is "wasted"

[37] Rezaei and Patterson (2015) used a similar setup to study the effects of transitivity failures on population values for transportation. However, the authors did not measure transitivity on a respondent level. Instead, single responses were removed if they met certain criteria (e.g., a response was linked to at least two failures of transitivity).

on irrational choices by each individual (Choi et al., 2014). However, these mechanisms are difficult, if not impossible, to implement when multiple domains of each bundle are on different scales. In the context of the TDL study, each choice task involves a trade-off between a longer lifespan and better health. Because respondents have heterogeneous preferences, there is no method to directly compare the severity of a failure of transitivity involving lifespan to one involving health. For example, which is failure of transitivity is more severe? Responses to two questions reveal a 5-year lifespan is preferred to a 10-year lifespan, or, responses reveal a health state with slight health problems is preferred to one with extreme health problems.

I develop a new method to measure the average severity of transitivity failures for each respondent by dividing the values from the Count Measure by the Order Measure ($tran_c/tran_o$). Using the Count Measure, it is likely that a response connected to a greater number of transitivity failures would display a larger dissonance from utility maximizing behavior. However, this would only be considered as a single failure of transitivity using the Order Measure. While this measure is certainly cruder than the financial measures mentioned above, it has a wider range of potential applications (e.g., in transportation and health).

4. Estimation Strategy and Methodology

4.1 Number of Opportunities for a Failure of Transitivity to Occur

The random selection of pairs provided to each respondent affects the number of chances (denoted, *chan*) that a failure of transitivity can occur. Comparisons can only be made when the health state in one pair weakly dominates the health state in the other across all domains. In the TDL study, many of the questions are not directly comparable. The health state in one pair may have more severe problems with anxiety/depression, while the health state in the other may present greater problems with mobility. Figure 4 displays a set of questions from the TDL study where one health state does not objectively dominate the other. Values for *chan* are likely correlated with the number of times that a respondent fails transitivity.

4.2 Willingness to Trade

Willingness to trade also has an effect on the potential number of times that an individual's responses can constitute a failure of transitivity. In each paired comparison question, individuals were required to state their preferences between a longer lifespan with health problems and a shorter lifespan with no health problems. For a failure in transitivity to occur, an individual must display a willingness to make tradeoffs between lifespan and quality of life (denoted, *trade*). When an individual splits the number of times they choose between longer lifespan and no health problems evenly, the potential number of times that transitivity can be broken also increases.

Previous research regarding preferences for health has found that a significant portion of the population will always choose the option with the longer lifespan (Fowler, Cleary, Massagli, Weissman, & Epstein, 1995; Nord, Daniels, & Kamlet, 2009).[38] Yet, many previous studies have failed to exclude these individuals from their analyses, confounding the results. The answers from individuals exhibiting lexicographic preference are uninformative with respect to the rationality test performed in this study; therefore, I exclude these respondents from the analysis. A comparison between individuals displaying lexicographic preferences and those with continuous preferences is provided in the appendix.

[38] Apart from the individuals who always choose the option with the longer lifespan, some individuals may always select the option with fewer health problems.

4.3 Empirical Model for Income and Transitivity

I employ a series of ordered probit and interval regression models to estimate the relationship between failures of transitivity and two real-world outcomes: income and health. Ordered probit models are often used when the dependent variable is observed as an ordered outcome. While they are useful for determining the sign and significance of a relationship between the independent and dependent variables, ordered probit models have one major limitation; it is often difficult to interpret the partial effects of each independent variable on the outcome of choice (Greene, 2012).

If the ordered levels have defined intervals (e.g. income categories), it may be preferable to use an interval regression approach instead.[39] Ordered probit models estimate the cut points between levels (e.g., slight and moderate pain), while interval models have the ability to directly measure the effect of the parameters on the outcome of interest (Wooldridge, 2010).

Focusing on income, the interval regression models are able to estimate changes in income associated with each additional failure of transitivity. The basic model for income is provided in Equation 4.1 below:

$$income_i = \beta_0 + \beta_1 tran_i + \beta_2 trade_i + \beta_3 chan_i + \beta_4 X_i + \beta_5 health_i + \varepsilon_i \qquad (4.1)$$

where *X* controls for education, age, employment status, race, gender, ethnicity, marital status, and geographic region of residence (e.g., New England). *tran* represents the number of times over the course of the survey that an individual's responses failed transitivity. Regressions for each model specification are run using the Count and Order Measures as dependent variables separately. Given that the Count Measure includes all failures of transitivity reported in the

[39] The overall setup for an interval regression is nearly identical to that of an ordered probit. Both models assume that the error term has a standard normal distribution. The main difference is that, for the interval regression models, the cut points between intervals have a defined value. With the interval regression model, we are able to interpret the model coefficients in the same manner as a linear regression.

Order Measure as well as the number of additional failures to which each response can be linked, it is expected that the coefficients for β_1 will be larger using the Order Measure.

trade is determined using the least frequently chosen of the two options (longer lifespan or better health) over the 30 paired comparisons. For example, if a respondent chose a longer lifespan with health problems for 20 of the 30 paired comparison tasks, they displayed a willingness to trade 10 times (i.e., the number of times a shorter lifespan with no health problems was selected). In this study, *trade* ranges from 1 to a maximum of 15. The relationship between *trade* and *tran* is likely non-linear in nature; as *trade* increases, *tran* is expected to increase at a decreasing rate.[40] To account for this, I also include $trade^2$ as an additional dependent variable in the regression models. Some specifications even include $trade^3$.

On average, respondents were provided 153 chances to fail transitivity with an interquartile range from 136 to 170. To account for individual differences, the regressions include both *chan* and $chan^2$.

4.4 Empirical Model for Health and Transitivity

Because the health questions lack a defined interval for each level of problems, the relationship between transitivity and health is evaluated using only ordered probit models. Equation 4.2 provides the basic model:

$$health_i = \beta_0 + \beta_1 tran_i + \beta_2 trade_i + \beta_3 chan_i + \beta_4 X_i + \beta_5 income_i + \varepsilon_i \qquad (4.2)$$

where $health$ represents responses to one of five health-related questions in the questionnaire (Figure 1). *income* is again a categorical variable, and the other independent variables in

[40] Using Figure 2 as an example, suppose all comparable health states provided to each respondent are ordered from most to least severe. As the disparity between the health states or lifespans increases, it is expected that the likelihood of a failure of transitivity will decrease.

Equation 4.2 are identical to those from Equation 4.1.

5. Data

5.1 Description of the TDL Study

The TDL Study was conducted from December 16, 2015 through January 11, 2016 using participants from a nationally representative online panel. A total of 3,909 respondents completed the survey. However, more than 17% (675) exhibited lexicographic preferences. 12% or respondents always chose the option with the longer lifespan, and 5% always selected the option with no health problems. After dropping these individuals, the final sample was reduced to 3,234 respondents.

5.2 Description of Independent Variables

To evaluate the relationship between transitivity and income/health, the regression models also control for an individual's education, age, gender, marital status, race, ethnicity, and region of residence. Apart from marital status and region of residence, the categories for each of these variables are provided in Table 1. The marital status variable categorized each individual as married, widowed, separated or divorced, never married, or living with a partner. In addition to its possible association with an individual's health, marital status is expected to be an important factor in determining household income. Given the potential for a second earner in the household, it is likely that individuals who are currently married or living with a partner have higher incomes than those who are single. Health and income may also vary by region. Using the U.S. Census definitions, each respondent was categorized into one of nine geographic regions according to their state of residence (U.S. Census Bureau, 2015).[41]

5.3 Description of Dependent Variables

5.3.1 Income

Respondents self-categorized their income into one of 11 groups.[42] The categories ranged from a household income of less than $15,000 per year ($income_i = 1$) to an income of at least $500,000 ($income_i = 11$). It is expected that individuals with higher household incomes will demonstrate greater levels of transitivity.

5.3.2 Health

There are well-established links between health and socioeconomic status (Smith, 1999), but few studies have examined the relationship between rationality and health. Many of the actions commonly associated with a healthy lifestyle must be practiced regularly to remain effective, and these actions are likely correlated with health status and literacy. Medicines should be taken following a prescribed regimen, exercise should be part of a daily routine, and safety precautions should be practiced at all times. Individuals who display higher levels of transitivity, a proxy for decision-making ability, may be more likely to practice behaviors that result in better health.

Separate regressions are run using responses to each domain of the EQ-5D-5L questionnaire (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) as the dependent variable. Each domain uses a similar five-level index, where a one level increase

[41] We could have also categorized individuals by state, but there were multiple states with less than 10 respondents. By comparison, the smallest region (East South Central) contained 178 respondents.
[42] Some individuals may consider income a sensitive subject; therefore, respondents were also allowed to state that they didn't know or could refuse to answer the question. Of the 3,234 completed surveys, 247 individuals (8%) chose to withhold this information. Missing is included as an income category for the regressions using health as the dependent variable.

corresponds to a greater severity of problems experienced (i.e., no problems, slight problems, moderate problems, severe problems, or extreme problems).

6. Results

Over the course of the experiment, respondents suffered an average of 2.70 failures of transitivity using the Count Measure and 0.99 failures of transitivity using the Order Measure. Conditional upon an individual trading between lifespan and health at least once, nearly 52% of respondents registered at least one failure of transitivity. Table 1 summarizes failures of transitivity across various demographic and socioeconomic groups. In general, older individuals, and those with higher education levels suffered fewer failures of transitivity than their younger, less educated counterparts.

As expected, individuals who displayed a greater willingness to trade between lifespan and health were more likely to experience failures of transitivity. Figure 5 displays the frequency individuals selected the option incorporating a longer lifespan with health problems along with the mean number of transitivity failures for each group. On average, individuals who traded more than 10 times suffered more than twice as many transitivity failures than those who traded five times or fewer (3.93 vs. 1.71 using the Count Measure and 1.43 vs. 0.64 using the Order Measure).

6.1 Transitivity and Income

Ordered probit and interval regression models are used to examine the relationship between income and transitivity. The results for the Count and Order Measures are presented in Tables 2 and 3. Estimates for the ordered probit model are displayed in Column 1, while Columns 2 to 5 display estimates using the interval approach.

Using the Count Measure of transitivity (Column 2 of Table 2), a standard deviation decrease in the number of non-transitive responses is associated with $3,100 increase in household income. If we restrict the sample to individuals who are currently working (Column 3), the effect rises to nearly $4,300.

In this study, we are unable to differentiate between single-income and multi-income families. The income question specifically asked, "What is your best estimate of your total income plus the total income of all family members from all sources?" Column 4 of Table 2 provides estimates for individuals who are currently married. Focusing on married individuals does not solve the single/dual income problem, but it does help to even the playing field by eliminating responses where there is no potential of a second income. The similarity between these results and those from the base model helps strengthen the argument that greater levels of transitivity are associated with positive real world outcomes. A standard deviation decrease in failures of transitivity is associated with a $3,000 increase in household income.

Column 5 provides estimates for individuals who were at least 26 years old at the time of the survey. Younger individuals may still be in school and, for those in the labor force, there may also be a smaller wage spread between education levels. Interestingly, this model results in the smallest coefficients for each failure of transitivity. However, the results remain highly significant.

In general, the estimates using the Order Measure of transitivity (Table 3) are largely congruent with those using the Count Measure. The coefficient estimates for each failure of transitivity are larger using the Order Measure, but the Count Measure captures more failures of transitivity. Using the base model (Column 2 of Table 3), a standard deviation decrease in failures of transitivity is associated with a $3,400 increase in household income. If we focus on individuals who are currently working (Column 3), the Count and Order Measures provide nearly identical results. A standard deviation increase in failures of transitivity corresponds to a $4,300 decrease in household income.

The largest difference between the Order and Count Measures for any of the regression models involved married individuals. A standard deviation decrease in failures of transitivity is associated with an increase in household income of more than $3,600 using the Order Measure (Column 4); 20% higher than the estimate using the Count Measure. Looking at individuals at least 26 years of age (Column 5), we again observe a smaller effect for transitivity on household income than in the base model.

6.1.1. Severity of Failures and Income

Given that the Order Measure of transitivity is nested within the Count Measure, it is not surprising to see that both measures provide similar estimates for the relationship between transitivity and income. The two measures are strongly correlated to one another (ρ=.88). The main benefit from including both measures in the study arises through our ability measure the average severity ($severity=tran_c/tran_o$) of transitivity failures for each respondent.

By itself, the average severity of transitivity failures provides limited information concerning respondent rationality. A respondent with high average severity is not necessarily more irrational than someone with a lower one. However, by also controlling for the Order measure of transitivity in the regression models, we can determine whether the severity or number of transitivity failures has a stronger relationship with household income.

Results from models evaluating the relationship between the severity of transitivity failures and income are displayed in Table 4. For these regressions, the sample was restricted to respondents who reported at least one failure of transitivity during the experiment. In the sample, values for *severity* ranged from 1 to 17 with a mean of 2.32. In other words, each irrational response (using the Order Measure) was often associated with multiple failures of transitivity. Higher values for *severity* represent a larger conflict with utility maximizing behavior.

Estimates using interval regression models are displayed in Columns 1 and 2, and those using ordered probits are presented in Columns 3 and 4 of Table 4. Even after controlling for the number of transitivity failures (using the Order Measure) the results show more severe failures of transitivity are associated with lower incomes. Conditional on the respondent suffering at least one failure of transitivity, a standard deviation increase in the average severity of each failure coincides with a $1,300 decrease in household income (Column 1). For respondents who are currently employed, the value increases to nearly $4,800 (Column 2). As with Tables 2 and 3, the ordered probit models (Columns 3 and 4) produce estimates of a similar sign and significance to the interval regression models.

Comparing the relationships between the severity and quantity of failures of transitivity and household income, it appears that the frequency of transitivity failures has a stronger association with income using the full sample (Column 1). A standard deviation decrease in the number of transitivity failures is associated with a $3,600 increase in household income, nearly three times the effect of a standard deviation decrease in severity. However, after restricting the sample to individuals who are currently employed, the severity of failures has a stronger relationship. This finding hints that both the frequency and severity of transitivity failures are important predictors for an individual's overall level of rationality.

6.2 Transitivity and Health

It is also intriguing to examine the relationship between failures of transitivity and current health. The results are displayed in Columns 1 to 5 of Tables 5 and 6 and show failures of transitivity are associated with two health outcomes in two domains: self-care and usual activities. The results are fairly similar for both transitivity measures; however, the relationship between health and transitivity appears to be more significant using the Order Measure.

The association between both transitivity measures and problems with self-care is highly significant ($p<0.01$), and it appears individuals experiencing problems with self-care are the driving force for the association found between failures of transitivity and problems with usual activities. Of 3,234 respondents included in the final sample, 156 reported problems with self-care; of these individuals, more than 90% also reported problems with usual activities. When self-care is added as an independent variable to the regression model for usual activities, or when respondents experiencing problems with self-care are excluded from the model, we no longer find any significant association between transitivity failures and usual activities.

A variety of other health-related dependent variables were also used to assess the relationship between failures of transitivity and health. Two of these include binary measures denoting whether the respondent is experiencing multiple (problems on at least two domains of the EQ-5D-5L) or intense health problems (severe or extreme problems with at least one domain). The results, estimated using a binary probit model, are displayed in Columns 6 and 7 of Tables 5 and 6. Although the estimates are insignificant for each of the transitivity measures, there does appear to be a relationship between failures of transitivity and intense health problems. Of the 3,234 respondents who displayed a willingness to trade between lifespan and health, only 175 were currently experiencing intense health problems on any domain. A larger sample size of these respondents may result in a significant finding.

6.2.1 Severity of Failures and Health

Table 7 displays the estimates when the severity of transitivity failures is added as an additional independent variable to the base model. Unlike the regression models for income, the coefficients for the quantity and severity of failures of transitivity have opposite signs when health is the dependent variable. An increase in the quantity of non-transitive responses is associated with greater problems with self-care and usual activities as well as a greater likelihood of experiencing intense health problems, but an increase in the severity of each failure of transitivity is associated with fewer health problems.

7. Discussion

7.1 General Findings

Nearly 52% of the respondents suffered at least one failure of transitivity during the experiment. Overall, these individuals averaged 2.70 failures of transitivity using the Count Measure and 0.99 failures using the Count Measure. I also find that failures of transitivity are associated with lower household income and greater health problems. These findings largely align with previous preference research on the topic. Focusing on income, I observe a positive relationship between transitivity and household income. A standard deviation decrease in the number of failures of transitivity is associated with a $3,100 to $3,400 increase in household income. When the sample is restricted to individuals that are currently employed, this number rises to $4,300.

Concerning the relationship between transitivity and health, I find problems with self-care has the strongest association with non-transitive preferences. There has been a substantial amount of research outside economics linking cognitive ability with the struggle to care for oneself (Cramm et al., 2013; Fong, Chan, & Au, 2001; Levinthal, Morrow, Tu, Wu, & Murray, 2008; Mottus et al., 2014), and my results offer additional confirmation for these earlier findings.

Restricting the sample to individuals who are currently working or below the age of 65 does not diminish the magnitude or significance of the estimates.

7.2 A Comparison of the Count and Order Measures of Transitivity

Both the Count and Order Measures use the same requirements to determine each failure of transitivity. The difference between the two measures arises through the manner in which each failure of transitivity is counted. The Order Measure determines the number of transitivity failures, whereas the Count Measure also accounts for the severity of each error. The fact that both measures produce similar results is somewhat expected. The main value provided by calculating both measures stems from the ability to examine both the quantity and severity of non-transitive responses. I discover that more severe transitivity violations are associated with lower incomes but not with greater health problems.

7.3 Robustness Checks

Previous studies have identified two concerns that may affect how the results from this study can be interpreted. The transitivity tests require preferences to remain stable throughout the experiment, yet it has been widely acknowledged that there may be a learning curve in many experimental settings as respondents adjust to an unfamiliar task (Johnson & Mathews, 2001; Lancsar & Louviere, 2006). It is unlikely that many individuals have ever thought about health in such a manner, and preferences may change if a respondent has not previously considered preferences for life and death.[43] To address this, the regression models were rerun after dropping

[43] The TDL offered three warm-up problems to help respondents get accustomed to the survey interface and type of questions that would be asked. Respondents were also allowed to return to previous questions to change any answers. However, given that respondents were provided a fixed amount of money for completing the study, we find it unlikely that many would return to change previous responses on account of a shift in preferences midway through the experiment.

responses to the first three questions. The results using the Count Measure of transitivity are displayed in Tables 8 and 9. As it can be seen, the results shown here are similar to those displayed in Tables 2 and 5. Failures of transitivity are still associated with lower household incomes and problems with self-care.

7.3.1 Transitivity and Moderate Health States

The second concern involves potential issues with the level descriptions of the EQ-5D-5L. To evaluate transitivity, the levels of each domain must have a dominant order; having moderate problems must be at least as detrimental as slight problems. One potential issue with this questionnaire presents itself through comparisons made between severe and extreme problems with pain/discomfort and anxiety/depression. The EQ-5D-5L orders “extreme problems” worse than “severe problems,” yet a previous study found many respondents viewed “severe problems” to be worse than “extreme problems” (Craig, Pickard, & Rand-Hendriksen, 2015). To account for this, a model specification was run that dropped all comparisons between health states where both had severe or extreme problems.[44] As shown in Tables 8 and 9, the results again appear to be highly consistent with the findings from the base models.

7.4 Study Limitations and Future Work

One of the major limitations of this study regards the requirement that individuals be willing to trade at least once during the experiment. Studies incorporating a wider variety of health states may reduce the prevalence of individuals displaying lexicographic preferences, but it is possible that some individuals may never be willing to trade between lifespan and better health. Logit models (displayed in the appendix) comparing individuals who traded at least once to those

[44] Comparisons between extreme/moderate and severe/moderate health states were still used, as there has been no evidence of any issues between these levels.

exhibiting lexicographic preferences found significant differences in the characteristics of each group, but it is possible that other variables not included in the TDL Study (e.g., religious preferences) also play a role. The fact that we no longer have a representative sample may reduce the generalizability of our results.

Another limitation of this study involves the amount of personal information available for each respondent. Ideally, future studies will obtain more information about each respondent's health (including history of diagnosed conditions, family history, and certain health behaviors) and socio-economic status (e.g., job description, individual income, or wealth). Within the current framework, it is difficult to prove a causal relationship between transitivity of preferences, as a proxy for decision-making ability, and individual outcomes.

8. Concluding Remarks

This study analyzes the transitivity of stated preferences for health as a possible proxy for decision-making ability. I employ a novel approach to measure individual preferences using two measures of transitivity. Using a large sample of respondents from the U.S. general population, I find failures of transitivity in a health valuation experiment to be associated with various socio-economic characteristics. Non-transitive patterns are associated with lower household incomes and poorer health.

References

Afriat, S. N. (1972). Efficiency estimation of production functions. *International Economic Review*, 568-598.

Al Sayah, F., Johnson, J. A., Ohinmaa, A., Xie, F., Bansback, N., & Canadian Eq, D. L. V. S. (2017). Health literacy and logical inconsistencies in valuations of hypothetical health states: results from the Canadian EQ-5D-5L valuation study. *Quality of Life Research, 26*(6), 1483-1492. doi:10.1007/s11136-016-1495-z

Andrade, M. V., Noronha, K., Kind, P., Reis, C. D., & de Carvalho, L. R. (2016). Logical Inconsistencies in 3 Preference Elicitation Methods for EQ-5D Health States: A Study in the Brazilian Population. *Medical Decision Making, 36*(2), 242-252. doi:10.1177/0272989x15601047

Badia, X., Roset, M., & Herdman, M. (1999). Inconsistent responses in three preference-elicitation methods for health states. *Social Science & Medicine, 49*(7), 943-950. doi:10.1016/s0277-9536(99)00182-3

Bleichrodt, H., & Prades, J. L. P. (2009). New Evidence of Preference Reversals in Health Utility

Breidert, C., Hahsler, M., & Reutterer, T. (2006). A review of methods for measuring willingness-to-pay. *Innovative Marketing*.

Brown, J. R., Kapteyn, A., Luttmer, E. F. P., & Mitchell, O. S. (2017). Cognitive Constraints on Valuing Annuities. *Journal of the European Economic Association, 15*(2), 429-462. doi:10.1093/jeea/jvw009

Choi, S., Kariv, S., Muller, W., & Silverman, D. (2014). Who Is (More) Rational? *American Economic Review, 104*(6), 1518-1550. doi:10.1257/aer.104.6.1518

Craig, B. M., Pickard, A. S., & Rand-Hendriksen, K. (2015). Do health preferences contradict ordering of EQ-5D labels? *Quality of Life Research, 24*(7), 1759-1765. doi:10.1007/s11136-014-0897-z

Craig, B. M., & Ramachandran, S. (2006). Relative risk of a shuffled deck: a generalizable logical consistency criterion for sample selection in health state valuation studies. *Health Economics, 15*(8), 835-848. doi:10.1002/hec.1108

Craig, B. M., Reeve, B. B., Brown, P. M., Cella, D., Hays, R. D., Lipscomb, J., . . . Revicki, D. A. (2014). US Valuation of Health Outcomes Measured Using the PROMIS-29. *Value in Health, 17*(8), 846-853. doi:10.1016/j.jval.2014.09.005

Cramm, J. M., Hartgerink, J. M., Steyerberg, E. W., Bakker, T. J., Mackenbach, J. P., & Nieboer, A. P. (2013). Understanding older patients' self-management abilities: functional loss, self-management, and well-being. *Quality of Life Research, 22*(1), 85-92. doi:10.1007/s11136-012-0131-9

de Bekker-Grob, E. W., Ryan, M., & Gerard, K. (2012). Discrete choice experiments in health economics: a review of the literature. *Health Economics, 21*(2), 145-172. doi:10.1002/hec.1697

Devlin, N. J., Hansen, P., Kind, P., & Williams, A. (2003). Logical inconsistencies in survey respondents' health state valuations - a methodological challenge for estimating social tariffs. *Health Economics, 12*(7), 529-544. doi:10.1002/hec.741

Dolan, P., & Kind, P. (1996). Inconsistency and health state valuations. *Social Science & Medicine, 42*(4), 609-615. doi:10.1016/0277-9536(95)00161-1

Echenique, F., Lee, S., & Shum, M. (2011). The Money Pump as a Measure of Revealed Preference Violations. *Journal of Political Economy, 119*(6), 1201-1223. doi:10.1086/665011

Engel, L., Bansback, N., Bryan, S., Doyle-Waters, M. M., & Whitehurst, D. G. T. (2016). Exclusion Criteria in National Health State Valuation Studies: A Systematic Review. *Medical Decision Making, 36*(7), 798-810. doi:10.1177/0272989x15595365

Euroqol. (2018). About the EQ-5D-5L. Retrieved from https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/

Fong, K. N. K., Chan, C. C. H., & Au, D. K. S. (2001). Relationship of motor and cognitive abilities to functional performance in stroke rehabilitation. *Brain Injury, 15*(5), 443-453. doi:10.1080/02699050010005940

Fowler, F. J., Cleary, P. D., Massagli, M. P., Weissman, J., & Epstein, A. (1995). The Role of Reluctance to Give up Life in the Measurement of the Values of Health States. *Medical Decision Making, 15*(3), 195-200. doi:10.1177/0272989x9501500301

Greene, W. H. (2012). *Econometric Analysis* (Seventh Edition ed.). Boston: Pearson Education Limited.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., . . . Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20*(10), 1727-1736. doi:10.1007/s11136-011-9903-x

Johnson, F. R., & Mathews, K. E. (2001). Sources and effects of utility-theoretic inconsistency in stated-preference surveys. *American Journal of Agricultural Economics, 83*(5), 1328-1333. doi:10.1111/0002-9092.00286

Lamers, L. M., Stalmeier, P. F. M., Krabbe, P. F. M., & Busschbach, J. J. V. (2006). Inconsistencies in TTO and VAS values for EQ-5D health states. *Medical Decision Making, 26*(2), 173-181. doi:10.1177/0272989x06286480

Lancsar, E., & Louviere, J. (2006). Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics, 15*(8), 797-811. doi:10.1002/hec.1104

Levinthal, B. R., Morrow, D. G., Tu, W. Z., Wu, J. W., & Murray, M. D. (2008). Cognition and health literacy in patients with hypertension. *Journal of General Internal Medicine, 23*(8), 1172-1176. doi:10.1007/s11606-008-0612-2

McIntosh, E., & Ryan, M. (2002). Using discrete choice experiments to derive welfare estimates for the provision of elective surgery: Implications of discontinuous preferences. *Journal of Economic Psychology, 23*(3), 367-382. doi:10.1016/s0167-4870(02)00081-8

Mottus, R., Johnson, W., Murray, C., Wolf, M. S., Starr, J. M., & Deary, I. J. (2014). Towards Understanding the Links Between Health Literacy and Physical Health. *Health Psychology, 33*(2), 164-173. doi:10.1037/a0031439

NICE. (2017). Position Statement on use of the EQ-5D-5L valuation set [Press release]. Retrieved from https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf

Nord, E., Daniels, N., & Kamlet, M. (2009). QALYs: Some Challenges. *Value in Health, 12*, S10-S15. doi:10.1111/j.1524-4733.2009.00516.x

Oppe, M., Devlin, N. J., van Hout, B., Krabbe, P. F. M., & de Charro, F. (2014). A Program of Methodological Research to Arrive at the New International EQ-5D-5L Valuation Protocol. *Value in Health, 17*(4), 445-453. doi:10.1016/j.jval.2014.04.002

Ozdemir, S., Mohamed, A. F., Johnson, F. R., & Hauber, A. B. (2010). Who Pays Attention in Stated-Choice Surveys? *Health Economics, 19*(1), 111-118. doi:10.1002/hec.1452

Prince, D. W., & Rubin, P. H. (2002). The effects of product liability litigation on the value of firms. *American Law and Economics Review, 4*(1), 44-87.

R. Varian, H. (1994). *Goodness-of-Fit for Revealed Preference Tests*.

Rezaei, A., & Patterson, Z. (2015). *Detecting, Non-Transitive, Inconsistent Responses in Discrete Choice Experiments*. Retrieved from https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2015-30.pdf

Ryan, M., Watson, V., & Entwistle, V. (2009). Rationalizing the 'Irrational':A Think Aloud Study of Discrete Choice Experiment Responses. *Health Economics, 18*(3), 321-336. doi:10.1002/hec.1369

Saelensminde, K. (2002). The impact of choice inconsistencies in stated choice studies. *Environmental & Resource Economics, 23*(4), 403-420.

San Miguel, F., Ryan, M., & Amaya-Amaya, M. (2005). 'Irrational' stated preferences: a quantitative and qualitative investigation. *Health Economics, 14*(3), 307-322. doi:10.1002/hec.912

Schwappach, D. L. B., & Strasmann, T. J. (2006). "Quick and dirty numbers"? The reliability of a stated-preference technique for the measurement of preferences for resource allocation. *Journal of Health Economics, 25*(3), 432-448. doi:10.1016/j.jhealeco.2005.08.002

Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives, 13*(2), 145-166. doi:10.1257/jep.13.2.145

U.S. Census Bureau. (2015). Geographic Terms and Concepts- Census Divisions and Census Regions. Retrieved from https://www.census.gov/geo/reference/gtc/gtc_census_divreg.html

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*: MIT press.

Yang, Z. H., van Busschbach, J., Timman, R., Janssen, M. F., & Luo, N. (2017). Logical inconsistencies in time trade-off valuation of EQ-5D-5L health states: Whose fault is it? *Plos One, 12*(9), 10. doi:10.1371/journal.pone.0184883

Table 2.1: Demographics

	Participants % (#)	Failures of Transitivity Count Measure Mean (Std. Dev)	Failures of Transitivity Order Measure Mean (Std. Dev)
Gender			
Male	50.40% (1630)	2.62 (5.36)	0.94 (1.32)
Female	49.60% (1604)	2.77 (5.32)	1.05 (1.41)
Age, years			
18-29	14.41% (466)	3.51 (6.50)	1.20 (1.47)
30-39	19.91% (644)	3.16 (6.36)	1.08 (1.53)
40-49	16.26% (526)	2.29 (4.74)	0.88 (1.25)
50-59	21.30% (689)	2.70 (5.22)	1.01 (1.41)
60-69	18.74% (606)	2.22 (4.28)	0.88 (1.26)
70+	9.37% (303)	2.13 (3.76)	0.89 (1.11)
Race			
African American/ Black	11.01% (356)	3.85 (6.75)	1.29 (1.64)
Asian/ Asian American	2.44% (79)	3.81 (8.14)	1.09 (1.73)
Caucasian/ White	82.28% (2661)	2.44 (4.88)	.94 (1.30)
Native American/ Inuit/ Aleut	0.74% (24)	3.33 (4.69)	1.21 (1.18)
Native Hawaiian/ Pacific Islander	0.25% (8)	11.13 (10.66)	2.38 (1.85)
Other	3.28% (106)	3.72 (6.77)	1.21 (1.52)
Ethnicity			
Non-Hispanic	87.91% (2843)	2.64 (5.28)	0.98 (1.36)
Hispanic	12.09% (391)	3.11 (5.77)	1.08 (1.44)
Work Status			
Working	61.56% (1991)	2.66 (5.42)	0.98 (1.37)
Looking for Work	4.27% (138)	4.24 (6.61)	1.33 (1.49)
Not Working, Not Looking for Work	9.28% (300)	2.93 (5.53)	1.02 (1.41)
Retired	23.50% (760)	2.39 (4.77)	0.95 (1.33)
Missing/Refused to Answer	1.39% (45)	3.13 (4.48)	1.27 (1.34)
Education			
No Diploma	2.04% (66)	3.24 (5.57)	1.17 (1.63)
High School Diploma/Equivalent	46.60% (1507)	2.92 (5.72)	1.04 (1.42)
Some College	10.14% (328)	2.85 (5.79)	1.05 (1.39)
Associates Degree/Equivalent	5.13% (166)	3.06 (5.90)	1.11 (1.50)
Bachelor's Degree	20.07% (649)	2.31 (4.43)	0.92 (1.25)
Advanced Degree (Masters, Doctorate)	16.02% (518)	2.26 (4.67)	0.86 (1.25)
Health			
Problems with Mobility	27.40% (886)	2.35 (5.41)	0.86 (1.38)
Problems with Self-Care	4.82% (156)	4.31 (8.41)	1.36 (1.88)
Problems with Usual Activities	19.23% (622)	2.89 (5.93)	1.06 (1.49)
Problems with Pain/Discomfort	53.40% (1727)	2.62 (5.23)	0.98 (1.37)
Problems with Anxiety/Depression	34.38% (1112)	2.85 (5.53)	1.04 (1.43)
Intense Health Problems	5.41% (175)	3.47 (7.27)	1.14 (1.66)
Household Income			
Less than $15,000	4.61% (149)	4.31 (6.70)	1.47 (1.71)
$15,000-$24,999	6.46% (209)	3.29 (6.42)	1.06 (1.48)
$25,000-$34,999	8.19% (265)	3.59 (7.49)	1.23 (1.73)
$35,000-$44,999	8.07% (261)	3.34 (6.26)	1.11 (1.49)
$45,000-$49,999	6.18% (200)	2.56 (5.15)	0.98 (1.31)
$50,000-$74,999	19.20% (621)	2.61 (4.97)	0.97 (1.31)
$75,000-$99,999	15.77% (510)	2.58 (4.89)	1.01 (1.35)
$100,000-$149,999	15.65% (506)	2.12 (4.30)	0.81 (1.16)
$150,000-$249,999	6.71% (217)	1.62 (3.75)	0.70 (1.10)
$250,000-$499,999	1.30% (42)	1.69 (2.86)	0.79 (0.95)
$500,000 or More	0.22% (7)	1.29 (1.98)	0.57 (0.79)
Missing/Refused to Answer	7.64% (247)	2.49 (4.67)	0.98 (1.28)
Sample Average		2.70 (5.34)	0.99 (1.37)

Table 2.2: Income and Transitivity-Count Measure

	(1) Ordered Probit Base	(2) Interval Regression Base	(3) Working	(4) Married	(5) At Least 26 Years Old
Failures of Transitivity	-0.0140***	-579.7***	-785.9***	-644.2***	-475.8***
	(0.00350)	(132.6)	(175.7)	(221.2)	(137.3)
Education					
No Diploma	-0.102	-2,615	-13,024***	1,281	-5,748
	(0.149)	(4,826)	(4,992)	(7,195)	(4,815)
High School Diploma/Equivalent	Base	Base	Base	Base	Base
	-	-	-	-	-
Some College	0.264***	12,175***	11,380***	16,143***	12,591***
	(0.0728)	(3,146)	(3,317)	(5,024)	(3,307)
Associate's Degree/Equivalent	0.291***	11,269***	13,059***	17,797***	12,928***
	(0.0841)	(3,242)	(4,229)	(5,035)	(3,377)
Bachelor's Degree	0.733***	30,794***	31,949***	38,339***	32,750***
	(0.0549)	(2,668)	(3,090)	(3,755)	(2,700)
Graduate Degree (Masters, Doctorate)	0.969***	43,573***	46,408***	50,686***	43,979***
	(0.0593)	(3,175)	(3,635)	(4,287)	(3,035)
Current Health					
Problems with Mobility	-0.0956**	-1,737	-2,790	-549.2	-2,535
	(0.0471)	(2,016)	(2,793)	(3,359)	(2,046)
Problems with Self-Care	-0.00767	2,231	454.8	4,129	2,116
	(0.0749)	(2,926)	(5,715)	(6,364)	(3,135)
Problems with Usual Activities	-0.0843	-5,087**	-1,324	-8,404**	-3,659
	(0.0549)	(2,406)	(3,599)	(3,622)	(2,507)
Problems with Pain/Discomfort	-0.0559*	-643.0	-739.6	379.0	-867.0
	(0.0332)	(1,422)	(1,961)	(2,037)	(1,418)
Problems with Anxiety/Depression	-0.0897***	-3,873***	-2,703*	-3,768**	-3,476***
	(0.0280)	(1,069)	(1,434)	(1,705)	(1,109)
Work Status					
Currently Working	Base	Base		Base	Base
	-	-		-	-
Looking for Work	-0.800***	-23,254***		-41,018***	-28,721***
	(0.115)	(3,677)		(5,359)	(3,679)
Not Working, Not Looking for Work	-0.509***	-11,029***		-10,588*	-13,885***
	(0.0850)	(3,722)		(5,673)	(3,979)
Retired	-0.387***	-13,346***		-11,737**	-13,809***
	(0.0652)	(3,589)		(4,775)	(3,584)
Missing/Refused to Answer	-1.252***	-31,031***		-24,558**	-31,776***
	(0.245)	(6,164)		(9,921)	(7,017)
Times Individual Traded Lifespan and Health	0.112**	5,846**	7,791**	6,323**	4,886**
	(0.0517)	(2,311)	(3,038)	(3,145)	(2,350)
Times Individual Traded Lifespan and Health^2	-0.0164**	-862.2**	-1,195***	-964.7**	-757.8**
	(0.00764)	(347.2)	(462.3)	(465.1)	(354.4)
Times Individual Traded Lifespan and Health^3	0.000692**	36.71**	53.60***	41.69**	33.94**
	(0.000325)	(15.06)	(20.33)	(19.78)	(15.41)
Chances	0.000847	77.06	-169.8	537.9	197.5
	(0.00645)	(273.6)	(348.6)	(379.6)	(258.3)
Chances^2	-5.19e-06	-0.382	0.365	-1.910	-0.789
	(2.03e-05)	(0.867)	(1.112)	(1.208)	(0.807)
Observations	2,987	2,987	1,873	1,702	2,809

Note: Gender, age, race, ethnicity, marital status, and region of residence are not reported.
Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.3: Income and Transitivity-Order Measure

	(1) Ordered Probit Base	(2) Interval Regression Base	(3) Working	(4) Married	(5) At Least 26 Years Old
Failures of Transitivity	-0.0600***	-2,461***	-3,173***	-2,802***	-2,107***
	(0.0141)	(565.6)	(777.5)	(850.8)	(584.3)
Education					
No Diploma	-0.103	-2,684	-13,220***	1,923	-5,790
	(0.150)	(4,828)	(4,904)	(7,150)	(4,806)
High School Diploma/Equivalent	Base	Base	Base	Base	Base
	-	-	-	-	-
Some College	0.263***	12,114***	11,208***	16,165***	12,492***
	(0.0726)	(3,143)	(3,305)	(5,017)	(3,300)
Associate's Degree/Equivalent	0.296***	11,465***	13,679***	18,127***	13,102***
	(0.0844)	(3,262)	(4,323)	(5,108)	(3,397)
Bachelor's Degree	0.740***	31,100***	32,605***	38,657***	32,975***
	(0.0548)	(2,662)	(3,091)	(3,753)	(2,695)
Graduate Degree (Masters, Doctorate)	0.971***	43,731***	46,743***	50,922***	44,084***
	(0.0592)	(3,180)	(3,655)	(4,271)	(3,040)
Current Health					
Problems with Mobility	-0.0972**	-1,796	-2,912	-714.6	-2,633
	(0.0474)	(2,032)	(2,831)	(3,374)	(2,051)
Problems with Self-Care	-0.00866	2,139	525.6	4,038	2,115
	(0.0751)	(2,930)	(5,715)	(6,382)	(3,139)
Problems with Usual Activities	-0.0777	-4,781**	-741.2	-7,744**	-3,313
	(0.0549)	(2,409)	(3,600)	(3,599)	(2,506)
Problems with Pain/Discomfort	-0.0563*	-658.2	-778.1	107.7	-888.0
	(0.0332)	(1,421)	(1,963)	(2,034)	(1,416)
Problems with Anxiety/Depression	-0.0901***	-3,888***	-2,789*	-3,744**	-3,525***
	(0.0280)	(1,070)	(1,436)	(1,709)	(1,111)
Work Status					
Currently Working	Base	Base		Base	Base
	-	-		-	-
Looking for Work	-0.794***	-22,908***		-40,494***	-28,461***
	(0.115)	(3,697)		(5,414)	(3,695)
Not Working, Not Looking for Work	-0.511***	-11,119***		-10,582*	-13,957***
	(0.0850)	(3,724)		(5,696)	(3,979)
Retired	-0.381***	-13,061***		-11,213**	-13,551***
	(0.0651)	(3,588)		(4,772)	(3,584)
Missing/Refused to Answer	-1.256***	-31,301***		-24,351**	-31,873***
	(0.245)	(6,152)		(10,126)	(6,992)
Times Individual Traded Lifespan and Health	0.00870	356.7	-177.4	172.6	-192.2
	(0.0187)	(870.4)	(1,150)	(1,176)	(871.9)
Times Individual Traded Lifespan and Health²	-0.000188	-1.611	57.81	7.040	38.07
	(0.00121)	(57.30)	(78.16)	(73.57)	(57.64)
Chances	0.000622	61.57	-196.8	531.3	185.2
	(0.00644)	(272.6)	(346.8)	(380.0)	(257.4)
Chances ²	-4.75e-06	-0.343	0.430	-1.892	-0.757
	(2.03e-05)	(0.863)	(1.105)	(1.208)	(0.804)
Observations	2,987	2,987	1,873	1,702	2,809

Note: Gender, age, race, ethnicity, marital status, and region of residence are not reported.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.4: Income-Severity and Frequency of Transitivity Failures

	(1) Interval Base	(2) Interval Working	(3) Probit Base	(4) Probit Working
Failures of Transitivity (Order Measure)	-2,610***	-2,089**	-0.0733***	-0.0590**
	(775.3)	(1,059)	(0.0208)	(0.0273)
Average Severity of Failure	-799.2	-2,890***	-0.0100	-0.0575***
	(644.3)	(806.9)	(0.0164)	(0.0194)
Education				
No Diploma	-6,434	-11,777*	-0.243	-0.353*
	(4,729)	(6,913)	(0.168)	(0.205)
High School Diploma/Equivalent	Base	Base	Base	Base
	-	-	-	-
Some College	11,494***	10,080**	0.224**	0.289**
	(4,449)	(4,233)	(0.0972)	(0.113)
Associate's Degree/Equivalent	10,462**	11,280*	0.222*	0.276*
	(4,653)	(6,226)	(0.122)	(0.151)
Bachelor's Degree	25,881***	25,358***	0.650***	0.675***
	(3,401)	(4,216)	(0.0785)	(0.0988)
Graduate Degree (Masters, Doctorate)	41,640***	42,064***	0.915***	0.985***
	(4,966)	(5,508)	(0.0852)	(0.105)
Current Health				
Problems with Mobility	3,300	5,602	0.0589	0.110
	(2,312)	(3,611)	(0.0622)	(0.0786)
Problems with Self-Care	3,207	3,974	0.0170	0.0676
	(3,673)	(6,732)	(0.0920)	(0.151)
Problems with Usual Activities	-9,219***	-12,055**	-0.218***	-0.265**
	(2,720)	(4,884)	(0.0707)	(0.111)
Problems with Pain/Discomfort	-682.1	-1,638	-0.0483	-0.0525
	(1,879)	(2,697)	(0.0448)	(0.0621)
Problems with Anxiety/Depression	-4,557***	-2,527	-0.112***	-0.0599
	(1,416)	(1,965)	(0.0385)	(0.0496)
Work Status				
Currently Working	Base		Base	
	-		-	
Looking for Work	-19,774***		-0.745***	
	(4,815)		(0.149)	
Not Working, Not Looking for Work	-10,937**		-0.525***	
	(4,739)		(0.108)	
Retired	-10,844**		-0.328***	
	(4,662)		(0.0873)	
Missing/Refused to Answer	-35,169***		-1.373***	
	(7,901)		(0.325)	
Times Individual Traded Lifespan and Health	376.9	375.7	0.00950	0.00222
	(1,259)	(1,638)	(0.0293)	(0.0379)
Times Individual Traded Lifespan and Health²	-1.372	21.41	0.000122	0.00116
	(74.58)	(99.58)	(0.00175)	(0.00229)
Chances	-51.18	-499.1	2.72e-05	-0.00824
	(414.3)	(544.5)	(0.00905)	(0.0111)
Chances ²	0.105	1.638	-1.91e-06	2.60e-05
	(1.365)	(1.812)	(2.88e-05)	(3.55e-05)
Observations	1,540	941	1,540	941

Note: Gender, age, race, ethnicity, marital status, work status, and region of residence are not reported.
Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.5: Health and Transitivity-Count Measure

	(1) Mobility	(2) Self-Care	(3) Usual Activities	(4) Pain or Discomfort	(5) Anxiety or Depression	(6) Multiple Problems	(7) Intense Problems
Failures of Transitivity	0.00372	0.0190***	0.00548	0.00122	0.00292	-0.00320	0.00791
	(0.00506)	(0.00643)	(0.00520)	(0.00466)	(0.00481)	(0.00448)	(0.00691)
Education							
No Diploma	0.0553	0.196	0.274*	0.0244	0.397***	0.300*	0.0626
	(0.162)	(0.212)	(0.144)	(0.162)	(0.149)	(0.161)	(0.223)
High School Diploma/Equivalent	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
Some College	-0.0730	-0.0864	0.0603	-0.112	0.104	0.000691	0.0159
	(0.0902)	(0.147)	(0.0901)	(0.0749)	(0.0746)	(0.0811)	(0.132)
Associate's Degree/Equivalent	0.00151	-0.0464	-0.0732	-0.0278	0.0163	0.0150	-0.0864
	(0.110)	(0.158)	(0.121)	(0.0919)	(0.101)	(0.107)	(0.190)
Bachelor's Degree	-0.133*	-0.157	-0.202**	-0.196***	-0.0802	-0.171**	-0.317**
	(0.0714)	(0.130)	(0.0785)	(0.0561)	(0.0625)	(0.0673)	(0.130)
Graduate Degree (Masters, Doctorate)	-0.104	-0.138	-0.0561	-0.188***	0.0297	-0.182**	-0.290**
	(0.0764)	(0.139)	(0.0799)	(0.0626)	(0.0673)	(0.0739)	(0.147)
Household Income							
Less than $25,000	0.313***	0.281***	0.342***	0.281***	0.178**	0.323***	0.277**
	(0.0854)	(0.108)	(0.0873)	(0.0779)	(0.0801)	(0.0851)	(0.115)
$25,000-$49,999	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
$50,000-$74,999	-0.222***	-0.246*	-0.0671	-0.117*	-0.0218	-0.0639	-0.0750
	(0.0759)	(0.129)	(0.0794)	(0.0640)	(0.0676)	(0.0725)	(0.122)
$75,000-$99,999	-0.293***	-0.314**	-0.227***	-0.266***	-0.152**	-0.189**	-0.358**
	(0.0820)	(0.143)	(0.0879)	(0.0699)	(0.0741)	(0.0794)	(0.156)
$100,000-$149,999	-0.222***	-0.324**	-0.168*	-0.173**	-0.178**	-0.151*	-0.191
	(0.0849)	(0.150)	(0.0902)	(0.0718)	(0.0770)	(0.0819)	(0.149)
$150,000+	-0.193*	-0.206	-0.372***	-0.172**	-0.312***	-0.252**	-0.328
	(0.109)	(0.220)	(0.129)	(0.0857)	(0.0967)	(0.104)	(0.235)
Missing/Refused to Answer	-0.0717	-0.0344	-0.0884	-0.114	-0.162*	-0.141	-0.178
	(0.107)	(0.160)	(0.109)	(0.0908)	(0.0953)	(0.0981)	(0.164)
Times Individual Traded Lifespan and	0.0916	-0.0722	0.111	-0.00297	0.0155	0.0134	0.0248
Health	(0.0674)	(0.108)	(0.0701)	(0.0570)	(0.0598)	(0.0638)	(0.105)
Times Individual Traded Lifespan and	-0.0138	0.00696	-0.0165	0.00112	-0.000718	-0.000152	-0.00694
Health^2	(0.00983)	(0.0159)	(0.0102)	(0.00830)	(0.00885)	(0.00932)	(0.0158)
Times Individual Traded Lifespan and	0.000583	-0.000223	0.000674	-5.98e-05	-2.54e-05	-3.87e-05	0.000316
Health^3	(0.000412)	(0.000668)	(0.000426)	(0.000348)	(0.000375)	(0.000391)	(0.000677)
Chances	0.00950	0.0152	0.0100	0.0116	0.00541	0.00501	0.0361**
	(0.00880)	(0.0146)	(0.00983)	(0.00741)	(0.00796)	(0.00839)	(0.0171)
Chances ^{2}	-3.15e-05	-4.38e-05	-3.22e-05	-3.59e-05	-1.99e-05	-1.35e-05	-0.000116**
	(2.77e-05)	(4.57e-05)	(3.11e-05)	(2.35e-05)	(2.51e-05)	(2.65e-05)	(5.44e-05)
Observations	3,234	3,234	3,234	3,234	3,234	3,234	3,226

Note: Gender, age, race, ethnicity, marital status, work status, and region of residence are not reported.
Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.6: Health and Transitivity-Order Measure

	(1) Mobility	(2) Self-Care	(3) Usual Activities	(4) Pain or Discomfort	(5) Anxiety or Depression	(6) Multiple Problems	(7) Intense Problems
Failures of Transitivity	0.0137	0.0844***	0.0336*	0.0102	0.0132	0.00355	0.0293
	(0.0189)	(0.0268)	(0.0199)	(0.0167)	(0.0176)	(0.0175)	(0.0288)
Education							
No Diploma	0.0491	0.185	0.265*	0.0237	0.396***	0.300*	0.0584
	(0.161)	(0.212)	(0.143)	(0.162)	(0.149)	(0.161)	(0.223)
High School Diploma/Equivalent	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
Some College	-0.0763	-0.0910	0.0560	-0.111	0.104	0.00200	0.0118
	(0.0901)	(0.146)	(0.0900)	(0.0749)	(0.0746)	(0.0811)	(0.132)
Associate's Degree/Equivalent	0.00137	-0.0425	-0.0740	-0.0283	0.0159	0.0145	-0.0855
	(0.110)	(0.158)	(0.121)	(0.0918)	(0.101)	(0.106)	(0.190)
Bachelor's Degree	-0.134*	-0.162	-0.203***	-0.196***	-0.0810	-0.168**	-0.320**
	(0.0712)	(0.130)	(0.0782)	(0.0561)	(0.0624)	(0.0673)	(0.129)
Graduate Degree (Masters, Doctorate)	-0.102	-0.139	-0.0515	-0.188***	0.0297	-0.181**	-0.287**
	(0.0762)	(0.140)	(0.0798)	(0.0626)	(0.0672)	(0.0738)	(0.147)
Household Income							
Less than $25,000	0.310***	0.283***	0.340***	0.282***	0.179**	0.322***	0.276**
	(0.0855)	(0.108)	(0.0875)	(0.0779)	(0.0801)	(0.0851)	(0.115)
$25,000-$49,999	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
$50,000-$74,999	-0.223***	-0.242*	-0.0658	-0.116*	-0.0216	-0.0620	-0.0751
	(0.0759)	(0.130)	(0.0794)	(0.0641)	(0.0676)	(0.0725)	(0.121)
$75,000-$99,999	-0.296***	-0.314**	-0.229***	-0.265***	-0.152**	-0.188**	-0.360**
	(0.0821)	(0.143)	(0.0880)	(0.0700)	(0.0740)	(0.0794)	(0.156)
$100,000-$149,999	-0.219**	-0.317**	-0.161*	-0.172**	-0.177**	-0.148*	-0.188
	(0.0849)	(0.150)	(0.0901)	(0.0717)	(0.0771)	(0.0819)	(0.148)
$150,000+	-0.190*	-0.204	-0.365***	-0.170**	-0.311***	-0.247**	-0.330
	(0.109)	(0.220)	(0.129)	(0.0857)	(0.0967)	(0.104)	(0.235)
Missing/Refused to Answer	-0.0775	-0.0302	-0.0951	-0.113	-0.161*	-0.139	-0.183
	(0.107)	(0.161)	(0.109)	(0.0910)	(0.0953)	(0.0980)	(0.164)
Times Individual Traded Lifespan and Health	0.00193	-0.0449	0.00649	0.00553	0.0188	0.0186	-0.0240
	(0.0235)	(0.0376)	(0.0246)	(0.0202)	(0.0213)	(0.0227)	(0.0383)
Times Individual Traded Lifespan and Health²	-8.37e-05	0.00198	-0.000598	-0.000281	-0.00130	-0.00108	0.000503
	(0.00152)	(0.00238)	(0.00159)	(0.00128)	(0.00137)	(0.00144)	(0.00247)
Chances	0.00915	0.0149	0.00945	0.0116	0.00538	0.00505	0.0358**
	(0.00874)	(0.0146)	(0.00975)	(0.00741)	(0.00795)	(0.00839)	(0.0170)
Chances ²	-3.01e-05	-4.21e-05	-3.00e-05	-3.58e-05	-1.97e-05	-1.36e-05	-0.000115**
	(2.76e-05)	(4.57e-05)	(3.09e-05)	(2.35e-05)	(2.51e-05)	(2.65e-05)	(5.40e-05)
Observations	3,234	3,234	3,234	3,234	3,234	3,234	3,226

Note: Gender, age, race, ethnicity, marital status, work status, and region of residence are not reported.
Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.7: Health- Severity and Frequency of Transitivity Failures

	(1) Self-Care	(2) Usual Activities	(3) Intense Problems
Failures of Transitivity (Order Measure)	0.102***	0.0596**	0.0413
	(0.0381)	(0.0301)	(0.0428)
Average Severity of Failure	-0.0454	-0.0338	-0.0192
	(0.0303)	(0.0229)	(0.0298)
Education			
No Diploma	0.352	0.422**	0.0289
	(0.260)	(0.187)	(0.309)
High School Diploma/Equivalent	Base	Base	Base
	-	-	-
Some College	0.0538	0.0238	0.0462
	(0.187)	(0.125)	(0.173)
Associate's Degree/Equivalent	0.319*	0.0164	0.0579
	(0.184)	(0.169)	(0.238)
Bachelor's Degree	-0.00940	-0.107	-0.394**
	(0.176)	(0.105)	(0.191)
Graduate Degree (Masters, Doctorate)	-0.0196	-0.144	-0.613**
	(0.189)	(0.115)	(0.239)
Household Income			
Less than $25,000	0.219	0.205*	0.118
	(0.138)	(0.117)	(0.155)
$25,000-$49,999	Base	Base	Base
	-	-	-
$50,000-$74,999	-0.291	-0.268**	-0.113
	(0.186)	(0.111)	(0.168)
$75,000-$99,999	-0.302	-0.376***	-0.443**
	(0.187)	(0.125)	(0.224)
$100,000-$149,999	-0.487**	-0.342***	-0.416*
	(0.211)	(0.122)	(0.220)
$150,000+	-0.185	-0.424**	-0.110
	(0.299)	(0.175)	(0.308)
Missing/Refused to Answer	0.0234	-0.0954	-0.0548
	(0.209)	(0.149)	(0.214)
Times Individual Traded Lifespan and Health	-0.0111	0.000970	-0.114**
	(0.0549)	(0.0370)	(0.0562)
Times Individual Traded Lifespan and Health²	-0.000153	-0.000426	0.00573*
	(0.00330)	(0.00222)	(0.00340)
Chances	0.0254	0.0236	0.0621**
	(0.0210)	(0.0148)	(0.0258)
Chances ²	-8.31e-05	-7.88e-05*	-0.000204**
	(6.58e-05)	(4.70e-05)	(8.34e-05)
Observations	1,672	1,672	1,648

Note: Gender, age, race, ethnicity, marital status, work status, and region of residence are not reported.
Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8: Robustness Checks-Income

	(1)	(2)	(3)	(4)
	Moderate Comparisons		Learning Curves	
	Base	**Working**	**Base**	**Working**
Failures of Transitivity	-673.3***	-877.5***	-692.2***	-890.6***
	(176.8)	(222.2)	(165.7)	(216.7)
Education				
No Diploma	-2,679	-13,343***	-2,506	-12,690**
	(4,832)	(4,989)	(4,810)	(4,968)
High School Diploma/Equivalent	Base	Base	Base	Base
	-	-	-	-
Some College	12,300***	11,539***	12,103***	11,322***
	(3,148)	(3,316)	(3,145)	(3,297)
Associate's Degree/Equivalent	11,198***	12,896***	11,356***	13,055***
	(3,253)	(4,252)	(3,243)	(4,230)
Bachelor's Degree	30,894***	32,121***	30,807***	31,823***
	(2,674)	(3,102)	(2,678)	(3,080)
Graduate Degree (Masters, Doctorate)	43,645***	46,497***	43,696***	46,483***
	(3,169)	(3,631)	(3,201)	(3,644)
Current Health				
Problems with Mobility	-1,714	-2,819	-1,612	-2,632
	(2,019)	(2,789)	(2,029)	(2,830)
Problems with Self-Care	2,255	463.9	2,070	433.4
	(2,923)	(5,691)	(2,913)	(5,652)
Problems with Usual Activities	-5,138**	-1,395	-4,876**	-1,103
	(2,409)	(3,601)	(2,404)	(3,589)
Problems with Pain/Discomfort	-639.0	-599.6	-677.4	-942.4
	(1,423)	(1,956)	(1,424)	(1,973)
Problems with Anxiety/Depression	-3,917***	-2,736*	-3,922***	-2,915**
	(1,068)	(1,434)	(1,068)	(1,435)
Work Status				
Currently Working	Base		Base	
	-		-	
Looking for Work	-23,326***		-23,008***	
	(3,685)		(3,683)	
Not Working, Not Looking for Work	-10,956***		-11,016***	
	(3,720)		(3,724)	
Retired	-13,515***		-13,161***	
	(3,589)		(3,589)	
Missing/Refused to Answer	-31,294***		-31,324***	
	(6,187)		(6,185)	
Times Individual Traded Lifespan and Health	5,681**	7,574**	3,441	7,010***
	(2,303)	(3,033)	(2,373)	(2,632)
Times Individual Traded Lifespan and Health²	-842.1**	-1,172**	-569.6	-1,206***
	(346.3)	(461.6)	(406.3)	(467.4)
Times Individual Traded Lifespan and Health³	35.86**	52.69***	27.39	60.58***
	(15.02)	(20.30)	(19.77)	(23.27)
Chances	17.18	-210.4	208.6	41.95
	(215.1)	(277.3)	(321.3)	(429.0)
Chances ²	-0.396	0.718	-0.958	-0.305
	(1.139)	(1.467)	(1.244)	(1.677)
Observations	2,987	1,873	2,987	1,873

Note: Gender, age, race, ethnicity, marital status, and region of residence are not reported.
Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.9: Robustness Checks for Self-Care

	(1) Moderate	(2) Learning Curves
Failures of Transitivity	0.0263***	0.0270***
	(0.00831)	(0.00812)
Education		
No Diploma	0.199	0.195
	(0.212)	(0.212)
High School Diploma/Equivalent	Base	Base
	-	-
Some College	-0.0943	-0.0852
	(0.147)	(0.145)
Associate's Degree/Equivalent	-0.0456	-0.0574
	(0.156)	(0.158)
Bachelor's Degree	-0.156	-0.151
	(0.131)	(0.132)
Graduate Degree (Masters, Doctorate)	-0.138	-0.135
	(0.139)	(0.139)
Household Income		
s than $25,000	0.280***	0.277**
	(0.108)	(0.108)
$25,000-$49,999	Base	Base
	-	-
$50,000-$74,999	-0.246*	-0.250*
	(0.130)	(0.130)
$75,000-$99,999	-0.314**	-0.321**
	(0.144)	(0.144)
$100,000-$149,999	-0.327**	-0.318**
	(0.151)	(0.151)
$150,000+	-0.202	-0.204
	(0.220)	(0.220)
Missing/Refused to Answer	-0.0353	-0.0292
	(0.159)	(0.159)
Times Individual Traded Lifespan and Health	-0.0717	-0.0328
	(0.107)	(0.104)
Times Individual Traded Lifespan and Health^2	0.00702	0.00529
	(0.0158)	(0.0181)
Times Individual Traded Lifespan and Health^3	-0.000228	-0.000345
	(0.000665)	(0.000892)
Chances	0.0126	0.00739
	(0.0114)	(0.0164)
Chances ^{2}	-5.73e-05	-2.04e-05
	(5.92e-05)	(6.18e-05)
Observations	3,234	3,234

Note: Gender, age, race, ethnicity, marital status, and region of residence are not reported.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1: EQ-5D-5L (UK English sample version)

Under each heading, please tick the **ONE** box that best describes your health **TODAY**

MOBILITY	
I have no problems in walking about	❑
I have slight problems in walking about	❑
I have moderate problems in walking about	❑
I have severe problems in walking about	❑
I am unable to walk about	❑
SELF-CARE	
I have no problems washing or dressing myself	❑
I have slight problems washing or dressing myself	❑
I have moderate problems washing or dressing myself	❑
I have severe problems washing or dressing myself	❑
I am unable to wash or dress myself	❑
USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*	
I have no problems doing my usual activities	❑
I have slight problems doing my usual activities	❑
I have moderate problems doing my usual activities	❑
I have severe problems doing my usual activities	❑
I am unable to do my usual activities	❑
PAIN / DISCOMFORT	
I have no pain or discomfort	❑
I have slight pain or discomfort	❑
I have moderate pain or discomfort	❑
I have severe pain or discomfort	❑
I have extreme pain or discomfort	❑
ANXIETY / DEPRESSION	
I am not anxious or depressed	❑
I am slightly anxious or depressed	❑
I am moderately anxious or depressed	❑
I am severely anxious or depressed	❑
I am extremely anxious or depressed	❑

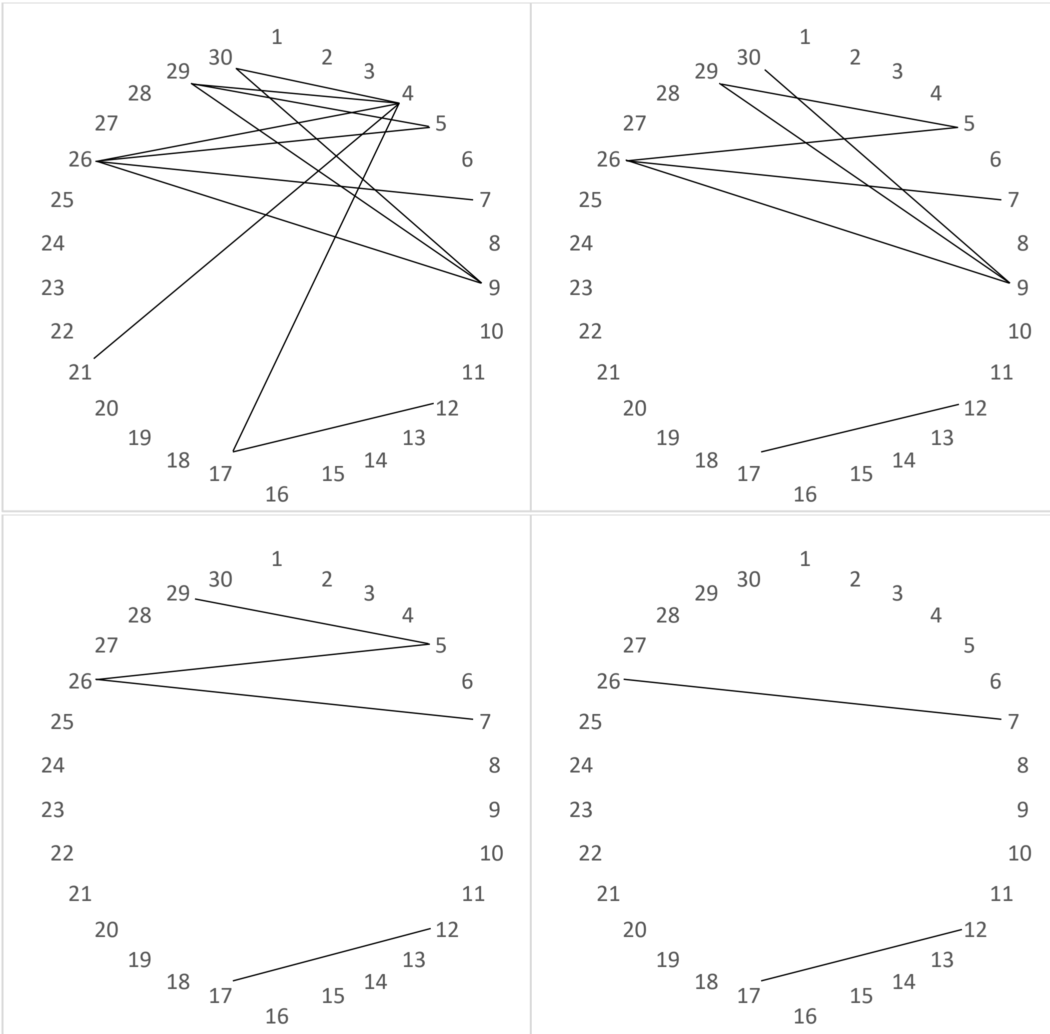
Figure 2.1: EQ-5D-5L Instrument

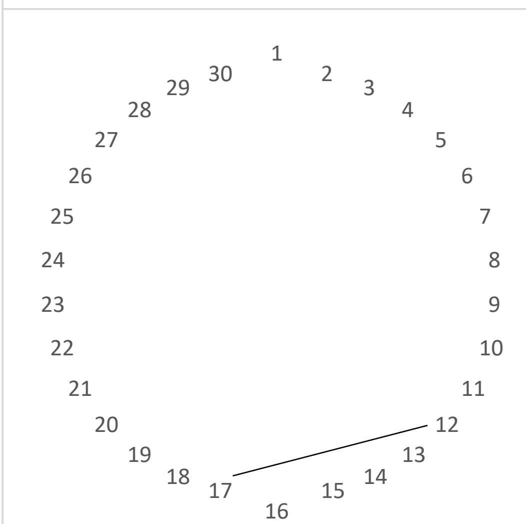
Figure 2.2: Defining a Failure of Transitivity in the TDL Study

For each question, *A* contains the option with a longer lifespan and health problems. *HS* is the health state, and *TA* is the amount of lag time with no health problems following time with health problems. *B* contains the option with a shorter lifespan with no health problems (*TB*).

All failures of transitivity require two necessary conditions:
1. Each domain of HS_2 must weakly dominate HS_1
2. The individual must also choose *A* for question 1 and *B* for question 2. In other words, an individual must prefer a more severe health state to time with no health problems in one question and then less time with no health problems to a less severe health state in another.

In addition, at least one of the following specifications must also be met:
3. $TA_2 \geq TA_1$ and $TB_1 \geq TB_2$
4. $TA_1 \leq TB_1$ or $TA_2 \geq TA_1$, and $TB_2 = 0$
5. $TB_1 - TA_1 \geq TB_2$





Using the Count Measure, the respondent has 12 failures of transitivity (each connected pair). Under the Order Measure, this is reduced to five failures of transitivity. In the Order Measure, I remove the question associated with the most failures until there are no remaining pairs (e.g., 4, 9, 5, 7, and 12). The average severity of each failure is 2.4 (12/5).

Figure 2.3: Comparing Transitivity Measures: Count vs. Order

Figure 2.4: Set of Pairs Where Dominance Cannot be Determined

One pair describes a health state with more pain/discomfort, while the other describes a health state with a higher level of anxiety/depression.

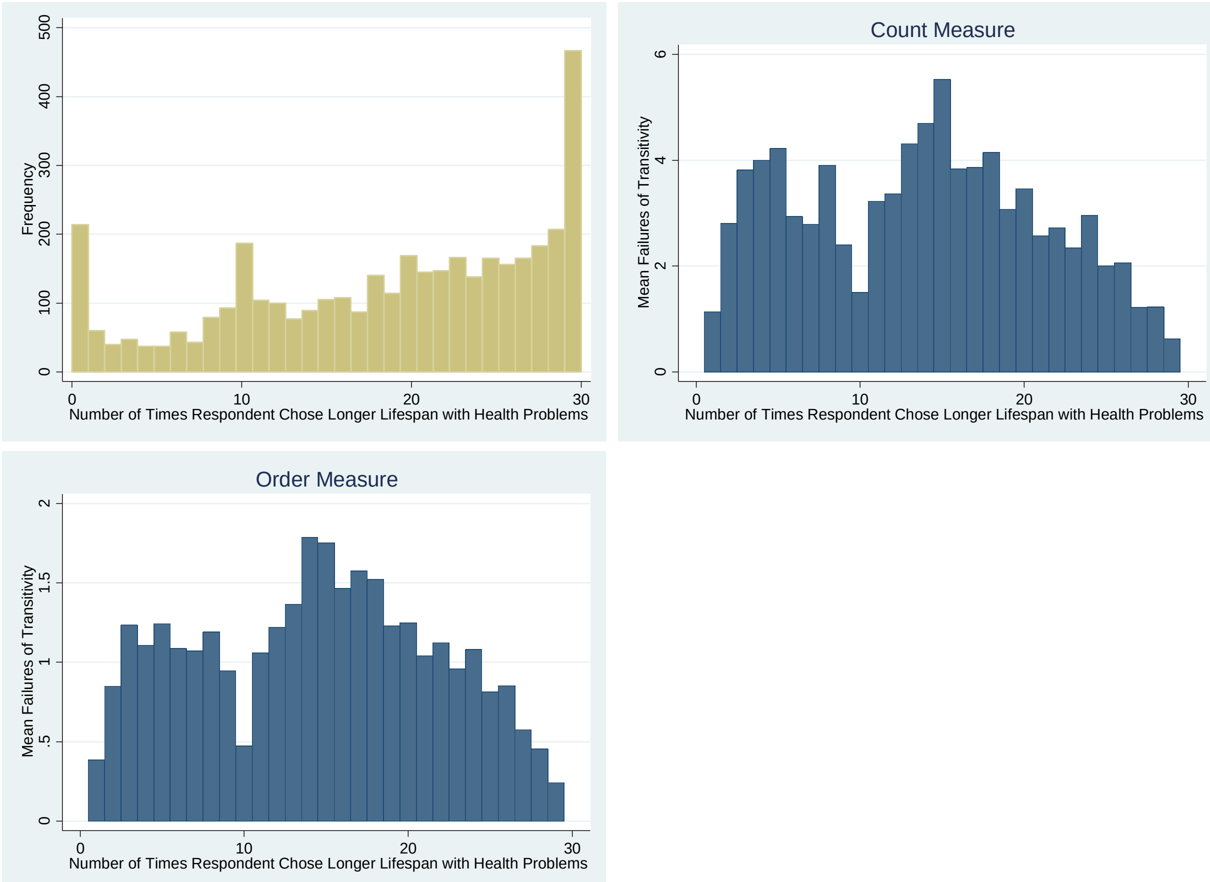
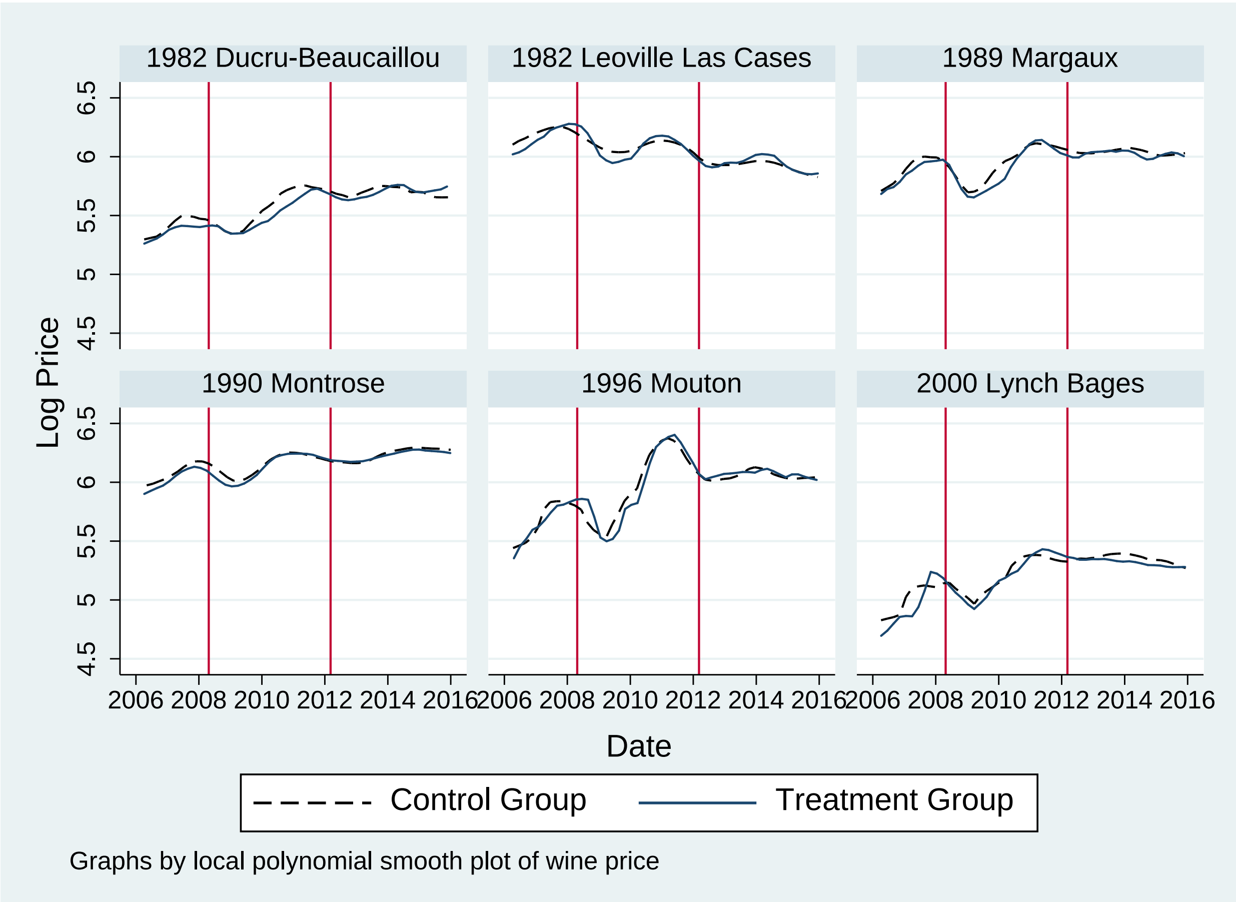


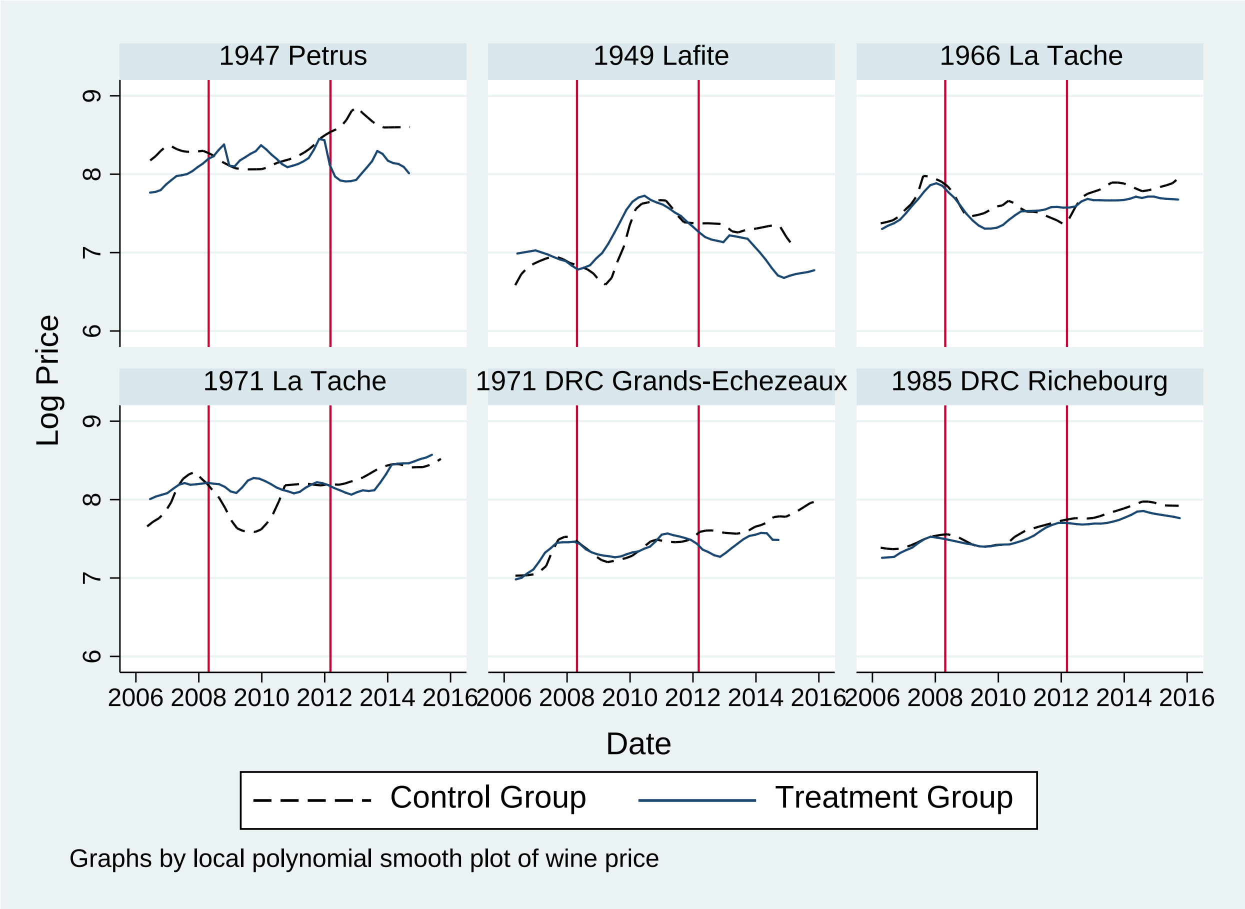
Figure 2.5: Willingness to Trade and Failures of Transitivity

Appendix A.1: Basic Difference-In-Difference Example

		Time		Δ Sales Price	Δ Volume
		Before	After		
Acker Merrall & Condit	Control	5.575	5.398	-0.177	
		(0.026)	(0.025)	(0.036)	
		[539]	[621]		15.21%
	Treated	7.505	7.486	-0.019	
		(0.023)	(0.025)	(0.034)	
		[1,183]	[1,235]		4.40%
	Difference			0.158	-10.82%
				(0.050)	
All Other Houses	Control	5.473	5.256	-0.217	
		(0.013)	(0.013)	(0.018)	
		[2,315]	[1,929]		-16.67%
	Treated	7.419	7.195	-0.224	
		(0.013)	(0.015)	(0.020)	
		[3,691]	[2,676]		-27.50%
	Difference			-0.007	-10.83%
				(0.027)	
Zachys/Sothebys	Control	5.511	5.225	-0.286	
		(0.016)	(0.017)	(0.023)	
		[1,523]	[1,215]		-20.22%
	Treated	7.449	7.190	-0.259	
		(0.016)	(0.018)	(0.024)	
		[2,501]	[1,648]		-34.11%
	Difference			0.027	-13.88%
				(0.033)	
DD Estimate	All			0.205	31.90%
				(0.039)	
	Zachys/ Sothebys			0.240	38.51%
				(0.042)	
DDD Estimate	All			0.165	0.01%
				(0.057)	
	Zachys/ Sothebys			0.131	3.07%
				(0.060)	



Appendix A.2: Wines Unlikely to Be Counterfeited



Appendix A.3: Wines Likely to Be Counterfeited

Appendix A.4: 2008 Difference-in-Difference Price Estimates- Exponential Regression

	(1) **Base DD**	(2) **North America**	(3) **Case and Large Format**	(4) **Very Expensive**
Acker	Base	Base	Base	Base
	-	-	-	-
Bonhams & Butterfields	-0.125***	-0.151***	-0.116***	-0.141***
	(0.0125)	(0.0132)	(0.0201)	(0.0234)
Hart Davis Hart	-0.0452***	-0.0491***	-0.0521***	-0.0732***
	(0.00906)	(0.00898)	(0.0120)	(0.0160)
Morrell & Company	-0.165***	-0.167***	-0.158***	-0.176***
	(0.0134)	(0.0133)	(0.0183)	(0.0359)
Sothebys	0.0156	0.0131	-0.0179	0.0232
	(0.0106)	(0.0108)	(0.0125)	(0.0189)
Zachys	-0.0644***	-0.0697***	-0.0894***	-0.0932***
	(0.00846)	(0.00822)	(0.0116)	(0.0140)
Time and House (DD)	**0.0834*** **	**0.0831*** **	**0.0537*** **	**0.0803*** **
	(0.0109)	**(0.0115)**	**(0.0166)**	**(0.0182)**
Asia	Base	Base	Base	Base
	-	-	-	-
Europe	-0.124***		-0.0873***	-0.160***
	(0.0156)		(0.0194)	(0.0287)
Internet	-0.227***		-0.326***	-0.269***
	(0.0169)		(0.0695)	(0.0398)
North America	-0.179***		-0.181***	-0.172***
	(0.0100)		(0.0136)	(0.0160)
Lot Size	0.00634***	0.00566***		0.0148***
	(0.00132)	(0.00136)		(0.00236)
Case Dummy	0.0327***	0.0441***		-0.0203
	(0.0109)	(0.0114)		(0.0196)
Observations	8,785	7,017	3,649	3,001
Number of Unique Wines	178	177	133	100

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

Appendix A.5: 2008 Triple Difference Price Estimates-Exponential Regression

	(1) Base DDD	(2) North America	(3) Burgundy	(4) Old Wines	(5) Semiannual Continent	(6) New Wines
Acker	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-
Bonhams & Butterfields	-0.109***	-0.125***	-0.211***	-0.185***	-0.0943***	-0.0724***
	(0.0124)	(0.0130)	(0.0391)	(0.0245)	(0.0125)	(0.0204)
Hart Davis Hart	-0.0270**	-0.0148	-0.155***	-0.0736***	-0.0290***	-0.0108
	(0.0109)	(0.0114)	(0.0366)	(0.0230)	(0.0108)	(0.0169)
Morrell & Company	-0.136***	-0.122***	-0.218***	-0.198***	-0.142***	-0.0845***
	(0.0130)	(0.0132)	(0.0503)	(0.0251)	(0.0128)	(0.0240)
Sothebys	0.0400***	0.0550***	-0.0535	0.00353	0.0249**	0.0341*
	(0.0123)	(0.0131)	(0.0422)	(0.0250)	(0.0124)	(0.0196)
Zachys	-0.0534***	-0.0411***	-0.190***	-0.0754***	-0.0736***	-0.0692***
	(0.0106)	(0.0111)	(0.0345)	(0.0217)	(0.0105)	(0.0168)
Time and House	0.0563***	0.0390**	0.0190	0.0995***	0.0431***	0.0521**
	(0.0135)	(0.0167)	(0.0432)	(0.0281)	(0.0137)	(0.0214)
Time and Fake	-0.00497	-0.00426	-0.0255	-0.00113	-0.00964**	0.00420
	(0.00431)	(0.00466)	(0.0155)	(0.00835)	(0.00452)	(0.00656)
Fake and House	0.00222	0.0275**	-0.171***	0.0106	-0.000914	-0.0361*
	(0.0126)	(0.0130)	(0.0373)	(0.0254)	(0.0126)	(0.0197)
Time, Fake, and House (DDD)	**0.0162**	**0.0456****	**0.136*****	**-0.0272**	**0.0146**	**0.0232**
	(0.0168)	**(0.0203)**	**(0.0479)**	**(0.0337)**	**(0.0172)**	**(0.0256)**
Asia	Base		Base	Base		Base
	-		-	-		-
Europe	-0.216***		-0.0812**	-0.254***		-0.212***
	(0.0120)		(0.0387)	(0.0266)		(0.0171)
Internet	-0.242***		-0.230***	-0.311***		-0.241***
	(0.0120)		(0.0279)	(0.0237)		(0.0179)
North America	-0.206***		-0.171***	-0.221***		-0.200***
	(0.00851)		(0.0160)	(0.0203)		(0.0117)
Lot Size	0.00219**	0.00135	0.00612**	0.00234	0.00185*	0.00578***
	(0.00107)	(0.00113)	(0.00249)	(0.00173)	(0.00109)	(0.00195)
Case Dummy	0.0389***	0.0545***	0.0112	0.0807***	0.0521***	-0.0284*
	(0.00823)	(0.00865)	(0.0240)	(0.0143)	(0.00839)	(0.0152)
Observations	14,189	10,837	2,390	5,555	14,189	4,679
Number of Unique Wines	286	282	106	141	286	86

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

Appendix A.6: 2012 Difference-In-Difference Price Estimates- Exponential Regression

	(1) **Base DD**	(2) **North America**	(3) **Case and Large Format**	(4) **Very Expensive**
Acker	Base	Base	Base	Base
	-	-	-	-
Bonhams & Butterfields	-0.131***	-0.196***	-0.0582***	-0.0891***
	(0.0105)	(0.0131)	(0.0154)	(0.0218)
Christies	0.0461***	-0.0417***	0.0599***	0.0896***
	(0.00829)	(0.00916)	(0.00876)	(0.0173)
Hart Davis Hart	-0.0735***	-0.103***	-0.0823***	-0.0483***
	(0.00679)	(0.00813)	(0.0102)	(0.0137)
Heritage Auctions	-0.128***	-0.150***	-0.121***	-0.104***
	(0.00936)	(0.0104)	(0.0170)	(0.0182)
K & L Wines	0.00914		-0.0861***	0.0617**
	(0.0101)		(0.0308)	(0.0286)
Morrell & Company	-0.117***	-0.105***	-0.0573***	-0.126***
	(0.0136)	(0.0159)	(0.0207)	(0.0317)
Sothebys	-0.0179**	-0.0572***	0.000529	-0.00951
	(0.00729)	(0.0130)	(0.00806)	(0.0144)
Spectrum Wine Auctions	-0.136***		-0.0827***	-0.136***
	(0.00897)		(0.0133)	(0.0160)
Zachys	-0.0778***	-0.0925***	-0.0764***	-0.0885***
	(0.00618)	(0.00874)	(0.00799)	(0.0127)
Time and House (DD)	**-0.0556*** **	**-0.0767*** **	**-0.0805*** **	**-0.0386** **
	(0.00780)	**(0.0107)**	**(0.00996)**	**(0.0164)**
Asia	Base		Base	Base
	-		-	-
Europe	-0.119***		-0.0973***	-0.168***
	(0.00787)		(0.00785)	(0.0163)
Internet	-0.178***		-0.120***	-0.269***
	(0.00938)		(0.0220)	(0.0244)
North America	-0.0985***		-0.0974***	-0.124***
	(0.00495)		(0.00661)	(0.00985)
Lot Size	4.95e-06	0.000148		0.00737***
	(0.000927)	(0.00111)		(0.00209)
Case Dummy	0.0667***	0.0593***		0.0305
	(0.00794)	(0.00992)		(0.0189)
Observations	15,471	6,090	6,657	4,908
Number of Unique Wines	178	177	143	100

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

Appendix A.7: 2012 Triple Difference Price Estimates- Exponential Regression

	(1) Base DDD	(2) North America	(3) Old Wines	(4) Case and Large Format	(5) New Wines
Acker	Base	Base	Base	Base	Base
	-	-	-	-	-
Bonhams & Butterfields	-0.0997***	-0.172***	-0.154***	-0.0463***	-0.0458***
	(0.00898)	(0.0127)	(0.0185)	(0.0112)	(0.0144)
Christies	0.0381***	-0.0246***	0.0608***	0.0450***	0.0442***
	(0.00631)	(0.00746)	(0.0144)	(0.00650)	(0.00638)
Hart Davis Hart	-0.0129*	-0.0552***	-0.0343**	-0.000179	0.0106
	(0.00737)	(0.0106)	(0.0165)	(0.00940)	(0.00953)
Heritage Auctions	-0.0981***	-0.130***	-0.115***	-0.0845***	-0.119***
	(0.00888)	(0.0119)	(0.0181)	(0.0130)	(0.0124)
K & L Wines	0.0419***		0.0236	-0.0159	0.0371***
	(0.00726)		(0.0156)	(0.0241)	(0.00999)
Morrell & Company	-0.0991***	-0.110***	-0.173***	-0.0618***	-0.0418***
	(0.0105)	(0.0151)	(0.0215)	(0.0154)	(0.0139)
Sothebys	0.0289***	-0.0321**	0.00195	0.0494***	0.0361***
	(0.00754)	(0.0128)	(0.0165)	(0.00893)	(0.00984)
Spectrum Wine Auctions	-0.119***		-0.193***	-0.0915***	-0.0546***
	(0.00641)		(0.0114)	(0.0100)	(0.00951)
Zachys	-0.0371***	-0.0624***	-0.0738***	-0.0257***	-0.0272***
	(0.00658)	(0.0105)	(0.0145)	(0.00800)	(0.00859)
Time and House	-0.0169**	-0.0693***	0.0137	-0.0132	-0.0244**
	(0.00788)	(0.0143)	(0.0187)	(0.00999)	(0.00993)
Time and Fake	-0.0327***	-0.0314***	7.22e-05	-0.0361***	-0.0412***
	(0.00343)	(0.00468)	(0.00783)	(0.00463)	(0.00458)
Fake and House	0.0464***	0.0272**	0.0497***	0.0653***	0.0227**
	(0.00798)	(0.0124)	(0.0167)	(0.00980)	(0.00985)
Time, Fake, and House (DDD)	**-0.0428*** **	**-0.00400**	**-0.0998*** **	**-0.0802*** **	**-0.0198**
	(0.0113)	**(0.0177)**	**(0.0241)**	**(0.0143)**	**(0.0140)**
Asia	Base		Base	Base	Base
	-		-	-	-
Europe	-0.115***		-0.195***	-0.0946***	-0.0748***
	(0.00591)		(0.0136)	(0.00582)	(0.00688)
Internet	-0.162***		-0.238***	-0.122***	-0.107***
	(0.00634)		(0.0120)	(0.0149)	(0.00876)
North America	-0.103***		-0.150***	-0.0964***	-0.0832***
	(0.00396)		(0.00808)	(0.00474)	(0.00518)
Lot Size	-0.00230***	-0.00183**	-0.00519***		-0.000981
	(0.000589)	(0.000906)	(0.00137)		(0.000662)
Case Dummy	0.0674***	0.0645***	0.120***		0.0342***
	(0.00525)	(0.00743)	(0.0124)		(0.00631)
Observations	25,078	9,686	9,031	11,697	8,688
Number of Unique Wines	286	284	141	249	86

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

Appendix A.8: List of Wines Included in the Dataset

	1929	1945	1947	1949	1952	1953	1955	1959	1961	1962	1964	1966	1969	1970	1971	1976	1978	1982	1985	1986	1988	1989	1990	1993	1995	1996	1999	2000	2002
Lafite		X		X		X	X	X	X			N		N				X		X		X	X			X		X	
Latour	X	X		X	X	X	X	X	X			N		N				X		X			M			M		X	
Mouton		X	X				X	X	X									X		M		N	X			N		X	
Margaux		X				X		X	X					N				X		N		N	X			M		X	
Haut Brion	X	X		X			X	X	X			N		N				M				X	X			N		M	
DRC La Tache								X		X		X			X	X	X		X		X		X		X	X	X		X
Ponsot Clos de la Roche																							X		N	N	N		N
DRC Romanee Conti										X	X	X			X	X	X		X		X		X		X	X	X		X
DRC Romanee St. Vivant																X	X		X				X		X	X	X		X
DRC Richebourg								X				X	X		X	X	X		X				X		X	X	X		X
Petrus		X	X					X	X					X				X				X	X		X			X	
Cheval Blanc			X				X	X	X		X	N		N				X				N	X		X			X	
Le Pin																		X	X	X			X						
Dujac Clos de la Roche																			X				X			N			N
Henri Jayer Echezeaux																	X		X				X						
DRC Echezeaux																			X		X		X		X	X	X		X
DRC Grands-Echezeaux															X		X		X		X		X		X	X	X		X
Ducru-Beaucaillou									N					N				N		N			N			N		N	
Calon Segur									N									N								N		N	
Leoville Barton																		N		N			N			N		N	
Leoville Poyferre																		N					N			N		N	
Lynch Bages									N					N				N				N	N			N		N	
Talbot																		N		N			N			N		N	
La Lagune																		N		N									
Beychevelle									N					N				N		N			N			N		N	
Les Forts de Latour														N				N					N			N		N	
Cos d'Estournel														N				N		N			N			N		N	
Mommessin Clos de Tart																			N				N			N	N		N
De Vogue Bonnes Mares																							N	N		N	N		N
Louis Latour Corton-Charlemagne																							N			N	N		
Bonneau Corton-Charlemagne																										N	N		N
Lafleur																		X					X			N		X	
Latour a Pomerol									X									N											
Trotanoy									X			N						N											
La Mission Haut Brion		X						X	X			N		N				X				X	N			N		X	
Leoville Las Cases									N			N		N				N					N			N		N	
Palmer		X						X	X					N				N					N					N	
Montrose									N									N					N			N		N	
Henri Jayer Richebourg																	X		X										
Henri Jayer Cros Parantoux																	X		X		X		X		X	X	X		
De Vogue Musigny								X		X	X	N			X		N										N		

Note: The X denotes that the wine is likely to be counterfeited, N denotes a wine is unlikely to be counterfeited, and M denotes wines excluded from the regression models.

Appendix B.1: Example set of 30 Paired Comparison Questions

Health states from the EQ-5D-5L are often described using a five-number nomenclature, where each number corresponds to the level of difficulties (no problems, slight problems, moderate problems, severe problems, and extreme problems) on each of the five domains. For example, 21354 corresponds to slight problems with mobility, no problems with self-care, moderate problems with usual activities, extreme pain or discomfort, and severe anxiety or depression. For the 30 questions below, the health problems are expected to last for 30 days, with the lag time following the health problems and alternate choice of time with no health problems also represented using days. Specifically, question 1 asks respondents to state their preferences between:

Option A) 30 days in health state 44424 followed by 33 days with no health problems
Option B) 1 day with no health problems

1. 44424 + 33 vs. 1	**16**. 33323 + 25 vs. 3
2. 44444 + 0 vs. 1	**17**. 33323 + 60 vs. 30
3. 44443 + 13 vs. 3	**18**. 33313 + 25 vs. 1
4. 44434 + 60 vs. 3	**19**. 33333 + 33 vs. 30
5. 44444 + 25 vs. 5	**20**. 33331 + 33 vs. 0
6. 44443 + 0 vs. 5	**21**. 55553 + 60 vs. 0
7. 44442 + 33 vs. 14	**22**. 55554 + 33 vs. 0
8. 44424 + 25 vs. 14	**23**. 55555 + 60 vs. 30
9. 44442 + 0 vs. 14	**24**. 55555 + 25 vs. 14
10. 44434 + 60 vs. 30	**25**. 55545 + 33 vs. 30
11. 33333 + 13 vs. 1	**26**. 55554 + 25 vs. 1
12. 33332 + 60 vs. 14	**27**. 55535 + 25 vs. 14
13. 33313 + 25 vs. 5	**28**. 55535 + 33 vs. 0
14. 33332 + 0 vs. 5	**29**. 55553 + 0 vs. 1
15. 33331 + 13 vs. 3	**30**. 55545 + 25 vs. 5

Appendix B.2: Negative Binomial Regressions Using Count Measure of Transitivity as the Dependent Variable

	(1) Base	(2) Drops Non-Traders	(3) Health Utility Value	(4) Intense Health Problems	(5) Years of Education	(6) Married	(7) Currently Working
Gender							
Female	-0.0903	-0.0568	-0.0617	-0.0579	-0.0554	-0.0111	0.0214
	(0.0690)	(0.0661)	(0.0657)	(0.0658)	(0.0663)	(0.0871)	(0.0870)
Age							
18-29	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
30-39	-0.00617	-0.0134	-0.0190	-0.0235	-0.00360	-0.101	0.00924
	(0.112)	(0.112)	(0.112)	(0.113)	(0.111)	(0.178)	(0.132)
40-49	-0.168	-0.229*	-0.250*	-0.246*	-0.216*	-0.445**	-0.117
	(0.139)	(0.132)	(0.131)	(0.131)	(0.131)	(0.188)	(0.155)
50-59	-0.125	-0.122	-0.148	-0.140	-0.108	-0.287	-0.0426
	(0.127)	(0.122)	(0.121)	(0.122)	(0.122)	(0.179)	(0.149)
60-69	-0.202	-0.219	-0.227	-0.238*	-0.206	-0.355*	0.00138
	(0.155)	(0.147)	(0.143)	(0.144)	(0.146)	(0.197)	(0.195)
70+	-0.292	-0.303*	-0.317*	-0.334*	-0.290*	-0.354	-0.0990
	(0.187)	(0.176)	(0.170)	(0.172)	(0.176)	(0.227)	(0.413)
Race							
African American/ Black	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
Asian/ Asian American	-0.130	-0.158	-0.151	-0.148	-0.144	-0.130	-0.234
	(0.238)	(0.233)	(0.234)	(0.234)	(0.234)	(0.312)	(0.314)
Caucasian/ White	-0.432***	-0.497***	-0.513***	-0.503***	-0.482***	-0.727***	-0.332***
	(0.107)	(0.103)	(0.102)	(0.102)	(0.102)	(0.160)	(0.126)
Native American/ Inuit/ Aleut	-0.0773	-0.169	-0.169	-0.174	-0.169	-0.0417	0.392
	(0.387)	(0.335)	(0.339)	(0.335)	(0.330)	(0.431)	(0.455)
Native Hawaiian/ Pacific Islander	1.153***	1.275***	1.299***	1.298***	1.299***	1.004*	1.398***
	(0.446)	(0.384)	(0.387)	(0.388)	(0.387)	(0.513)	(0.406)
Other	-0.0817	-0.206	-0.188	-0.159	-0.201	-0.527*	0.128
	(0.264)	(0.238)	(0.237)	(0.237)	(0.236)	(0.294)	(0.286)
Ethnicity							
Hispanic	0.142	0.182*	0.186*	0.179*	0.172*	0.212	0.0440
	(0.118)	(0.104)	(0.105)	(0.106)	(0.104)	(0.149)	(0.138)
Marital Status							
Married	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
Widowed	-0.0658	0.0393	0.0451	0.0399	0.0447		0.387
	(0.194)	(0.206)	(0.205)	(0.205)	(0.206)		(0.402)
Divorced/Separated	0.0959	0.0454	0.0555	0.0582	0.0465		0.0364
	(0.130)	(0.120)	(0.121)	(0.122)	(0.120)		(0.165)
Never Married	0.0908	0.0632	0.0585	0.0578	0.0580		0.0956
	(0.0962)	(0.0939)	(0.0944)	(0.0944)	(0.0944)		(0.117)
Living with a Partner	-0.0487	-0.0588	-0.0761	-0.0665	-0.0596		-0.123
	(0.128)	(0.121)	(0.119)	(0.120)	(0.122)		(0.146)
Education							
No Diploma	-0.00128	-0.00671	-0.00620	0.000605		0.168	0.446
	(0.210)	(0.205)	(0.204)	(0.202)		(0.281)	(0.272)
High School Diploma/Equivalent	Base	Base	Base	Base		Base	Base
	-	-	-	-		-	-
Some College	-0.115	-0.128	-0.124	-0.110		-0.123	0.0521
	(0.120)	(0.115)	(0.115)	(0.116)		(0.152)	(0.158)
Associate's Degree/Equivalent	-0.244*	-0.149	-0.123	-0.107		-0.210	-0.231
	(0.137)	(0.142)	(0.144)	(0.145)		(0.185)	(0.198)
Bachelor's Degree	-0.259**	-0.272***	-0.259***	-0.262***		-0.425***	-0.171
	(0.103)	(0.0948)	(0.0950)	(0.0950)		(0.120)	(0.117)
Graduate Degree (Masters, Doctorate)	-0.336***	-0.329***	-0.316***	-0.315***		-0.459***	-0.234*
	(0.104)	(0.0976)	(0.0973)	(0.0974)		(0.126)	(0.131)
Years of Education					-0.0485***		
					(0.0138)		
Current Health							
Problems with Mobility	-0.0808	-0.0766			-0.0770	-0.0699	0.0633
	(0.0725)	(0.0689)			(0.0689)	(0.0937)	(0.111)
Problems with Self-Care	0.270***	0.267***			0.267***	0.289**	0.335*
	(0.0903)	(0.0891)			(0.0894)	(0.145)	(0.179)

Appendix B.2 (Continued)							
Problems with Usual Activities	-0.0152	0.0201			0.0174	0.0201	-0.0250
	(0.0858)	(0.0813)			(0.0815)	(0.112)	(0.139)
Problems with Pain/Discomfort	0.00789	-0.0169			-0.0151	-0.0331	-0.191**
	(0.0598)	(0.0540)			(0.0544)	(0.0698)	(0.0776)
Problems with Anxiety/Depression	0.0339	0.0319			0.0325	-0.0754	-0.0625
	(0.0557)	(0.0479)			(0.0479)	(0.0615)	(0.0628)
Health Utility Value			-0.756**				
			(0.312)				
Intense Health Problems				0.192			
				(0.151)			
Household Income							
Less than $25,000	0.0432	0.0963	0.0956	0.102	0.0959	0.0882	0.0942
	(0.119)	(0.118)	(0.118)	(0.118)	(0.118)	(0.199)	(0.183)
$25,000-$49,999	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
$50,000-$74,999	-0.107	-0.0937	-0.0873	-0.0982	-0.103	0.115	-0.287**
	(0.107)	(0.103)	(0.102)	(0.102)	(0.103)	(0.130)	(0.141)
$75,000-$99,999	-0.106	-0.118	-0.104	-0.115	-0.122	0.0315	-0.253*
	(0.116)	(0.108)	(0.108)	(0.108)	(0.109)	(0.139)	(0.134)
$100,000-$149,999	-0.203*	-0.247**	-0.235**	-0.246**	-0.261**	-0.0625	-0.567***
	(0.119)	(0.111)	(0.111)	(0.111)	(0.111)	(0.137)	(0.148)
$150,000+	-0.499***	-0.519***	-0.520***	-0.529***	-0.526***	-0.246	-0.763***
	(0.163)	(0.149)	(0.147)	(0.148)	(0.150)	(0.187)	(0.185)
Missing/Refused to Answer	-0.183	-0.180	-0.180	-0.159	-0.189	0.00695	-0.404**
	(0.137)	(0.128)	(0.128)	(0.130)	(0.127)	(0.171)	(0.191)
Work Status							
Currently Working	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
With a Job, but Not at Work	-0.215	-0.258	-0.226	-0.222	-0.250	-0.127	
	(0.206)	(0.189)	(0.189)	(0.190)	(0.190)	(0.273)	
Looking for Work	0.286*	0.130	0.145	0.160	0.140	0.472**	
	(0.169)	(0.148)	(0.148)	(0.149)	(0.148)	(0.215)	
Working, but Not for Pay,	-0.491	-0.677*	-0.632*	-0.611*	-0.659*	-0.0820	
at a Family Business	(0.395)	(0.349)	(0.353)	(0.359)	(0.349)	(0.449)	
Not Working and Not Looking for Work	-0.119	-0.123	-0.146	-0.136	-0.113	-0.0960	
	(0.123)	(0.117)	(0.116)	(0.117)	(0.118)	(0.147)	
Retired	0.0179	-0.00897	-0.0424	-0.0210	-0.00213	-0.0155	
	(0.120)	(0.111)	(0.110)	(0.111)	(0.112)	(0.135)	
Refused to Answer	-0.132	0.0186	0.00964	-0.00371	0.0393	0.277	
	(0.268)	(0.280)	(0.288)	(0.290)	(0.280)	(0.446)	
Don't Know	-0.0171	0.0989	0.180	0.204	0.114	0.284	
	(0.259)	(0.272)	(0.323)	(0.325)	(0.269)	(0.428)	
Duration of Health Problems							
1 Day	Base	Base	Base	Base	Base	Base	Base
	-	-	-	-	-	-	-
7 Days	-0.196	-0.287**	-0.280**	-0.279**	-0.300**	-0.462***	-0.102
	(0.130)	(0.120)	(0.120)	(0.120)	(0.119)	(0.154)	(0.164)
30 Days	0.0325	-0.0130	-0.00479	0.00538	-0.0174	-0.100	0.0929
	(0.119)	(0.116)	(0.116)	(0.116)	(0.117)	(0.148)	(0.156)
9 Weeks	-0.524***	-0.542***	-0.531***	-0.531***	-0.550***	-0.689***	-0.538***
	(0.123)	(0.116)	(0.116)	(0.116)	(0.116)	(0.155)	(0.155)
12 Months	-0.127	-0.162	-0.163	-0.161	-0.170*	-0.236*	-0.122
	(0.106)	(0.103)	(0.103)	(0.103)	(0.103)	(0.132)	(0.136)
10 Years	-0.812***	-0.891***	-0.865***	-0.854***	-0.906***	-0.928***	-0.808***
	(0.179)	(0.170)	(0.169)	(0.169)	(0.170)	(0.224)	(0.238)
Times Individual Traded Lifespan and Health	0.535***	0.195***	0.197***	0.198***	0.195***	0.254***	0.253***
	(0.0273)	(0.0343)	(0.0343)	(0.0345)	(0.0344)	(0.0449)	(0.0443)
Times Individual Traded Lifespan and Health^2	-0.0251***	-0.00646***	-0.00659***	-0.00662***	-0.00644***	-0.00982***	-0.00980***
	(0.00173)	(0.00206)	(0.00205)	(0.00206)	(0.00206)	(0.00265)	(0.00266)
Number of Observations	3,909	3,234	3,234	3,234	3,234	1,845	1,867
Alpha	2.997	2.684	2.690	2.697	2.687	2.518	2.764

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

Appendix B.3: Willingness to Trade and Binary Transitivity Models

	(1) Logit Willingness To Trade	(2) NegBin II Willingness To Trade	(3) NegBin II Drops Non-Traders	(4) Logit At Least One Failure	(5) Logit Drops Non-traders
Education					
No Diploma	-0.149	-0.126	-0.0930	0.0822	0.0990
	(0.274)	(0.0991)	(0.0775)	(0.247)	(0.276)
High School Diploma/Equivalent	Base	Base	Base	Base	Base
	-	-	-	-	-
Some College	0.224	0.0168	-0.0210	0.108	0.0661
	(0.154)	(0.0433)	(0.0338)	(0.124)	(0.130)
Associate's Degree/Equivalent	-0.0569	0.00642	0.0224	-0.0619	-0.00995
	(0.181)	(0.0596)	(0.0439)	(0.165)	(0.177)
Bachelor's Degree	0.547***	0.0208	-0.0587**	0.0689	-0.0474
	(0.134)	(0.0339)	(0.0273)	(0.102)	(0.104)
Graduate Degree (Masters, Doctorate)	0.425***	0.0490	-0.0171	-0.131	-0.213*
	(0.147)	(0.0374)	(0.0297)	(0.115)	(0.115)
Current Health					
Problems with Mobility	0.0952	0.0253	0.0101	-0.0606	-0.0744
	(0.104)	(0.0268)	(0.0206)	(0.0792)	(0.0831)
Problems with Self-Care	-0.0275	-0.0483	-0.0414	0.340***	0.382***
	(0.126)	(0.0434)	(0.0353)	(0.111)	(0.120)
Problems with Usual Activities	-0.142	-0.0295	-0.00574	-0.0430	-0.0227
	(0.117)	(0.0327)	(0.0247)	(0.0998)	(0.105)
Problems with Pain/Discomfort	0.162**	0.0356	0.00805	0.0258	0.000431
	(0.0778)	(0.0218)	(0.0169)	(0.0635)	(0.0649)
Problems with Anxiety/Depression	-0.0273	-0.00534	0.000998	0.00630	0.0163
	(0.0637)	(0.0180)	(0.0137)	(0.0521)	(0.0550)
Household Income					
Less than $25,000	-0.134	-0.00723	0.0259	-0.152	-0.120
	(0.153)	(0.0488)	(0.0359)	(0.134)	(0.143)
$25,000-$49,999	Base	Base	Base	Base	Base
	-	-	-	-	-
$50,000-$74,999	0.137	0.00642	-0.0145	-0.0462	-0.0709
	(0.135)	(0.0390)	(0.0301)	(0.112)	(0.117)
$75,000-$99,999	0.242	0.0514	0.0151	-0.0747	-0.0993
	(0.151)	(0.0415)	(0.0325)	(0.121)	(0.125)
$100,000-$149,999	0.373**	0.0697	0.0159	-0.125	-0.171
	(0.160)	(0.0427)	(0.0338)	(0.127)	(0.130)
$150,000+	0.405*	0.105**	0.0450	-0.332**	-0.355**
	(0.210)	(0.0528)	(0.0422)	(0.163)	(0.163)
Missing/Refused to Answer	-0.134	-0.00807	0.0215	-0.103	-0.0870
	(0.170)	(0.0541)	(0.0410)	(0.146)	(0.156)
Times Individual Traded Lifespan and Health				0.173***	0.0854***
				(0.00761)	(0.00901)
Number of Observations	3,909	3,909	3,234	3,909	3,234
Alpha		0.730	0.203		

Note: Sex, race, ethnicity, marital status, duration of time with health problems, and work status are not reported.
Robust standard errors are in parentheses
*** p<0.01, ** p<0.05, * p<0.1

ABOUT THE AUTHOR

John graduated with a B.S.B.A. in Economics in August 2010 from the University of West Florida. Prior to entering the Ph.D. program in the Department of Economics at the University of South Florida, John worked as a math teacher at Woodham Middle School in Pensacola, FL. While a graduate student at USF, John also worked as a Research Coordinator in the Department of Health Outcomes and Behavior at Moffitt Cancer Center.