Safety and Effectiveness: The FDA's Approach to Risk in Prescription Medication

By

Elissa Philip Gentry

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Law and Economics

May 2016

Nashville, Tennessee

Approved:

W. Kip Viscusi, Ph.D.
James Blumstein, LL.B.
Daniel Gervais, Ph.D., LL.M
Melinda J. Beeuwkes Buntin, Ph.D.

To my loving parents, my amazing siblings, and my incredibly supportive husband. To God be the glory.

ACKNOWLEDGEMENTS

This work would not have been possible without the support of the Vanderbilt Law and Economics Ph.D. program and its founders, Professors Joni Hersch and W. Kip Viscusi. Without the program's financial, professional, and administrative support, I would not be where I am today. I especially want to thank Amy Cates and Laurel Donahue for tirelessly indulging my questions and advising me.

I especially would like to thank my advisor W. Kip Viscusi for all his support and advice throughout the years, both in the dissertation and in other projects. I have learned so much from working for and with him, and I am so thankful for having been given the opportunity to do so.

I am so grateful for my dissertation committee: Professors James Blumstein, Daniel Gervais, and Melinda Buntin each provided such valuable insight and feedback. Their input significantly improved my work, and I am profoundly grateful for the time and effort they invested in me.

I would also like to thank the faculty affiliated with the Law and Economics program for their feedback during numerous presentations: Professors Joni Hersch, Paige Skiba, Andrew Daughety, Jennifer Reinganum, Kathryn Anderson, Jennifer Shinall, and Peter Savelyev.

I am so thankful for the support of my peers in the program; I especially want to thank Jake Byl, Caroline Cecot, Ben McMichael, Sam Miller, and Blair Bullock for always providing sage advice and helpful feedback. I want to thank Henri Rautonen, Jean Xiao, Erin Meyers, Scott DeAngelis, and Danielle Drory for their help in conducting the experiment. I am particularly grateful to Professor Cindy Kam for valuable advice on the experiment and for the use of the RIPS lab.

I would also like to thank my friends and family. I am indebted to my parents and siblings, Philip and Susan Simon, Joanna Clayton, Drew Clayton, Jonathan Philip, and Michaela Philip for their endless support in this process. Finally, I could not have done this without the encouragement and advice of my husband Matthew Gentry. Thank you for reassuring me and providing perspective. I do not know what I would do without you in my corner.

TABLE OF CONTENTS

PAGE

DEDICATION……iii

ACKNOWLEDGEMENTS……iv

LIST OF TABLES……viii

LIST OF FIGURES……xi

INTRODUCTION AND OVERVIEW……1

CHAPTER

1. VALUATIONS OF AMBIGUITY IN PRESCRIPTION DRUG RISKS: EVIDENCE BY FRAMING AND RISK TYPE……3
I. INTRODUCTION……3
II. BACKGROUND……4
III. VALUATIONS OF AMBIGUITY IN SAFETY AND INEFFECTIVENESS……8
A. Experimental Design and Hypotheses……8
B. Data and Sample……17
IV. THEORETICAL AND EMPIRICAL MODELS……20
V. EMPIRICAL RESULTS……23
A. Evaluating Weights on Risk Estimates……23
B. Deviations from Neutrality……30
C. Discussion……32
1. Risk Framing……32
2. Safety vs. Ineffectiveness……34
VI. CONCLUSION……36
VII. REFERENCES……38
VIII. APPENDIX:……43
A. Dropping the First Round……43
B. Dropping Errors……47
C. Full Results……51
D. Random Effects for the Full Sample……53

2. EMPIRICAL EVIDENCE OF RISK PENALTIES FOR NTI DRUGS……51
I. INTRODUCTION……54
II. EXPLORING THE PERCEIVED RISKS OF NTI DRUGS……56
III. MODEL……61
IV. DATA……64
V. RESULTS……68
A. Descriptive Statistics……68
B. Price Regression……69
C. Substitutability Amongst Versions……77

VI. CONCLUSION 85
VII. REFERENCES 87
VIII. APPENDIX 92

3. RELINQUISHMENT OF INAPPROPRIATE OFF-LABEL USES: THE EFFECT OF THE FALSE CLAIMS ACT 98
I. INTRODUCTION 101
II. THE CURRENT LANDSCAPE OF OFF-LABEL DRUG USE 104
A. The Importance of Appropriate Off-Label Usage 104
B. Regulating Off-Label Uses through the False Claims Act 108
III. THEORETICAL MODEL 112
IV. AVERAGE EFFECT OF FCA SUITS ON RELINQUISHMENT 114
A. Data 115
B. Differences-in-Differences Analysis 117
V. NEURONTIN: A CASE STUDY 121
A. Testing the Relinquishment Hypotheses 122
B. Linear Probability Model 133
C. Triple-Differences 142
VI. CONCLUSION: CONSEQUENCES OF THE FALSE CLAIMS ACT 150
A. FCA as a Informational Learning Mechanism 151
B. Implications of Heterogeneity in Relinquishment by Payment 152
C. Conclusion 155
VII. REFERENCES 157

LIST OF TABLES

CHAPTER 1

Table 1 Experimental Design. 9

Table 2 Variable Definitions. 18

Table 3 Demographic Information. 20

Table 4. Fixed Effects Regression for Certainty Equivalents. 23

Table 5. Fixed Effects Regressions for Certainty Equivalents, by Order. 24

Table 6. Fixed Effects Regressions for Certainty Equivalents. 25

Table 7. OLS Regressions for Certainty Equivalents. 27

Table 8. Deviation from Ambiguity Neutrality[a] 31

Table A-1. Fixed Effects Regression for Certainty Equivalents. 43

Table A-2. Fixed Effects Regressions for Certainty Equivalents, by Order. 43

Table A-3. Fixed Effects Regression for Certainty Equivalents. 44

Table A-4. OLS Regressions for Certainty Equivalents. 45

Table A-5. Deviation from Ambiguity Neutrality. [a] 46

Table A-6. Fixed Effects Regression for Certainty Equivalents. 47

Table A-7. Fixed Effects Regressions for Certainty Equivalents, by Order. 47

Table A-8. Fixed Effects Regression for Certainty Equivalents. 48

Table A-9. OLS Regressions for Certainty Equivalents. 49

Table A-10. Deviation from Ambiguity Neutrality. [a] 50

Table A-11. Full Results for Table 8: Deviation from Ambiguity Neutrality. 51

Table A-12. Random Effects Regression for Certainty Equivalents. 53

CHAPTER 2

Table 1. Payment Per Unit, by NTI and Generic/Brand-Name Status[a] 69

Table 2. Price Sensitivity by Payer, Fixed Effects 71

Table 3. Price Sensitivity by Payer, Fixed Effects, Adjustment for Rebates. 75

Table 4. Probability of Mixing in a Given Period, Linear Probability Model. 79

Table 5. Version Loyalty: Linear Probability Model. 80

Table 6. Likelihood of Switching Between Bundles, Linear Probability Model. 83

Table A-1. Price Sensitivity by Payer, Fixed Effects, Tablets. 92

Table A-2. Price Sensitivity by Payer, Adjustment for Rebates, (Tablets only), Fixed Effects... 93

Table A-3. Price Sensitivity by Payer, Fixed Effects 94

Table A-4. Price Sensitivity by Payer, Fixed Effects, Adjustment for Rebates. 95

Table A-5. Price Sensitivity by Payer, Fixed Effects, Tablets. 96

Table A-6. Price Sensitivity by Payer, Adjustment for Rebates, (Tablets only), Fixed Effects... 97

Table A-7. Probability of Mixing in a Given Period, Probit. 98

Table A-8 Predicted Probabilities of Mixing, from Table A-13 98

Table A-9. Version Loyalty: Probit. 99

Table A-10 Predicted Probabilities from Table 15. 100

CHAPTER 3

Table 1. Number of Prescriptions by Payer and Year, 2005–2010: OLS Regressions 119

Table 2. Descriptive Statistics, 1998-2008 123

Table 3. Information Shock Hypothesized to Lead to Relinquishment. 124

Table 4. Likelihood of Prescribing Neurontin, Seemingly Unrelated Regressions and Linear Probability Models, 1998-2008. 136

Table 5. The Likelihood of Prescribing Neurontin, Separate OLS regression, Intervals. 140

Table 6. Number of Purchases[a], 1998-2008 143

Table 7. Number of Prescriptions, by Month, Year, and Coverage: Difference in differences. 144

TABLE OF FIGURES

CHAPTER 1

Figure 1. Sample Survey Screen 11

CHAPTER 3

Figure 1. Percent Prescriptions by Year 125

Figure 2 Percent of Neurontin Prescriptions by Payer and Year 131

INTRODUCTION AND OVERVIEW

The FDA ensures that prescription drugs in America are generally "safe and effective" for the uses for which they are prescribed. Given the inherent tradeoff between rigorously vetting drugs and delaying alternative treatments for patients, however, there are areas in which prescription risks are left unregulated or are unable to be precisely quantified. In these areas where drug risks can be heterogeneous, physicians' ability to assess, and the market's ability to price, risk is particularly important. This dissertation focuses on the legal framework surrounding consumer drug choice.

Chapter 1 begins by measuring consumers' preferences with respect to ambiguity in the risk that a drug is safe or ineffective in the presence of framing effects. It uses an incentivized experiment on Vanderbilt undergraduates to understand consumers' preferences if they were given explicit information about drug risks. These deviations from ambiguity neutrality depend on context: the Chapter finds that framing effects result in ambiguity-seeking behavior in the "loss domain" and ambiguity-averse behavior in the "gains domain." The results suggest that alleviating some ambiguity through uniform recommendations or physician intermediaries might actually lead to patients choosing better treatment options.

Chapter 2 empirically examines the consumption of narrow therapeutic index (NTI) drugs, an area in which there was a perceived but unregulated risk. The Chapter tests whether third-party payers contributions are consistent with patients' preferences against risky drugs, particularly drugs in which switching is costly. The results support this idea, as there is evidence of a price penalty for NTI status in all first- and third-party payers' contributions. The gap between brand-name and generic drugs is also smaller for NTI drugs than for non-NTI drugs, consistent with costly switching. Finally, the Chapter examines consumption behavior and finds

evidence of version loyalty, consistent with costly switching. The Chapter concludes that third-party payers might be adequate agents of patients in taking risk preferences into account in their pricing decisions even absent direct government regulation.

Chapter 3 measures the effect of a False Claims Act settlement on relinquishment of inappropriate off-label uses. It finds a significant average drop in off-label prescriptions after an FCA suit is settled. It further utilizes a case study of neurontin to disentangle the response of different payers to both scientific and legal information shocks. It finds heterogeneity in relative relinquishment by payer and finds evidence consistent with the relinquishment being spurred by the litigation process.

This dissertation seeks to understand the relationship between risk preferences and regulation in the context of pharmaceutical drugs. Chapter 1 discusses the underlying issue of ambiguity in pharmaceutical drug risks. Chapters 2 and 3 examine two contexts in which drug risks were uncertain or not well-quantified and discuss whether explicit regulation is necessary. Chapter 2 finds that the market seems sensitive to risk even in the absence of government regulation. Chapter 3 finds that the False Claims Act can provide industry signals as to the appropriateness of off-label uses but cautions that this function may be undermined by further expansion of the FCA. The results provide an insight into how the legal framework surrounding pharmaceutical regulation interacts with consumer risk preferences and where explicit regulation is necessary and where the market might be well-equipped to compensate risk.

CHAPTER 1

VALUATIONS OF AMBIGUITY IN PRESCRIPTION DRUG RISKS: EVIDENCE BY FRAMING AND RISK TYPE

I. Introduction

Prescription drug use is a widespread phenomenon in the United States. Almost half of the population has used at least one prescription drug within the past 30 days.[1] Prescription drug use is almost always associated with some level of risk. While the Food and Drug Administration (FDA) requires drugs to be both safe and effective (21 U.S.C. 355(j); Wittich et al. 2012), these drugs are not risk free. Some drug uses, however, have higher and more ambiguous risks than others. How consumers react to this ambiguity is important, as it potentially affects their decisions over uncertain treatments.

One context in which risks are particularly ambiguous is in off-label uses of drugs. A drug is used "off-label" when it is prescribed for a disease, population, or dosage for which it is not approved (Wittich et al. 2012). For example, a drug approved for the control of seizures might be prescribed for the treatment of bipolar disorder. The FDA allows physicians broad discretion in how to prescribe a drug. Since companies do not have incentives to submit these uses for FDA approval, rigorous randomized double-blind controlled studies may be less likely to be available for off-label uses. Instead, physicians may rely on pharmaceutical representatives and available scientific information to determine whether off-label prescriptions are appropriate. This results in risks that are less precisely defined and ambiguous.

In this Chapter, I use an incentivized experiment to explore differences in ambiguity attitudes based on framing effects. A person's attitudes toward ambiguity usually fall into one of

[1] The CDC reports 48.7% of people (men and women) from 2009–2012 have taken at least one prescription drug in the last 30 days (CDC 2014).

three categories: he may prefer ambiguity ("ambiguity-seeking"), avoid ambiguity ("ambiguity-averse"), or be indifferent to ambiguity ("ambiguity-neutral"). Framing effects are important, given that consumers' prior treatment experiences likely provide a reference point with respect to which current attitudes are formed. Additionally, the experiment measures differences in ambiguity attitudes between the risk that the drug is ineffective and the risk the drug is unsafe. A person's attitudes toward ambiguity are difficult to empirically ascertain given asymmetric information on drug benefits; this experiment seeks to elicit these through an experiment where this information is explicit. Understanding these attitudes is important: if consumers are ambiguity-averse or ambiguity-seeking they may consume a suboptimal mix of drugs. If consumers are predictably influenced by ambiguity in particular frames or for particular risks, the government might be able to implement policies to reduce perceived ambiguity and help patients choose a treatment option that better maximizes their expected outcomes.

II. Background

This study seeks to incorporate the effect of ambiguity into consumer valuations of pharmaceutical risks. Drug risks are inherently hard to state with certainty—even FDA-approved drugs have additional risks that are discovered years after approval. Ambiguity is particularly relevant for off-label treatments. Off-label drug uses are uses for which a drug did not receive FDA approval. The risks of off-label treatments are not necessarily unknown, as there are studies available on the treatment's safety and effectiveness. However, these studies are often case studies or small sample studies, which do not provide the same level of "certainty" as a double-blind controlled study. In a survey of off-label mentions, Radley et al. (2006) find that only 27% were supported by strong scientific evidence. In contrast, on-label uses have a better chance of identifying their risks, since double-blind controlled studies are often necessary for

FDA approval.[2] This study examines the effect of ambiguity on perception of pharmaceutical risks, focusing on how the effect of ambiguity varies with framing and by type of risk.

Framing is important in the medical context, particularly if patients exhibit reference dependence in their preferences and utilities. Consider two patients choosing between an approved treatment with a certain likelihood of success and an off-label treatment with an ambiguous likelihood of success: One patient was well-treated on their previous drug but no longer has access to it and is looking for a new treatment. The other patient had a bad experience with her previous drug and discontinued its use. These patients, though facing similar problems, may respond differently to the ambiguously risky off-label treatment simply due to their past treatment experience.

Similarly, the type of risk faced might affect reactions to ambiguity. This Chapter focuses on two types of risk: safety and ineffectiveness risks. For the purpose of this Chapter, a "safety risk" will be defined as the risk that the drug causes an adverse effect, and an "ineffectiveness risk" will be defined as the risk that a drug does not alleviate the symptoms of a patient's disease. There are legitimate reasons for consumers to value safety and effectiveness differently. Consumers might prefer the ailment they have experienced (their disease's symptom) to an ailment they have not experienced (a drug's side effect). Similarly, consumers may prefer not to be actively involved in the harm, as discussed in the omission bias and related literature (e.g., Ritov and Baron 1990; Zeckhauser and Viscusi 1990; Connolly and Reb 2003[3]).[4] Consumers

[2] Junod, Suzanne White (2014). *FDA and Clinical Trials: A Short History*, FOOD AND DRUG ADMINISTRATION, *available at* http://www.fda.gov/AboutFDA/WhatWeDo/History/Overviews/ucm304485.htm#_edn52.

[3] In contrast, Connolly and Reb (2003) find little evidence of omission bias in vaccination decisions.

[4] Kees, Bone, Kozup, and Ellen (2008) find evidence of the omission bias in a study of black box warnings and risk presentation on prescription drug websites. Omission bias has also been

may prefer to let their disease worsen through inaction rather than affirmatively harm themselves with a drug, even though it has a chance of alleviating their symptoms.[5]

Several studies have focused on ambiguity in medical decision-making contexts. Many concentrate on ambiguity in disease-specific contexts: Han et al. (2011) find evidence of both ambiguity-aversion and ambiguity-seeking in the context of individualized colorectal cancer risk, with participants less likely to experience ambiguity-aversion if they were optimistic or experienced visual communication of probabilities.[6] Similarly, Han et al. (2007) study the effect of ambiguity perceptions in mammography recommendations and find that perceived ambiguity is associated with lower likelihood of receiving a mammography. Curley, Eraker, and Yates (1984) examine ambiguity-aversion in patients' choice of treatment, finding that confidence and context were significant in whether a patient was ambiguity-averse. A similar literature exists for physicians' choice of treatment (Curley, Young, and Yates 1989; Gerrity, DeVellis, and Earp 1990). Viscusi (1999) asks judges whether they would choose to market a superior CAT scan drug with uncertain risk properties; he finds that the majority of judges chose the old drug with the known but higher risk. Additionally, he finds higher valuations of one's own life are correlated with higher likelihoods of approving the uncertain risk drug.[7]

---

studied in the context of physicians decided treatment regimes. Through a series of vignettes, Aberegg, Haponik, and Terry (2005) find that pulmonary and critical care physicians exhibit omission bias.

[5] Zeckhauser and Viscusi (1990) note that the issue of omission might exist for two reasons: first, consumers might react adversely to the idea that the government is comfortable with risk to the public. Second, consequences of risks of commission can be more visible than risks of omissions.

[6] Han et al. (2009) do a similar analysis.

[7] While there has been work on ambiguity in medical decision making, another line of literature more rigorously characterizes ambiguity as a Bayesian learning process (Viscusi and Magat 1992; Viscusi 1997). Similarly, previous literature discusses the effect of ambiguity on risk perceptions, including the effect of extreme estimates (Viscusi 1997), the range and mean of estimates (Viscusi and Chesson 1999), different sources of information (Viscusi 1997), and asymmetric risks (Viscusi, Magat, and Huber 1991b).

The effects of framing and safety/effectiveness on ambiguity have largely been studied separately. The effect of framing on ambiguity has often been studied outside the medical context. Einhorn and Hogarth (1985) acknowledge framing in their descriptive model, proposing that participants anchor on an initial estimate and then adjust in response to ambiguity. Using experiments, Kahn and Sarin (1988) find evidence of ambiguity-aversion in the gains domain and ambiguity-seeking in the loss domain at high mean probabilities; they find the reverse at low mean probabilities. Viscusi and Chesson (1999) also see evidence of this reversal and estimate the crossover point between high and low mean probabilities. Unlike previous studies, however, they use a Bayesian learning model as the underlying framework.

The prior literature on safety and effectiveness has mostly focused on the relative values of safety and effectiveness in various disease-specific contexts (e.g., osteoporosis (de Bekker-Grob et al. 2008), hip fracture (Telser and Zweifel 2002), and asthma (McTaggart-Cowan et al. 2008)). These studies often produce mixed results as to whether patients value safety or effectiveness more. Johnson et al. (2007) find that patients with Crohn's Disease were often willing to accept higher adverse event risks in exchange for improved efficacy. In contrast, a study of antiepilepsy drugs finds that patients were willing to pay less to reduce the chance of seizure symptoms than to avoid side effects such as feeling sick, skin rash, or hair loss (Lloyd, McIntosh, and Price 2005).

Perhaps the most relevant prior study is Bier and Connell (1994), which examines ambiguity in the context of pharmaceutical consumption. Their experimental design does vary safety and ineffectiveness risks; however, the study focuses on the effect of optimism and message framing. Bier and Connell find that subjects were ambiguity-seeking in positive frames, especially for more optimistic subjects, and ambiguity neutral in negative frames. This runs

counter to previous framing research, which finds ambiguity-aversion in the gains frame (Kahn and Sarin 1988; Hogarth 1989). Bier and Connell, additionally, do not characterize ambiguity as a Bayesian learning process.

This Chapter contributes to the literature by looking at framing and safety/effectiveness in conjunction. In order to do this, the study employs a model of Bayesian learning to examine the effect of framing on the perception of safety and ineffectiveness risks in the presence of ambiguity.

III. Valuations of Ambiguity in Safety and Ineffectiveness

This experiment[8] examines the effect of ambiguity on the perception of the risk that a drug is ineffective or unsafe. The experiment studies ambiguity effects in consumer drug choice, using the pairwise comparison methodology established in Viscusi, Magat, and Huber (1991a); Viscusi (1997); and Viscusi and Chesson (1999). Consumers will choose between a drug with a certain risk and a drug with an ambiguous risk until they reach a point of indifference.

A. Experimental Design and Hypotheses

Participants are faced with a pairwise choice between two drugs: One of the drugs has a known, certain probability of risk while the other has two different estimates of risk, resulting in an "ambiguous" risk. Participants gain and lose money based on the risks associated with the drug they choose. Table 1 outlines the experimental design.

Participants are assigned to Group 1 (*Ineffectiveness*) or Group 2 (*Safety*). Each group faces ambiguity in the risk of experiencing headaches: participants in the *Ineffectiveness* group experience the risk of headaches as a symptom of what the survey calls Disease A, while participants in the *Safety* group experience the risk of headaches as the side effect of the drug in

[8] This experiment went through Vanderbilt University IRB approval (study #150977).

question, called Xites. Both groups face a certain, stable risk of experiencing nausea:

ineffectiveness experiences it as a Xites side effect and *Safety* as a Disease A symptom.

Table 1 Experimental Design.

	Group 1 (Ineffectiveness)		**Group 2 (Safety)**	
Disease Symptoms	Headaches		Nausea	
Drug Side Effects	Nausea		Headaches	
	Drug C	**Drug D**	**Drug C**	**Drug D**
Risk that Drug Fails to Alleviate Symptoms	Certain	Ambiguous	Certain, static	Certain, static
Risk that Drug Causes a Side Effect	Certain, static	Certain, static	Certain	Ambiguous

Participants in the *Ineffectiveness* group received the following instruction:

> You will be given an option of two drugs: one drug has a known probability of preventing your symptoms and the other has an ambiguous probability. Please make a decision between the two drugs in each round. After each round, based on the risks associated with the drug chosen, a health event will occur that will either earn or deduct money. We are interested in the point at which you are indifferent between the two drugs.
>
> You were diagnosed with Disease A, which causes you to experience intense headaches 4-5 times a week. These headaches leave you sensitive to light and incapacitate you. No over-the-counter medication alleviates your symptoms. Your current prescription medication for Disease A does not control your symptoms. Instead, it also produces intolerable side effects. For these reasons, your physician refuses to prescribe it anymore. Instead, he suggests that you switch to a new type of drug, Xites.
>
> There are many types of Xites drugs. While Xites can control Disease A, they are also associated with certain side effects. In particular, Xites can cause bouts of nausea. These bouts cause sensitivity to smell and often confine you to bed. Because of drug interactions, additional drugs cannot be taken to alleviate these adverse side effects. Over the next slides you will see a pair of Xites prescription drugs. You must choose one of the two drugs, or indicate that you are indifferent between the two. If you indicate that you are indifferent, one of the drugs will be randomly selected.

Similarly, participants in the *Safety* group received the following instruction:

> You will be given an option of two drugs: one drug has a known probability of causing side effects and the other has an ambiguous probability. Please make a decision between the two drugs in each round. After each round, based on the risks associated with the drug chosen, a health event will occur that will either earn or deduct money. We are interested in the point at which you are indifferent between the two drugs.

> You were diagnosed with Disease A, which causes you to experience bouts of nausea 4-5 times a week. Each bout causes sensitivity to smell and often confines you to bed. No over-the-counter medication alleviates your symptoms. Your current prescription medication for Disease A does not control your symptoms. Instead, it also produces intolerable side effects. For these reasons, your physician refuses to prescribe it anymore. Instead, he suggests that you switch to a new type of drug, Xites.
>
> There are many types of Xites drugs. While Xites can control Disease A, they are also associated with certain side effects. In particular, Xites can cause intense headaches. These headaches cause sensitivity to light and incapacitation. Because of drug interactions, additional drugs cannot be taken to alleviate these adverse side effects. Over the next slides you will see a pair of Xites prescription drugs. You must choose one of the two drugs, or indicate that you are indifferent between the two. If you indicate that you are indifferent, one of the drugs will be randomly selected.

Headaches and nausea were chosen since they seem similar in severity, so as not to create a comparison effect between a minor symptom and a severe side effect (or vice versa). Using this design, I can compare the ambiguity effect for the same ailment (headaches) in different contexts.

For each group, Drug C has a certain probability of incurring the side effect of headaches, while Drug D has two estimates. *Low Risk* corresponds to the low estimate of risk for Drug D, while *High Risk* corresponds to the high estimate of risk for Drug D. Depending on the participants' choice, the computer will update the certain probabilities for the Drug C until an indifference point is reached between the two drugs. A sample screen that participants might see is included as Figure 1.

The initial probability associated with Drug C is equal to either *High Risk* or *Low Risk* for Drug D. For example, for Drug D with ambiguous risk estimates 100/1000 and 300/1000, some participants initially faced a Drug C option with a certain probability of 100 out of 1000 (a circumstance noted by *Ascending*=1) while others faced a certain probability of 300 out of 1000 (a circumstance noted by *Ascending*=0).

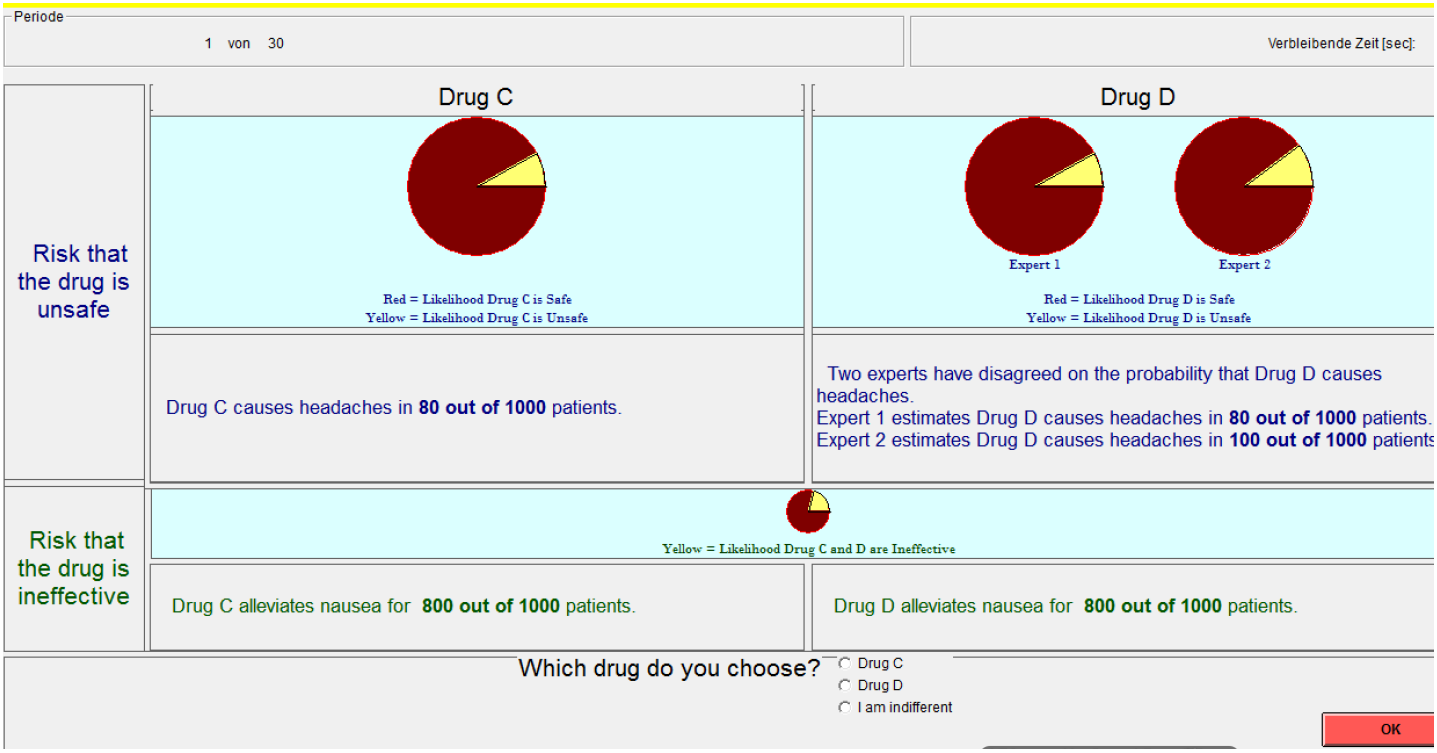

Figure 1. Sample Survey Screen

Hypotheses. The experiment tests two hypotheses: The first involves framing effects with respect to a probability reference point. Previous literature suggests that participants will display ambiguity-averse behavior in gains and ambiguity-seeking behavior in losses, for sufficiently high mean probabilities (Kahn and Sarin 1988; Viscusi and Chesson 1999). Similarly, Viscusi, Magat, and Huber (1987) examine reference dependence in probabilities and find that the decrease in value generated by an increase in risk from a reference probability point is larger than the corresponding increase in value generated by a comparable decrease in risk. increases in risks from a reference probability point were valued significantly greater than comparable decreases. This Chapter will discuss framing with respect to reference dependence.

Hypothesis 1: Consumers facing an *Ascending*=1 will exhibit ambiguity-seeking behavior and those facing Ascending=0 will exhibit ambiguity-averse behavior.

For this Chapter, the appropriate reference point is a participant's initial certain probability (the probability associated with Drug C). When *Ascending*=1, participants are better off with the certain probability than with the gamble, forcing them to initially choose Drug C. With each period that they choose Drug C, the certain probability of harm increases and becomes decreasingly attractive. Situations in which *Ascending*=1 can thus be considered within a "loss domain." Conversely, when *Ascending*=0, participants are better off with the gamble (Drug D), since their certain probability is equal to the gamble's highest estimate of risk. Since a participant's certain probability of harm decreases (becomes more attractive) for each round that they chose Drug D, this can be seen as a "gains" domain. If participants see Drug C's certain probability as their reference point, prior literature would suggest that *Ascending*=0 would exhibit ambiguity-averse behavior, and *Ascending*=1 would exhibit ambiguity-seeking behavior.

The described reference dependence has analogies to prospect theory. Prospect theory suggests that value is determined by changes in utility, rather than final level (Kahneman and

Tversky 1979). Prospect theory predicts that people are risk averse in gains but risk seeking in losses. Notably, Kahneman and Tversky explain prospect theory in the context of gambles over outcomes (risk aversion) instead of gambles over probabilities (ambiguity-aversion).

Hypothesis 2: Consumers are relatively more ambiguity-averse with respect to *Ineffectiveness* than in *Safety*.

The second hypothesis predicts that people will have different attitudes toward ambiguity in safety and ineffectiveness risks. There are a couple of reasons to think that people may have different perceptions of ambiguity in each context. First, a person might be more familiar with the symptom he already knows, such as the symptom of his disease, than the symptom he does not, such as a new side effect. Prior experience with the symptom of his disease might make the person find the risk of more symptoms more salient. In contrast, a person might be more optimistic about their ability to "beat the odds" with an unfamiliar ailment. Second, the idea of loss aversion provides a possible explanation. Tversky and Kahneman (1991) describe a value function centered on an endowment point corresponding to a person's original allocation of a good. Value decreases more due to losses from the endowment point than it increases due to equivalent gains. If the familiarity with his symptoms causes a person to form an endowment point around *Ineffectiveness*, such that he feels attached to the level of symptom management he currently has, loss aversion suggests that participants will be less willing to experience an increase in symptoms.[9]

Methodology. As stated above, a participant faces a choice between a drug with a certain probability of harm and a drug with an ambiguous probability of harm. After each drug choice,

[9] It is also conceivable that participants feel an endowment in being side effect free, which would imply more ambiguity-aversion in *Safety* than *Ineffectiveness*. The argument here is that prior experience with symptoms might make the endowment effect stronger for *Ineffectiveness* than *Safety*.

the computer reports which health outcome occurred based on the probabilities associated with the drug chosen. Participants were paid based on their performance in the experiment. Each participant began with $10 in his account. In any round that the harm does not occurs, participants earn money; when a harm occurs, a penalty is deducted. Five rounds were randomly chosen at the end of the experiment to be paid out (added and subtracted from their original account total). [10] This way, participants are incentivized to choose the drug that they feel minimizes their chance of an adverse health outcome. Since the rounds were randomly paid, there is little concern that participants would lose interest after they have hit their target award level, or become disheartened after their account level drops sufficiently low.

Each risk is signified as in terms of the numerator x out of a population of 1000 people. This technique for portraying risk has been found to be more comprehensible than providing a decimal (Viscusi, Magat, and Huber 1991a; Coaster et al. 2011). The base of 1000 instead of 100 was chosen in order to be able to portray probabilities such as .135 as whole number numerators, providing more diversity in probabilities displayed. For simplicity, since the denominator is always the same, I will only list the numerator. Additionally, when participants viewed the "ineffectiveness" risk, they actually saw that the drug alleviated a symptom for (1000-x) out of 1000 people. Despite this, throughout the remainder of this analysis, these probabilities are analyzed as x/1000, for comparability with the safety risk. For example, even though a participant sees that a drug has a certain chance of alleviating symptoms for 700 out of 1000 people, the analysis will see this as a 300 out of 1000 risk of being ineffective.[11] Participants

[10] These rounds were drawn with replacement, so there is a possibility that a round could be randomly selected twice. However, each round had an equal likelihood of this, so this should not affect expectations.

[11] This wording decision was made after a focus group test, in which participants said they thought the positive wording for ineffectiveness was more natural and not confusing. The probabilities are inverted for analysis so that both probabilities are risks of harm.

received (low, high) risk estimates of (100, 300), (120, 280), (100, 900), (250, 750), (10, 200), (52 158), (80, 100), (45, 135), (615, 735), and (405, 945).[12] These were chosen as scaled-up version of some of the probabilities used in Viscusi (1997), along with a mean-preserving spread.

The initial choice is a rationality check. For example, a participant might first face a choice between Drug C, which has a certain risk of being ineffective of 300, or Drug D, which has a risk of either 100 or 300. The dominant choice is to choose Drug D, since its high estimate of the risk is equal to the certain probability associated with Drug C (and its low estimate is lower than Drug C's certain probability). If participants instead choose the strictly dominated option (in this example choosing Drug C), they receive a prompt asking them if they are sure about their decision. The rationality check continues throughout the experiment; a prompt will be triggered if the participant tries to find indifference between the certain risk and a risk above the higher risk estimate or below the lower risk estimate.[13] Participants are able to choose a

[12] For the first round, only two probability pairs were drawn for *Ascending*=0. This should not matter in aggregate for two reasons: The remaining probability pairs were eventually drawn with roughly equal probability in the subsequent rounds. The probability pairs drawn in the first round were also drawn in subsequent rounds, ensuring that the probability pair was not conflated with a particular round. Additionally, this error was symmetric across *Ineffectiveness* and *Safety*. In case this is a problem, however, the Appendix contains tables dropping the first round (only considering rounds in which all probability pairs are drawn with nonzero probability). Additionally, I run a robustness check in which regressions are weighted by the inverse frequency of risk pairs by *Ascending*.

[13] This rationality check goes on throughout the experiment, although sometimes it was only shown right after a strictly dominated choice was chosen. This seems to happen for (52, 158) because the rationality check is triggered when a certain choice is equal or less than the lower estimate or if the certain check is equal or greater than the upper bound. Based on the way z-tree processes decimals, when intervals involved decimals, equality was not achieved and the rationality check was not triggered until the next round. This should not be a problem, however, since this was only a learning device and no choice is restricted based on the rationality check. In case the decimal issue affected rounding such that participants did not see exact probabilities (e.g., they see 63 instead of 62.6), I do a robustness check in which I drop all observations facing (52, 158); the results remain qualitatively the same.

dominated option if they click through the prompt; the prompt merely alerts them to the possibility of a better choice.[14]

Pie charts were added to visually signify the risk associated with each drug. In order to aid in making the harms salient, graphics were also added to the health outcomes. For example, if the health outcome was that the participant experienced headaches, they saw a stick figure with arrows radiating from its head. Additionally, the text was color-coded, such that the risks for ineffectiveness were in one color and risks for safety in another. When an outcome involved safety, the outcome was reported in its respective color. These measures were taken to ensure that participants link each risk to their outcomes and can differentiate between risks.

The certain probability at which participants choose "I am indifferent" is considered the indifference point, or the *Certainty Equivalent*. However, sometimes participants switch from choosing Drug C to Drug D (or vice versa). Instead of just considering the midpoint of these probabilities as the indifference point, the program allows the participant to make one more choice between Drugs C and D, telescoping to find an indifference point before moving to new probabilities. Suppose a participant facing the gamble of (100, 300) and a certain risk of 280 chooses the gamble. If next period she prefers the certain probability of 260 to the gamble, this counts as a switch. The program will then offer her a choice between the gamble of (100, 300) and 270. If she chooses the gamble, her *Certainty Equivalent* is 265.

After making their drug choices, participants completed a demographic survey, including questions on previous prescription drug experience, health experience, risk attitudes, and cognitive reflective abilities.

[14] The prevalence of violations of the rationality check is discussed on page 23.

B. Data and Sample

I ran this experiment on Vanderbilt University undergraduates over a period of five days. The experiment involved 175 participants, each completing 30 periods. Each drug choice a participant makes is considered a separate period—each participant completes 5 practice periods and 25 true periods. The practice periods are not considered in the analysis, resulting in 4,375 observations for 175 participants. A round, in contrast, is defined by finding indifference between the *Certainty Equivalent* and the gamble. Participants may complete a different number of rounds depending on how many periods it takes to find indifference. The 4,375 periods resulted in 787 rounds (and in turn, 787 *Certainty Equivalents*).

The variables gathered from the experiment are summarized in Table 2. *Low Risk*, *High Risk*, *Certainty Equivalent*, *Safety*, and *Ascending* are defined above. *Cumulative Harm* denotes the number of times a participant experiences a penalty during a round. Other variables from the survey include demographic attributes such as age (*Age*), sex (*Male)*, reported race (*Race),* political affiliation (*Republican, Democrat, Independent,* or *Other*), and smoking status (*Smoker*).

Table 2 Variable Definitions.

Variable	Definition
Low Risk	The low estimate of risk for Drug D.
High Risk	The high estimate of risk for Drug D.
Certainty Equivalent	The certain probability in which participant is indifferent to gamble.
Deviation	$(\frac{Certainty\ Equivalent - Low\ Risk}{High\ Risk - Low\ Risk}) - 0.50.$ Reports deviation from ambiguity neutral decision.
Safety	Group experience ambiguity in risk that drug is unsafe (omitted is *Ineffectiveness,* the group experiencing ambiguity in the risk drug is ineffective).
Cumulative Harm	The number of times a participant experiences a penalty during a round.
Ascending	Participant experiences a certain probability equal to the Low Risk initially, and the gamble is strictly dominated. If *Ascending*=0, participant's initial certain probability is equal to High Risk.
Age	Participant's reported age.
Experienced	Participant reports experiencing headaches or nausea at some point, either past or present.
Risk Averse	In a choice of 14 horse races, after winning big in the 4th round, participant chooses to take their winnings rather than continue playing.
Race	Reported race: options are white, African American, Asian, Hispanic, and other.
Student	Reports full-time student status.
Male	Reported sex as male.
Political	Reported political party: Options are Republican, Democrat, Independent, or other.
Off Label	Reported using a drug off-label.
Prescription Use	Reports using a prescription drug, either past or present.
Smoker	Reports smoking regularly or occasionally.
Econ	Reports undergraduate or graduate major in economics.
Premed	Reports major in premedical studies.
CRT	Answers all of the CRT questions correctly.

If students reported using a prescription drug, either in the present or in the past, *Prescription Use* =1. This variable serves as a measure of familiarity with prescription drugs and, presumably, the risks associated with them. Similarly, *Off Label* =1 if students report knowing that they use an off-label treatment. Finally, *Experienced*=1 if students report experiencing headaches or nausea, either in the present or past. This is important, as students with experience with each ailment might take the probabilities more seriously. Students were questioned about their majors: I differentiate those choosing economics degrees (either undergraduate or graduate — *Econ*) or premed (*Premed*), as these majors might be more familiar dealing with probabilities or drug risks.

Part of the demographic survey included answering three questions from the Cognitive Reflection Test (Frederick 2005), which is designed to differentiate between intuitive System 1 and methodical System 2 thinking. System 1 processes are spontaneous or intuitive responses, while System 2 processes require mental effort and "execution of learned rules." If the participant answers all of the three questions correctly, *CRT*=1. Finally, a measure of "risk aversion" is included; students are told that there is a series of 14 horse races and are told that they won big in the 4th race. They are asked whether they want to continue playing or leave with their earnings. If participants indicate that they would take their winnings, *Risk Averse*=1.[15]

The demographics of the participants are summarized in Table 3. Several significant facts stand out: The majority of participants are undergraduate freshmen, as the mean age is 18.55. The sample is 72% male, which is higher than the underlying population. However, there does seem to be significant experience with prescription drugs, as around 70% report using a prescription drug in either the past or present. This may be a function of prevalence of birth control pills or ADHD medication. Finally, around 70% of the sample experienced headaches or nausea at some point. The sample is evenly split across political groups. Interestingly, around 22% of the sample answered the three CRT questions correctly.

[15] This measure was taken from Kam and Simas (2010).

Table 3 Demographic Information.

Variable	Mean
Age	18.55
Experienced	0.69
Risk Averse	0.70
Race	
White	0.61
African American	0.06
Asian	0.22
Hispanic	0.07
Other	0.04
Male	0.72
Political	
Republican	0.30
Democrat	0.33
Independent	0.29
Other	0.08
Off Label	0.13
Prescription Use	0.68
Smoker	0.18
Econ	0.27
Premed	0.22
CRT	0.22

IV. Theoretical and Empirical Models

This Chapter uses the above data to estimate the following model, established by Viscusi (1997) to understand Bayesian learning in the context of uncertainty. The model forms equivalence between risk perceptions and known probabilities. The risk perception *R* can be seen as follows:

$$(1)\ R = \frac{\gamma}{\gamma + \zeta + \zeta^*} q + \frac{\zeta}{\gamma + \zeta + \zeta^*} r + \frac{\zeta^*}{\gamma + \zeta + \zeta^*} r^*,$$

a function of prior risk perceptions, q, and the two risk estimates, *r* and *r**. The weights associated with each estimate are signified by ζ and ζ^*, respectively. The experiment compares two drugs, one with a known risk, *S*, and one with an ambiguous risk. This ambiguous risk corresponds to some risk perception, *R*. As noted above, the computer continues to update *S* until

the participant is indifferent between the two drugs. At this indifference point, the following is true:

$$(2)\ \ (1-R)\times U(Healthy) + R\times U(Injured) = (1-S)\times U(Healthy) + S\times U(Injured),$$

Viscusi (1997) uses this model to discuss the informational content of signals r, r*, as they come from different sources (either industry or government sources). Here, this experimental setup explores whether r and r* vary based on the type of risk (e.g., safety risk vs. ineffectiveness risk) or the context (in the loss domain or the gains domain) in which the participant experiences ambiguity.

Since the experiment elicits the point at which a participant is indifferent between the uncertain and certain risks, this simplifies to

$$(3)\ \ S = \frac{\gamma}{\gamma+\zeta+\zeta^*}q + \frac{\zeta}{\gamma+\zeta+\zeta^*}r + \frac{\zeta^*}{\gamma+\zeta+\zeta^*}r^* + \varepsilon.$$

Using the data described in Section (b), I try to estimate Equation (3). Since I do not observe each individual's prior beliefs, I include person–fixed effects, which should capture invariant differences between participants regarding prior expectations of risk. I thus estimate the following fixed effects regression:

$$(4)\ \ S = \beta_1 r + \beta_2 r^* + \gamma_i + \varepsilon.$$

Equation (4) is based on Equation (3), where $\beta_1 = \frac{\zeta}{\gamma+\zeta+\zeta^*}$, $\beta_2 = \frac{\zeta^*}{\gamma+\zeta+\zeta^*}$, and γ_i is the participant fixed effect. Here, r is the low estimate of risk while r* is the high estimate of risk. If $\beta_1 + \beta_2 = 1$,[16] the following statements can be made: $\beta_1 = \beta_2$ would suggest ambiguity neutral behavior; that is, participants put equal weight on the high and low estimates. In contrast, if $\beta_1 > \beta_2$, participants place greater weight on the low estimate in determining their indifference

[16] If $\beta_1 + \beta_2 > 1$ or $\beta_1 + \beta_2 < 1$, this analysis becomes more complicated.

point, suggesting ambiguity-seeking behavior. If $\beta_1 < \beta_2$, participants place greater weight on the high estimate, suggesting ambiguity-averse behavior.

Since individuals are assigned to *Safety* or *Ineffectiveness* alone, a fixed-effect regression cannot estimate this treatment effect.[17] Thus, in conjunction, an OLS regression is run:

$$(5)\ S = \beta_1 r + \beta_2 r^* + Safety'\beta_3 + X'\delta + \varepsilon,$$

where X is a vector of demographic variables. X also includes variables such as including *Prescription Use*, *Off-Label*, *CRT*, and *Experienced*, *Risk Averse*, *Smoker*, *Econ*, and *Premed*. Errors are clustered by participant.

Equations (4) and (5) are important in estimating $\frac{\zeta}{\gamma+\zeta+\zeta^*}$ and $\frac{\zeta^*}{\gamma+\zeta+\zeta^*}$. However, the interpretation of the variables in *X* for Equation (5) is not straightforward in this context. In order to understand what factors cause participants to move away from the ambiguity-neutral choice, I transform the observed *Certainty Equivalent* into a measure of deviations from the ambiguity-neutral *Certainty Equivalent*. The OLS regression is as follows:

$$(6)\ Deviation = \beta_1 Safety + X'\delta + \varepsilon\ ,$$

where *Deviation* is defined as $abs((\frac{Certainty\ Equivalent - Low\ Risk}{High\ Risk - Low\ Risk}) - 0.50)$, as defined in Table 2. Since an ambiguity neutral participant would choose the mean of the two risk estimates, *Deviation* would then be 0. *Deviation* for an ambiguity-seeking or ambiguity-averse participant would be nonzero. Thus, factors that minimize deviations from ambiguity neutrality will have negative coefficients, while factors that promote a movement toward either ambiguity-seeking or ambiguity-aversion will have positive coefficients.

[17] A random effects model is included in the Appendix, Table A-12. This controls for unobserved heterogeneity but does allow for estimation of a treatment effect. As shown, the results are very similar to the OLS results in Table 7

V. Empirical Results

A. Evaluating Weights on Risk Estimates

Table 4 lists the results for Equation (4). There is weak evidence that participants are slightly ambiguity-averse in *Ineffectiveness* but not in *Safety* given the lower weight on *Low Risk*, relative to the weight on *High Risk*. Column (1) lists the weights corresponding to the low and high estimates of risk for all the observations, while Columns (2) and (3) list the weights for the *Ineffectiveness* and *Safety* groups, respectively. The estimated coefficients seem almost identical for both groups, indicating that, on the whole, participants took the average of the two risk estimates. F-test statistics are provided in Table 4 and all other fixed-effects tables; these test whether the coefficient on *High Risk* is equal to that of *Low Risk*. The F-tests indicate that the *Low Risk* coefficient is indistinguishable from the *High Risk* coefficient for *Safety*; however, the *Low Risk* coefficient is often significantly different than the coefficient on *High Risk* for the *Ineffectiveness* group.[18] This suggests that, on the whole, consumers are slightly ambiguity-averse in *Ineffectiveness,* providing support for Hypothesis 1.

Table 4. Fixed Effects Regression for Certainty Equivalents.

	All	Effectiveness	Safety
Low Risk	0.46***	0.43***	0.47***
	(0.04)	(0.04)	(0.06)
High Risk	0.53***	0.57***	0.50***
	(0.03)	(0.04)	(0.04)
Observations	787	372	415
R-squared	0.85	0.89	0.81
F-test Low Risk=High Risk (p-value)	1.39 (0.2397)	3.22 (0.0765)	0.14 (0.7085)
Number of id	175	85	90
Robust standard errors clustered by participant, in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.			

[18] However, this is not a consistent result, as Table A-6 does not find the difference significant – neither does Column (8) of Table A-12.

The effects of framing are quite stark. Table 4 separates each of the groups by *Ascending* status. The first two columns involve rounds in which the certain risk initially equaled the low estimate of the ambiguous risk (*Ascending*=1). The last two columns, conversely, involve rounds in which the certain risk initially equaled the high estimate of the ambiguous risk (*Ascending*=0). There is a marked difference between the weights accorded to the low and high estimates in these two circumstances. When *Ascending* =1, both groups seem to be somewhat ambiguity-seeking, placing more weight on the low estimate of risk than on the high. In contrast, Columns (3) and (4) show that the weights associated with *Ascending*=0 are consistent with ambiguity-aversion. The low estimate is only weighted about 24%, as opposed to 76–77% for the high estimate.

Table 5. Fixed Effects Regressions for Certainty Equivalents, by Order.

	Ascending =1		Ascending = 0	
	Ineffectiveness	Safety	Ineffectiveness	Safety
Low Risk	0.59***	0.76***	0.24***	0.24***
	(0.05)	(0.05)	(0.05)	(0.05)
High Risk	0.37***	0.24***	0.77***	0.76***
	(0.04)	(0.03)	(0.04)	(0.04)
Observations	192	200	180	215
R-squared	0.95	0.94	0.96	0.95
F-test Low Risk=High Risk (p-value)	6.30 (0.0142)	37.67 (0.0000)	31.01 (0.0000)	39.07 (0.0000)
Number of id	78	82	79	87
Robust standard errors clustered by participant, in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.				

In order to formally test the effect of *Ascending*, Table 6 includes *Ascending* as a main effect and as interaction terms with both *Low Risk* and *High Risk*.

Table 6. Fixed Effects Regressions for Certainty Equivalents.

	Ineffectiveness	Safety
Low Risk	0.26***	0.22***
	(0.06)	(0.05)
High Risk	0.77***	0.79***
	(0.05)	(0.04)
Ascending*Low Risk	0.32***	0.53***
	(0.09)	(0.07)
Ascending*High Risk	-0.39***	-0.54***
	(0.07)	(0.05)
Ascending	17.79	21.91**
	(11.19)	(9.52)
Constant	-22.35**	-17.67**
	(8.67)	(8.82)
Observations	372	415
R-squared	0.95	0.94
Number of id	85	90
Robust standard errors clustered by participant, in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.		

The interactions between *Ascending* and the two risk estimates are both significant and the resulting weights are very similar to those in Table 5.

The interaction between *Ascending* and *Safety*/*Ineffectiveness* is also instructive. Recall that the point estimates in Table 5 Columns (1)–(2) seem to be consistent with stronger ambiguity-seeking for the *Safety* group than for the *Ineffectiveness* group, as *Safety* placed 76% weight on the low risk estimate, as opposed to 59% for *Ineffectiveness*. Whether or not this difference is actually significant is not tested in this table. There seems to be no significant difference between *Safety* and *Ineffectiveness* within *Ascending*=0, as the point estimates are almost identical.

Since individuals are assigned to *Safety* or *Ineffectiveness* alone, a fixed effect regression cannot estimate the effect of *Safety* versus *Ineffectiveness*.[19] Equation (5) addresses this by

[19] A random effects model is included in the Appendix, Table A-12. This controls for unobserved heterogeneity but does allow for estimation of a treatment effect. As shown, the results are very similar to the OLS results in Table 7.

running an OLS regression, controlling for demographics that might influence an individual's priors. This attempts to estimate Equation (3) while formally testing the effect of *Safety* versus *Ineffectiveness*. Table 7 lists the results for Equation (5). Column (1) lists the results for all rounds in which *Ascending*=1 while Column (2) lists the results for all rounds in which *Ascending*=0. Column (3) includes all observations.

Table 7. OLS Regressions for Certainty Equivalents.

	(1)	(2)	(3)
	Ascending =1 (Loss Domain)	**Ascending =0 (Gains Domain)**	**Total**
Safety	15.89**	6.74	3.60
	(7.64)	(10.08)	(9.86)
Ascending			20.81**
			(9.24)
Low Risk	0.60***	0.28***	0.28***
	(0.04)	(0.05)	(0.05)
High Risk	0.37***	0.76***	0.76***
	(0.03)	(0.04)	(0.04)
Low Risk * Safety	0.12**	-0.06	-0.06
	(0.06)	(0.07)	(0.07)
High Risk * Safety	-0.11**	0.02	0.03
	(0.04)	(0.06)	(0.06)
Ascending * Safety			3.05
			(12.03)
Low Risk * Ascending			0.30***
			(0.07)
High Risk * Ascending			-0.39***
			(0.06)
Low Risk * Safety * Ascending			0.20**
			(0.10)
High Risk * Safety * Ascending			-0.14*
			(0.08)
Experience with Headaches or Nausea	-0.45	18.72*	10.73*
	(8.58)	(9.81)	(6.39)
Risk Averse	8.80	12.51	10.90*
	(8.05)	(8.90)	(6.07)
Male	2.72	8.63	3.48
	(7.98)	(10.63)	(5.93)
Uses Drug Off-Label	15.65	-3.12	7.75
	(10.81)	(11.98)	(8.11)
Prescription Use	9.83	-11.19	-1.28
	(9.64)	(8.28)	(5.80)
Smoker	21.67***	-9.59	2.37
	(7.34)	(9.55)	(6.35)
Economics Major	8.53	-1.79	1.60
	(9.42)	(9.27)	(6.44)
Premed Major	19.68**	-17.63*	-0.88
	(9.07)	(9.65)	(6.26)
CRT	27.66***	-5.49	11.06*
	(9.64)	(8.86)	(5.79)
Observations	392	395	787
R-squared	0.93	0.95	0.94

Robust standard errors in parentheses, clustered by participant. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Other variables included but not reported are age, race, and political affiliation.

The results in Table 7 are very similar to those in Table 5. The weights reported in Columns (1)–(2) mirror those in Table 5. The interaction terms in Columns (1) and (2) test the significance of the difference between *Safety* and *Ineffectiveness* for *Ascending*=1 and Ascending=0, respectively. The interaction terms *Low Risk * Safety* and *High Risk * Safety* are significant for *Ascending*=1 but not for Column (2), indicating that the weights for *Safety* and *Ineffectiveness* are different in the loss domain but not the gains domain.

Column (3) includes all observations and includes interaction effects to differentiate observations with *Ascending*=0 and *Ascending*=1. The resulting weights are similar to the prior estimates. Those with *Ineffectiveness* =1, *Ascending*=0 place 28% weight on the low weight and 76% on the high weight. Those with *Safety*=1, *Ascending*=0 have roughly the same weight, as *Low Risk * Safety* and *High Risk * Safety* are insignificant. Those with *Ineffectiveness* =1,[20] *Ascending*=1 place 58% weight on the low estimate and 37% on the high estimate.[21] Finally, those with *Safety* =1, *Ascending* =1 place 78% weight on the low estimate and 23% on the high estimate. These weights line up closely with those reported in Table 5; Table 6 formally determines that the difference between *Safety* and *Ineffectiveness* is insignificant for *Ascending*=0 but significant for *Ascending*=1.

One possible concern is that participants were bored and just clicked through to get out of the rounds. This is unlikely to be an issue, given the incentive-compatibility device. In particular, there are only 47 occurrences of choosing indifference immediately and most participants who do this only do it a few times (9 did this one time, 4 did it twice, 5 do it 3 times, and 3 did this 5 times). However, if this is an issue, one way to deal with this is to drop strictly dominated choices: options in which 1) participants chose a gamble or were indifferent when the certainty

[20] *Ineffectiveness* is the omitted category in Table 7.
[21] This weight is computed as (.28 + .30 = .58) and (.76 - .39 = .37).

probability was lower than (or equal to) the lower estimate, 2) participants chose the certain probability or were indifferent when the certain probability is higher than (or equal to) than the highest estimate, and 3) whenever the certain probability has been pushed higher than the higher estimate or below the lower estimate. In total, these errors constituted around 4% of the observations and 9.3% of the certainty equivalents. The results dropping these errors are reported in the Appendix[22] and are largely similar to the original results, suggesting that this is not the driver of the results.

Another possibility is that participants in *Ascending*=0 were just uncomfortable by being at the high estimate of the ambiguous risk. They may have chosen indifference as a way to escape the scenario. Similarly, it is possible that participants were impatient and chose indifference earlier than they rationally would to get to new probabilities (resulting in ambiguity-seeking for *Ascending*=1 and ambiguity-aversion for *Ascending*=0). Again, the incentive-compatible device does alleviate these concerns. Additionally, as a robustness check, I only consider *Certainty Equivalents* that are triggered when a participant switches from one drug to another, rather than explicitly choosing indifference. Participants switching between drugs seem less likely to be trying to escape a scenario, as it would be easier to directly choose "I am indifferent." Only considering these choices in the fixed effects regressions produces qualitatively similar results, suggesting that these findings are not an artifact of trying to escape the scenario.

Other notable results include the strong effect of *CRT* and *Premed*. However, as noted above, these variables are included more as controls and will be discussed more in-depth in the following section.

[22] The tables include Table A-6–Table A-10.

B. Deviations from Neutrality

As previously stated, Table 4–Table 7[23] were useful in deriving ζ and ζ^*, corresponding to Equation (3). However, the interpretation of the effect of demographic variables on risk preferences is not straightforward in this context. The following section addresses this by focusing on the observed deviation from the ambiguity-neutral choice. Recall that *Deviation* is defined as $abs((\frac{Certainty\ Equivalent - Low\ Risk}{High\ Risk - Low\ Risk}) - 0.50)$, such that either ambiguity-averse or ambiguity-seeking behavior would result in a nonzero *Deviation*. This analysis will not distinguish between ambiguity-seeking and ambiguity-averse behavior but will see them both as deviations from ambiguity-neutral behavior. In this specification, variables with negative coefficients tend to result in more ambiguity-neutral decisions while variables with positive coefficients result in more deviation. The results for Equation (6) are listed in Table 8.

Interestingly, a wider range between the two risk estimates is negatively associated with deviations from the ambiguity neutral choice. This suggests that for a small enough range, participants might have seen the two estimates as essentially equal, displaying some nonlinearities in perceptions of risk. In contrast, higher mean values are positively associated with deviations. This means that when participants face higher risks of harm, they are more likely to exhibit deviations from ambiguity neutrality.

[23] ζ and ζ^* are calculated in Table 7 using a linear combination of several interaction terms.

Table 8. Deviation from Ambiguity Neutrality[a]

	(1)	(2)	(3)
	Ascending =1 (Loss Domain)	**Ascending =0 (Gains Domain)**	**Total**
Safety	0.02	0.01	0.02
	(0.02)	(0.02)	(0.02)
Ascending			0.01
			(0.01)
Range/100	-0.01***	-0.00	-0.01**
	(0.00)	(0.00)	(0.00)
Mean/100	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Experience with Headaches or Nausea	0.00	0.03	0.02
	(0.03)	(0.02)	(0.02)
Risk Averse	-0.01	-0.03	-0.02
	(0.03)	(0.02)	(0.02)
Male	-0.05**	0.02	-0.01
	(0.02)	(0.03)	(0.02)
Uses Drug Off-Label	-0.05*	0.02	-0.01
	(0.03)	(0.03)	(0.03)
Prescription Use	-0.07**	-0.03	-0.04**
	(0.03)	(0.02)	(0.02)
Smoker	-0.10***	-0.03	-0.06***
	(0.03)	(0.03)	(0.02)
Economics Major	-0.05	0.00	-0.02
	(0.03)	(0.03)	(0.02)
Premed Major	-0.06**	-0.02	-0.04
	(0.03)	(0.03)	(0.02)
CRT	-0.10***	-0.06**	-0.08***
	(0.03)	(0.03)	(0.02)
Constant	0.40	-0.05	0.17
	(0.28)	(0.23)	(0.17)
Observations	392	395	787
R-squared	0.27	0.20	0.19

Robust standard errors in parentheses, clustered by participant. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Other variables included but not reported are age, sex, race, and political affiliation.

[a]Deviation from Ambiguity Neutrality = ((Certainty Equivalent-Low Risk)/(High Risk-Low Risk))-.50.

The results in Table 8 seem to bolster the idea that consumers can be rational if properly motivated. First, variables dealing with general experience with prescription drugs or medical risk led participants to choose a more ambiguity-neutral indifference point. *Prescription Use* was negative in all specifications, and significant in *Ascending*=1 and *Total*. Similarly, *Premed Major* is negative when *Ascending*=1. Both of these situations might make participants more

comfortable assessing medical risk, either through personal experience or through their course of work.

Second, factors associated with rationality are associated with a more ambiguity neutral indifference choice. While *CRT* does not measure rationality, it does measure the propensity of participants to engaged in System 2 thinking. As noted above, System 2 thinking requires mental effort while System 1 thinking uses intuitive responses. Participants utilizing System 1 thinking might have been strongly influenced by a high or low estimate while participants utilizing System 2 thinking might have thought through the expected value of the gamble. The results support this theory: *CRT* is negative and significant, strongly associated with moving toward ambiguity neutrality.

These results are not without limitation. Since participants were faced with a chronic, non-life-threatening ailment, the results here may not generalize to a context in which there is a larger probability of death. However, the interest in this question was behavior in this mid-range of ailments.

C. Discussion

These findings suggest evidence consistent with the two hypotheses tested. First, there is evidence that ambiguity preferences vary with risk framing. Second, for *Ascending*=1, there is a significant difference in ambiguity attitudes between *Safety* and *Ineffectiveness*. The implications of each of these results are discussed below.

1. Risk Framing

The results are consistent with the first hypothesis: the framing of the risks significantly affects preferences over ambiguity. Specifically, when participants face an increasingly better certain option (*Ascending*=0), they are more likely to be ambiguity-averse in the options that they choose. In contrast, when participants face a worsening certain option (*Ascending*=1), they

are more likely to be ambiguity-seeking. If a patient previously consumed a well-performing drug but is offered a choice between a gamble and a drug with a higher certain risk of harm, she can be seen as being in the "loss" domain. Similarly, if a patient previously consumed a poorly performing drug and is offered a choice between a gamble and a drug with a lower certain risk of harm, she can be seen as being in the "gains" domain.

Consumers exhibit deviations from ambiguity neutrality in both domains. These deviations counteract each other in the aggregate, as shown in Table 4, but once stratified by *Ascending*, the deviations become clear: those in the loss domain exhibit ambiguity-seeking behavior while those in the gains domain exhibit ambiguity-averse behavior. The switch between ambiguity-aversion and ambiguity-seeking is consistent with the theory for large probability risks developed in Viscusi and Chesson (1999) and the results found in prior literature.

These findings have interesting implications for medical care. A patient's previous experience of care may cause them to deviate from ambiguity neutrality in predictable ways. If a patient had a good experience with her previous treatment, she is more likely to choose an ambiguously risky drug over a drug with a certain risk higher than her original (but not as high as the expected value of the risky drug). In contrast, if a patient had a poor experience with a previous treatment, she is more likely to choose the drug with the certain risk over an ambiguously risky drug, even though the certain risk is not as low as the expected value of the ambiguously risky drug.

The factors that temper these deviations from neutrality are important for policy purposes. First, familiarity with medical risks seems to result in more ambiguity neutral choices. *Prescription Use* is often significant in minimizing deviations, suggesting that true consumers might make more ambiguity-neutral decisions. Additionally, *Premed* is generally significant, indicating that personal or academic familiarity with drug risks help. This may also reflect a

general belief in the efficacy of medical care. Second, factors associated with rational thinking predictably minimize deviations. Participants who correctly answered all three CRT questions were more likely to choose a more ambiguity-neutral option. This is intuitive, since the CRT questions were designed the idea of distinguishing between System 1 and System 2 thinking (Frederick 2005).

It is possible that the reference-dependence effect is stronger in this context than in some real-life contexts, depending on whether consumers experience framing within the context of a single drug or across drugs. In the experiment, the "same drug" keeps updating its certain probabilities instead of patient choosing between two new drugs each round. This might be less applicable when consumers are moving from a good treatment to a worse (but separate) treatment (or vice versa). In contrast, these results might be more applicable when consumers experience framing in the context of a single drug. For example, a consumer may experience the loss domain if a drug that they are taking becomes less beneficial over time (e.g., due to patients building up a tolerance); a consumer may experience the gains domain if a drug becomes more beneficial over time (e.g., given adjustments to dosage). However, one can argue that consumers' reference points are simply indexed to level of care and are not segmented by treatment regime, making these results applicable in both contexts.

2. Safety vs. Ineffectiveness

The secondary result is that there is weak evidence suggesting that patients are slightly ambiguity-averse in *Ineffectiveness* but not necessarily in *Safety*. Moreover, the interaction between *Safety*/*Ineffectiveness* and framing is very instructive. Participants were more ambiguity-seeking when faced with *Safety* than *Ineffectiveness* in situations where *Ascending*=1. In situations where *Ascending*=0, this difference is insignificant.

It is unclear why the difference between *Safety* and *Ineffectiveness* is insignificant when *Ascending*=0: First, this may be an issue of insignificant power such that a true difference is unable to be detected. Alternatively, it is possible that there is no actual difference between *Safety* and *Ineffectiveness* when *Ascending*=0. This might mean that the ambiguity-aversion triggered by the gains framework actually overwhelms any initial difference in ambiguity preferences between *Safety* and *Ineffectiveness*. Given the results, I cannot distinguish between these explanations.

The results do suggest that when *Ascending*=1, in the loss domain, people are more ambiguity-seeking for *Safety* than for *Ineffectiveness*. This may suggest that facing *Safety* risks simply exaggerates responses to ambiguity triggered by the domain.[24] The underlying theory is intuitive: safety effects involve new ailments that are not inherent to the disease previously experienced. These new ailments are possibly more uncertain, intensifying the reaction to ambiguity. I do not, however, see the same reaction in the gains domain, as the *Safety* and *Ineffectiveness* weights are similar. Another explanation might be simply that the loss domain triggers ambiguity-seeking in both *Safety* and *Ineffectiveness* uniformly. Since participants in *Ineffectiveness* are initially slightly ambiguity-averse, the resulting effect is that they are less ambiguity-seeking than participants in *Safety*.

In sum, the experiment provides support for the two stated hypotheses. As predicted by the second hypothesis, there is weak evidence that participants are more ambiguity-averse for ineffectiveness risks than safety risks. Consistent with the first hypothesis, there are predictable

[24] These results are consistent with some post-hoc evidence on the difference between safety and ineffectiveness. Even though Bier and Connell (1994) did not design their study to compare ambiguity across safety and ineffectiveness, they note that an informal post-hoc analysis found that their ambiguity-framing effects were found mostly in scenarios involving side effects. They later note that this might be confounded with the higher probabilities associated with side effects, and they later rule out the thought that ambiguity effects were dependent on the division between safety and ineffectiveness.

reactions to ambiguity based on framing effects: participants exhibited ambiguity-seeking behavior in a "loss" domain and ambiguity-averse behavior in the "gains" domain. Moreover, the interaction between these two effects is intriguing: the results suggest that participants are significantly more ambiguity-seeking for safety risks than ineffectiveness risks in the loss domain.

VI. Conclusion

The purpose of this study was to measure consumers' ambiguity attitudes for risks that a drug will be unsafe or ineffective in the presence of framing effects. The results confirm that sensitivity to ambiguity is affected by risk framing, with participants exhibiting ambiguity-seeking behavior in the loss domain and ambiguity-averse behavior in the gains domain. Moreover, the results suggest that patients are more ambiguity-seeking for risks of safety than for risks of ineffectiveness in the loss domain.

These results do have practical implications for FDA policy, particularly for the regulatory regime governing off-label treatments. As stated above, off-label uses seem to trigger more ambiguous risks than most on-label uses. The results of this study suggest that if this is true, the predicted behavior of patients facing ambiguously risky off-label treatments depend strongly on context. If the patient taking the off-label treatment in a loss domain frame of mind (e.g., if their previous medication or condition was good relative to the available "certain" treatments), they might generally be more ambiguity-seeking; in fact, they might be more ambiguity-seeking with respect to safety, relative to ineffectiveness. Put in terms of the original example, if the drug that controls a patient's symptoms becomes unavailable and all other approved drugs have a relatively lower likelihood of being effective, the patient might be more likely to try an uncertain off-label treatment. The patient might be even more likely to try an uncertain off-label treatment if his previous drug produced very low side effects and all other

approved drugs have a relatively higher likelihood of producing side effects (as opposed to controlling symptoms). In contrast, if the patient is switching away from a drug he can no longer tolerate, he might favor approved treatments with marginally better results rather than pursuing ambiguous off-label treatments.

If consumers are not ambiguity neutral, they run the danger of consuming a suboptimal mix of drugs. They may not take a treatment with preferable, but ambiguous risks or may take overly risky treatments. Given that the direction of this reaction depends strongly on context (e.g., the experience with prior treatment), and that prior treatment is not immediately visible, it is difficult to develop a uniform plan for debiasing. Given these results, there might be value in reducing perceived ambiguity.

One solution might be to depend on physicians to synthesize ambiguous results and provide a medical recommendation to their patients. This would alleviate the bias as long as physicians are more ambiguity neutral in their recommendations to patients.[25] However, it is unclear how much better physicians are at handling ambiguity, though the *Premed* results seem optimistic. Another possibility is that the government or a private organization could assign an ambiguity-neutral rating on types of treatments. For example, for an off-label treatment with a mix of small case studies and anecdotes, the organization could form an ambiguity-neutral estimate of the demonstrated safety or efficacy. Some drug compendia, like DRUGDEX, do provide a coarse tier of "strength of the evidence" for off-label uses. Expanding this and linking it to an impartial organization might be a step forward.

Future work might concentrate on whether similar sensitivity toward ambiguity is observed when participants are making recommendations for someone else's treatment.

[25] Physicians summarizing results in a biased way would still reduce ambiguity, although it would introduce other distortions into the decision.

Additionally, it would be interesting to understand whether a "tier" synthesis of ambiguous results make consumers more ambiguity neutral. Either way, this experiment suggests that some intervention in synthesizing ambiguity on behalf of the consumer might improve patient care.

VII. References

Aberegg, Scott K., Edward F. Haponik, and Peter B. Terry. 2005. Omission Bias and Decision Making in Pulmonary and Critical Care Medicine. *CHEST Journal, 128* (3), 1497–1505.

Bier, Vicki M., and Brad L. Connell. 1994. Ambiguity Seeking in Multi-Attribute Decisions: Effects of Optimism and Message Framing. *Journal of Behavioral Decision Making, 7*(3), 169–182.

Centers for Disease Control and Prevention (CDC). 2014. *Health, United States, 2014*, *available at* http://www.cdc.gov/nchs/data/hus/hus14.pdf#085.

Coaster, Mariam, Baxter P. Rogers, Owen D. Jones, W. Kip Viscusi, Kristen L. Merkle, David H. Zald, and John C. Gore. 2011. Variables Influencing the Neural Correlates of Perceived Risk of Physical Harm. *Cognitive, Affective, & Behavioral Neuroscience, 11*(4), 494–507.

Connolly, Terry, and Jochen Reb. 2003. Omission Bias in Vaccination Decisions: Where's the Omission? Where's the Bias? *Organizational Behavior and Human Decision Processes, 91*(2), 186–202.

Curley, Shawn P., Stephen A. Eraker, and J. Frank Yates. 1984. An Investigation of Patient's Reactions to Therapeutic Uncertainty. *Medical Decision Making 4*(4), 501–511.

Curley, Shawn P., Mark J. Young, and J. Frank Yates 1989. Characterizing Physicians' Perceptions of Ambiguity. *Medical Decision Making, 9*(2), 116–124.

de Bekker-Grob, E. W., Marie-Louise Essink-Bot, Willem Jan Meerding, H. A. P. Pols, B. W. Koes, and E. W. Steyerberg. 2008.. Patients' Preferences for Osteoporosis Drug

Treatment: A Discrete Choice Experiment. *Osteoporosis International, 19*(7), 1029–1037.

Einhorn, Hillel J., and Robin M. Hogarth. 1985. Ambiguity and Uncertainty in Probabilistic Inference. *Psychological Review, 92*(4), 433.

Frederick, Shane. 2005. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Gerrity, Martha S., Robert F. DeVellis, and Jo Anne Earp. 1990. Physicians' Reactions to Uncertainty in Patient Care: A New Measure and New Insights. *Medical Care*, 724–736.

Han, Paul KJ, William MP Klein, Tom Lehman, Bill Killam, Holly Massett, and Andrew N. Freedman. 2011. Communication of Uncertainty Regarding Individualized Cancer Risk Estimates Effects and Influential Factors. *Medical Decision Making, 31*(2), 354–366.

Han, Paul KJ, William MP Klein, Thomas C. Lehman, Holly Massett, Simon C. Lee, and Andrew N. Freedman. 2009. Laypersons' Responses to the Communication of Uncertainty Regarding Cancer Risk Estimates. *Medical Decision Making*.

Han, Paul KJ, Sarah C. Kobrin, William MP Klein, William W. Davis, Michael Stefanek, and Steven H. Taplin. 2007. Perceived Ambiguity about Screening Mammography Recommendations: Association with Future Mammography Uptake and Perceptions. *Cancer Epidemiology Biomarkers & Prevention, 16*(3), 458–466.

Hogarth, Robin M. 1989. Ambiguity and Competitive Decision Making: Some Implications and Tests. *Annals of Operations Research, 19*(1), 29–50.

Johnson, F. Reed, Semra Özdemir, Carol Mansfield, Steven Hass, David W. Miller, Corey A. Siegel, and Bruce E. Sands. 2007. Crohn's Disease Patients' Risk-Benefit Preferences: Serious Adverse Event Risks versus treatment efficacy. *Gastroenterology, 133*(3), 769–779.

Junod, Suzanne White. 2014. *FDA and Clinical Trials: A Short History*, FOOD AND DRUG ADMINISTRATION, *available at* http://www.fda.gov/AboutFDA/WhatWeDo/History/Overviews/ucm304485.htm#_edn52

Kahn, Barbara E., and Rakesh K. Sarin. 1988. Modeling Ambiguity in Decisions under /uncertainty. *Journal of Consumer Research*, 265–272.

Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica, 47*(2), 263–91. doi:10.2307/1914185.

Kam, Cindy D., and Elizabeth N. Simas. 2010. Risk Orientations and Policy Frames. *The Journal of Politics,72*(2), 381–96. doi:10.1017/s0022381609990806.

Kees, Jeremy, Paula Fitzgerald Bone, John Kozup, and Pam Scholder Ellen. 2008. Barely or Fairly Balancing Drug Risks? Content and Format Effects in Direct-to-Consumer Online Prescription Drug Promotions. *Psychology & Marketing, 25*(7), 675–691.

Lloyd Andrew, Emma McIntosh, and Martin Price. 2005. The Importance of Drug Adverse Effects Compared with Seizure Control for People with Epilepsy: A Discrete Choice Experiment. *Pharmacoeconomics, 23*(11), 1167–1181.

McTaggart-Cowan, Helen M., Peilin Shi, J. Mark FitzGerald, Aslam H. Anis, Jacek A. Kopec, Tony R. Bai, Judith A. Soon, and Larry D. Lynd. 2008. An Evaluation of Patients' Willingness to Trade Symptom-Free Days for Asthma-Related Treatment Risks: A Discrete Choice Experiment. *Journal of Asthma, 45*(8), 630–638.

Radley, David C., Stan N. Finkelstein, and Randall S. Stafford. 2006. "Off-Label Prescribing Among Office-Based Physicians." *Archives of Internal Medicine* 166(9): 1021–1026, available at: http://archinte.jamanetwork.com.proxy.library.vanderbilt.edu/article.aspx?articleid=410250.

Ritov, Ilana, and Jonathan Baron. 1990. Reluctance to Vaccinate: Omission Bias and Ambiguity. *Journal of Behavioral Decision Making, 3*(4), 263–277.

Telser, Harry, and Peter Zweifel. 2002. Measuring Willingness-to-Pay for Risk Reduction: An Application of Conjoint Analysis. *Health Economics, 11*(2), 129–139.

Tversky, Amos, and Daniel Kahneman. 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics, 106*(4), 1039–61. http://www.jstor.org/stable/2937956.

Viscusi, W. Kip. 1997. Alarmist Decisions with Divergent Risk Information. *The Economic Journal, 107*(445), 1657–1670.

Viscusi, W. Kip. 1999. How Do Judges Think about Risk? *American Law and Economics Review, 1*(1), 26–62.

Viscusi, W. Kip and Harrell Chesson. 1999. Hopes and Fears: The Conflicting Effects of Risk Ambiguity. *Theory and Decision, 47*(2), 153–178.

Viscusi, W. Kip, and Wesley A. Magat. 1992. Bayesian decisions with ambiguous belief aversion. *Journal of Risk and Uncertainty, 5*(4), 371–387.

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber. 1987. An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks. *The RAND Journal of Economics*, 465–479.

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber. 1991. Pricing Environmental Health Risks: Survey Assessments of Risk–Risk and Risk–Dollar Trade-Offs for Chronic Bronchitis. *Journal of Environmental Economics and Management, 21*, 32–51 [Viscusi, Magat, and Huber 1991a].

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber. 1991. Communication of Ambiguous Risk Information. *Theory and Decision, 31*, 159–173 [Viscusi, Magat, and Huber 1991b].

Wittich, Christopher M., Christopher M. Burkle, and William L. Lanier. 2012. Ten Common Questions (And Their Answers) About Off-Label Drug Use. *Mayo Clinic Proceedings*, *87*, 982–83.

Zeckhauser, Richard J., and W. Kip Viscusi. 1990. *Risk Within Reason*. National Emergency Training Center.

VIII. Appendix:

A. Dropping the First Round

Table A-1. Fixed Effects Regression for Certainty Equivalents.

	All	Ineffectiveness	Safety
Low Risk	0.43***	0.42***	0.43***
	(0.05)	(0.05)	(0.07)
High Risk	0.55***	0.59***	0.52***
	(0.04)	(0.04)	(0.05)
Observations	612	287	325
R-squared	0.85	0.90	0.80
F-test Low Risk=High Risk (p-value)	2.16 (0.1434)	3.89 (0.0520)	0.55 (0.4586)
Number of id	175	85	90
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1			

Table A-2. Fixed Effects Regressions for Certainty Equivalents, by Order.

	Ascending =1		Ascending = 0	
	Ineffectiveness	Safety	Ineffectiveness	Safety
Low Risk	0.59***	0.85***	0.21***	0.22***
	(0.07)	(0.05)	(0.06)	(0.05)
High Risk	0.37***	0.17***	0.76***	0.77***
	(0.05)	(0.04)	(0.05)	(0.04)
Observations	143	153	144	172
R-squared	0.96	0.96	0.97	0.95
F-test Low Risk=High Risk (p-value)	4.11 (0.0464)	64.22 (0.0000)	28.48 (0.0000)	39.73 (0.0000)
Number of id	73	72	73	83
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1				

Table A-3. Fixed Effects Regression for Certainty Equivalents.

	Ineffectiveness	Safety
Low Risk	0.26***	0.20***
	(0.06)	(0.05)
High Risk	0.76***	0.81***
	(0.05)	(0.04)
Ascending*Low Risk	0.32***	0.60***
	(0.10)	(0.07)
Ascending*High Risk	-0.37***	-0.61***
	(0.08)	(0.06)
Ascending	2.19	15.92
	(14.71)	(10.76)
Constant	-13.95	-13.62
	(10.29)	(9.56)
Observations	287	325
R-squared	0.96	0.95
Number of id	85	90
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1		

Table A-4. OLS Regressions for Certainty Equivalents.

	(1)	(2)	(3)
	Ascending =1 (Loss Domain)	**Ascending =0 (Gains Domain)**	**Total**
Safety	22.86***	7.98	1.40
	(7.54)	(10.00)	(9.56)
Ascending			8.88
			(10.20)
Low Risk	0.59***	0.28***	0.28***
	(0.04)	(0.06)	(0.06)
High Risk	0.38***	0.75***	0.75***
	(0.03)	(0.04)	(0.04)
Low Risk * Safety	0.15**	-0.05	-0.06
	(0.07)	(0.07)	(0.07)
High Risk * Safety	-0.16***	0.03	0.04
	(0.05)	(0.06)	(0.06)
Ascending * Safety			15.06
			(12.91)
Low Risk * Ascending			0.30***
			(0.08)
High Risk * Ascending			-0.36***
			(0.06)
Low Risk * Safety * Ascending			0.23**
			(0.10)
High Risk * Safety * Ascending			-0.21**
			(0.08)
Experience with Headaches or Nausea	13.57	21.99*	21.50***
	(9.39)	(11.62)	(7.57)
Risk Averse	14.03	17.12	18.14**
	(9.32)	(10.62)	(7.08)
Male	7.40	19.62*	11.75*
	(8.03)	(11.51)	(6.52)
Uses Drug Off-Label	6.36	4.84	5.06
	(9.71)	(13.87)	(7.81)
Prescription Use	6.13	-14.09	-5.55
	(10.24)	(9.81)	(6.51)
Smoker	18.79**	-8.29	-1.86
	(8.19)	(12.04)	(7.75)
Economics Major	8.30	-12.66	-5.58
	(10.52)	(10.46)	(7.37)
Premed Major	19.40**	-25.50**	-4.79
	(9.81)	(11.38)	(6.89)
CRT	25.52**	-5.42	8.18
	(11.35)	(10.92)	(6.74)
Constant	62.15	-86.08	2.65
	(64.94)	(97.22)	(58.16)
Observations	296	316	612
R-squared	0.94	0.95	0.94

Robust standard errors in parentheses, clustered by participant. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Other variables included but not reported are age, race, and political affiliation.

Table A-5. Deviation from Ambiguity Neutrality. [a]

	(1)	(2)	(3)
	Ascending =1 (Loss Domain)	**Ascending =0 (Gains Domain)**	**Total**
Safety	0.03	0.03	0.03
	(0.02)	(0.02)	(0.02)
Ascending			0.01
			(0.01)
Range/100	-0.01***	-0.00	-0.01**
	(0.00)	(0.00)	(0.00)
Mean/100	0.01***	0.01***	0.01***
	(0.01)	(0.00)	(0.00)
Experience with Headaches or Nausea	-0.01	0.03	0.01
	(0.03)	(0.03)	(0.03)
Risk Averse	-0.01	-0.01	-0.01
	(0.03)	(0.03)	(0.02)
Male	-0.06*	0.03	-0.00
	(0.03)	(0.03)	(0.02)
Uses Drug Off-Label	-0.06*	0.03	-0.01
	(0.03)	(0.04)	(0.03)
Prescription Use	-0.07**	-0.03	-0.04
	(0.03)	(0.03)	(0.02)
Smoker	-0.12***	-0.03	-0.06**
	(0.03)	(0.03)	(0.02)
Economics Major	-0.04	-0.01	-0.03
	(0.03)	(0.03)	(0.02)
Premed Major	-0.07**	-0.04	-0.05**
	(0.03)	(0.03)	(0.03)
CRT	-0.10***	-0.07**	-0.08***
	(0.03)	(0.03)	(0.02)
Constant	0.48	0.08	0.23
	(0.29)	(0.21)	(0.17)
Observations	296	316	612
R-squared	0.27	0.21	0.19

Robust standard errors in parentheses, clustered by participant. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Other variables included but not reported are age, race, and political affiliation.

[a]Deviation from Ambiguity Neutrality =abs(((Certainty Equivalent-Low Risk)/(High Risk-Low Risk))-.50).

B. Dropping Errors

Table A-6. Fixed Effects Regression for Certainty Equivalents.

	All	Ineffectiveness	Safety
Low Risk	0.50***	0.45***	0.54***
	(0.04)	(0.06)	(0.05)
High Risk	0.50***	0.55***	0.45***
	(0.03)	(0.04)	(0.03)
Observations	714	345	369
R-squared	0.89	0.90	0.88
F-test Low Risk=High Risk (p-value)	0.01 (0.9181)	1.12 (0.2922)	1.43 (0.2357)
Number of id	174	84	90
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1			

Table A-7. Fixed Effects Regressions for Certainty Equivalents, by Order.

	Ascending =1		Ascending = 0	
	Ineffectiveness	Safety	Ineffectiveness	Safety
Low Risk	0.57***	0.75***	0.26***	0.35***
	(0.05)	(0.04)	(0.05)	(0.04)
High Risk	0.38***	0.26***	0.75***	0.67***
	(0.03)	(0.03)	(0.04)	(0.03)
Observations	176	177	169	192
R-squared	0.96	0.96	0.96	0.97
F-test Low Risk=High Risk (p-value)	6.35 (0.0138)	53.10 (0.0000)	27.34 (0.0000)	19.61 (0.0000)
Number of id	77	82	78	84
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1				

Table A-8. Fixed Effects Regression for Certainty Equivalents.

	Ineffectiveness	Safety
Low Risk	0.28***	0.33***
	(0.06)	(0.04)
High Risk	0.74***	0.69***
	(0.04)	(0.04)
Ascending*Low Risk	0.29***	0.38***
	(0.07)	(0.06)
Ascending*High Risk	-0.36***	-0.42***
	(0.05)	(0.05)
Ascending	15.42	17.77*
	(9.83)	(9.86)
Constant	-17.73**	-11.26
	(8.57)	(8.05)
Observations	345	369
R-squared	0.96	0.95
Number of id	84	90
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1		

Table A-9. OLS Regressions for Certainty Equivalents.

	(1)	(2)	(3)
	Ascending =1 (Loss Domain)	**Ascending =0 (Gains Domain)**	**Total**
Safety	13.84*	10.92	9.30
	(7.96)	(9.31)	(9.10)
Ascending			17.21**
			(8.18)
Low Risk	0.60***	0.30***	0.30***
	(0.04)	(0.05)	(0.06)
High Risk	0.38***	0.73***	0.74***
	(0.03)	(0.04)	(0.04)
Low Risk * Safety	0.09*	0.03	0.03
	(0.05)	(0.07)	(0.07)
High Risk * Safety	-0.09**	-0.05	-0.04
	(0.04)	(0.05)	(0.05)
Ascending * Safety			-2.13
			(11.03)
Low Risk * Ascending			0.29***
			(0.07)
High Risk * Ascending			-0.35***
			(0.05)
Low Risk * Safety * Ascending			0.08
			(0.09)
High Risk * Safety * Ascending			-0.05
			(0.07)
Experience with Headaches or Nausea	-10.50	22.15***	7.75
	(8.39)	(8.43)	(5.99)
Risk Averse	11.61	9.18	11.39**
	(7.21)	(8.03)	(5.66)
Male	5.48	1.57	2.42
	(8.00)	(9.01)	(5.41)
Uses Drug Off-Label	12.17	-5.48	5.29
	(10.19)	(10.92)	(7.32)
Prescription Use	18.29*	-3.78	5.39
	(9.38)	(7.23)	(5.48)
Smoker	17.44**	-5.96	2.47
	(7.07)	(9.09)	(5.83)
Economics Major	3.98	-2.55	-0.76
	(7.87)	(7.97)	(5.33)
Premed Major	12.17	-6.74	-0.37
	(8.20)	(9.19)	(5.90)
CRT	21.85**	-10.07	7.05
	(9.20)	(6.16)	(4.85)
Constant	10.02	89.90	39.25
	(73.41)	(76.49)	(43.65)
Observations	353	361	714
R-squared	0.95	0.96	0.95

Robust standard errors in parentheses, clustered by participant. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Other variables included but not reported are age, race, and political affiliation.

Table A-10. Deviation from Ambiguity Neutrality. [a]

	(1)	(2)	(3)
	Ascending =1 (Loss Domain)	**Ascending =0 (Gains Domain)**	**Total**
Safety	0.02	0.00	0.01
	(0.02)	(0.02)	(0.02)
Ascending			-0.00
			(0.01)
Range/100	-0.01**	-0.01*	-0.01***
	(0.00)	(0.00)	(0.00)
Mean/100	0.01***	0.02***	0.01***
	(0.00)	(0.00)	(0.00)
Experience with Headaches or Nausea	0.03	0.03	0.03*
	(0.02)	(0.02)	(0.02)
Risk Averse	-0.03	-0.03	-0.03*
	(0.02)	(0.02)	(0.02)
Male	-0.05**	-0.02	-0.03
	(0.02)	(0.03)	(0.02)
Uses Drug Off-Label	-0.01	0.03	0.01
	(0.03)	(0.03)	(0.02)
Prescription Use	-0.06***	-0.01	-0.03*
	(0.02)	(0.02)	(0.02)
Smoker	-0.06**	-0.01	-0.03
	(0.02)	(0.03)	(0.02)
Economics Major	-0.03	-0.01	-0.02
	(0.02)	(0.02)	(0.02)
Premed Major	-0.02	-0.01	-0.01
	(0.02)	(0.03)	(0.02)
CRT	-0.07***	-0.06***	-0.07***
	(0.02)	(0.02)	(0.02)
Constant	0.21	0.19	0.18
	(0.29)	(0.17)	(0.18)
Observations	353	361	714
R-squared	0.28	0.20	0.21

Robust standard errors in parentheses, clustered by participant. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Other variables included but not reported are age, race, and political affiliation.

[a]Deviation from Ambiguity Neutrality =abs(((Certainty Equivalent-Low Risk)/(High Risk-Low Risk))-.50).

C. Full Results

Table A-11. Full Results for Table 8: Deviation from Ambiguity Neutrality.

	(3)	(1)	(2)
	Ascending =1 (Loss Domain)	Ascending =0 (Gains Domain)	Total
Safety	0.02	0.01	0.02
	(0.02)	(0.02)	(0.02)
Ascending			0.01
			(0.01)
Range/100	-0.01***	-0.00	-0.01**
	(0.00)	(0.00)	(0.00)
Mean/100	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
Age	-0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)
Experience with Headaches or Nausea	0.00	0.03	0.02
	(0.03)	(0.02)	(0.02)
Risk Averse	-0.01	-0.03	-0.02
	(0.03)	(0.02)	(0.02)
Race			
African American	0.17***	0.05	0.12**
	(0.05)	(0.05)	(0.05)
Asian	0.09***	0.07**	0.08***
	(0.03)	(0.03)	(0.02)
Hispanic	-0.03	0.07	0.02
	(0.03)	(0.04)	(0.03)
Other	0.07	0.08*	0.06*
	(0.05)	(0.05)	(0.03)
Male	-0.05**	0.02	-0.01
	(0.02)	(0.03)	(0.02)
Political			
Democrat	-0.01	0.11***	0.06**
	(0.03)	(0.03)	(0.03)
Independent	-0.03	-0.00	-0.01
	(0.03)	(0.03)	(0.03)
Other	0.09**	0.03	0.06*
	(0.04)	(0.05)	(0.04)
Uses Drug Off-Label	-0.05*	0.02	-0.01
	(0.03)	(0.03)	(0.03)
Uses Prescription Drug	-0.07**	-0.03	-0.04**
	(0.03)	(0.02)	(0.02)
Smoker	-0.10***	-0.03	-0.06***
	(0.03)	(0.03)	(0.02)
Economics Major	-0.05	0.00	-0.02
	(0.03)	(0.03)	(0.02)

Premed Major	-0.06**	-0.02	-0.04
	(0.03)	(0.03)	(0.02)
CRT	-0.10***	-0.06**	-0.08***
	(0.03)	(0.03)	(0.02)
Constant	0.40	-0.05	0.17
	(0.28)	(0.23)	(0.17)
Observations	392	395	787
R-squared	0.27	0.20	0.19
Robust standard errors in parentheses, clustered by individual *** p<0.01, ** p<0.05, * p<0.1.			

D. Random Effects for the Full Sample

Table A-12. Random Effects Regression for Certainty Equivalents.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ascending=1		**Ascending=0**					
	Ineffectiveness	**Safety**	**Ineffectiveness**	**Safety**	**Ineffectiveness**	**Safety**	**All**	**All**
Safety							4.88	11.26
							(9.33)	(8.05)
Ascending					19.95**	22.73***	19.99**	
					(8.92)	(7.75)	(8.88)	
Low Risk	0.58***	0.73***	0.27***	0.23***	0.28***	0.23***	0.28***	0.44***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)	(0.04)
High Risk	0.37***	0.26***	0.76***	0.78***	0.76***	0.78***	0.76***	0.55***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Ascending*Low Risk					0.31***	0.50***	0.31***	
					(0.08)	(0.06)	(0.08)	
Ascending*High Risk					-0.39***	-0.53***	-0.39***	
					(0.06)	(0.05)	(0.06)	
Safety*Low Risk							-0.05	0.01
							(0.07)	(0.07)
Safety*High Risk							0.02	-0.04
							(0.06)	(0.05)
Ascending*Safety							2.73	
							(11.78)	
Ascending*Safety*Low Risk							0.19*	
							(0.10)	
Ascending*Safety*High Risk							-0.14*	
							(0.08)	
Observations	192	200	180	215	372	415	787	787
Number of id	78	82	79	87	85	90	175	175
Robust standard errors clustered by participant in parentheses. *** p<0.01, ** p<0.05, * p<0.1								

CHAPTER 2

EMPIRICAL EVIDENCE OF RISK PENALTIES FOR NTI DRUGS

I. Introduction

New drug advances can often outpace regulatory processes designed to control drug risks. The FDA sometimes allows a drug to go through an accelerated approval process for the sake of increasing access to the drug. These abbreviated processes might not fully reveal the risk that a drug is unsafe or ineffective, resulting in higher or more uncertain risks. In such situations, physicians' ability to assess, and the market's sensitivity to, risk is particularly important.

In order to incentivize generic manufacturers to apply for approval, the FDA allows follow-on products to go through an abbreviated new drug approval (ANDA). 21 U.S.C. 355(j). These generic products face less stringent approval standards: manufacturers must only prove bioequivalence to a "reference-listed" drug rather than safety and effectiveness of their own drug.[26] Bioequivalence does not require exact replication of the reference-listed drug. Instead, healthy subjects receive the reference drug and the test drug and two one-sided tests are conducted, measuring if the test drug is significantly less bioavailable than the reference and vice versa. Bioequivalence is satisfied if there is no more than a "difference of 20% in the area under the curve (AUC) and peak concentration (Cmax) between the reference and test drugs" (Singh et al. 2014).

Some are doubtful that this abbreviated approval process can adequately control risk, particularly for a certain group of drugs called narrow therapeutic index (NTI) drugs. The FDA's proposed definition of NTI drugs are "those drugs where small differences in dose or blood

[26] In contrast, reference drugs have to provide " 'substantial evidence' that the drug 'will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling' " (Junod 2014).

concentration may lead to dose and blood concentration dependent, serious therapeutic failures or adverse drug reactions" (Yu 2011). Given the narrow range in which the drug provides therapeutic benefit and the allowed difference between products that are technically "bioequivalent," changes from a brand-name to a generic or from a generic to another generic can result in adverse events.[27] This perceived heterogeneity in risk was not previously regulated by the FDA, as the FDA treated all generic drugs as equivalent to brand-name drugs. However, the FDA has recently acknowledged the increased risk associated with NTI drugs. In 2010, the FDA Advisory Committee for Pharmaceutical Science proposed instituting stricter regulations for NTI bioequivalence (Yu 2011). More recently, scientists at the FDA have published new bioequivalence limits for NTI drugs. (Yu et al. 2015).

This Chapter discusses narrow therapeutic index drugs as an example in which there was a perceived, yet unregulated, risk. There are several similar issues spurred by medical innovations: critics have suggested that traditional generic approval processes may be insufficient to evaluate generic versions of nonbiological complex drugs (NBCD), making the risks associated with them more uncertain (Gottlieb 2014). The FDA recently tried to approve generic versions of nonbiological complex drugs but found it difficult to verify therapeutic equivalence. While the FDA approved a generic version of IV iron medication in 2011, in 2013 it requested an independent study to verify that the generic version was safe and effective (Gottlieb 2014). Some criticized the FDA because the bioequivalence process was insufficient to guarantee safety and effectiveness upon approval and called for the FDA to develop a better approval process for that type of drug. Given that the FDA will continue to face issues in which

[27] Crawford et al. (2006) briefly mention the additional risk from switching from one generic to another.

there is an unregulated but perceived risk, this Chapter examines NTI drugs as an example, as they elicit similar issues and have more data on market experience.

This Chapter uses NTI drugs as a case study to examine whether, in the absence of regulation, the market is sensitive to this risk. Assuming that consumers are less willing to pay for riskier products, does this preference translate into a price penalty for riskier drugs? If so, abbreviated approval processes might provide broader access to drugs without imposing uncompensated risks. If instead the market is insensitive to this risk, the FDA might be justified in imposing stronger regulation.

The Chapter observes a price penalty for NTI drugs. Consistent with consumer preferences when switching is costly, the Chapter also finds that the gap between brand-name and generic drugs is smaller for NTI drugs than for non-NTI drugs. Finally, this Chapter examines consumption behavior and finds that consumers make purchasing decisions that are consistent with awareness of this risk. The results of this Chapter have ramifications beyond the context of NTI drugs and might inform by analogy the appropriateness of requiring extra FDA regulation on generic versions of other types of products.

The rest of the Chapter proceeds as follows: Section II explores the current literature on narrow therapeutic index substitution and Section III outlines the conceptual model underlying consumer choice. Section IV describes the data, and Section V discusses the results. Section VI concludes.

II. Exploring the Perceived Risks of NTI Drugs

The issue of narrow therapeutic index drugs is nested in a larger discussion about generic substitution. In general, generic versions of brand-name drugs are thought to be therapeutically

equivalent.[28] Generic substitution is also often associated with increased cost savings. Fischer and Avorn (2003) estimate that more generic substitution could save Medicaid at least $229 million. Similarly, Haas et al. (2005) use the Medical Expenditure Panel Survey (MEPS) to estimate that generic substitution would produce potential savings of $46 per adult below the age of 65 and $78 per adult above the age of 65. From a policy perspective, Shrank et al. (2010) find that the rates of generic substitution for states requiring patient consent for generic substitution were 25% lower than those not requiring patient consent; they note that increased substitution could lead to lower costs for state Medicaid programs. Despite the general consensus being that generic substitution should save money for patients and third-party payers, a few studies have suggested that this might not always hold true. In a recent meta-analysis, 64% of the economic comparisons reviewed found that generic substitution actually increased costs (Gothe et al. 2015). Particularly, they conclude that generic substitution might not decrease costs for drug categories such as antiepileptics or immunosuppressants. In a review of studies on antiepileptic drugs, Duh et al. (2009) find similar results. They note that many studies attribute this increase in costs to the need for other medical attention (such as higher rates of hospitalization and longer hospital stays) while taking generic versions of antiepileptic drugs.

Even if generic substitution were financially preferable, patients have traditionally harbored distrust of generic drugs. There have been several studies outlining demographic and sociological reasons patients perceive generic drugs to be riskier than their brand-name counterparts. For example, Omojasola et al. (2012) find that minorities were more likely to think that generic drugs were associated with more side effects; relatedly, Shrank et al. (2007) finds that patients living in the poorest zip codes were less likely to initially receive generic drugs.

[28] For a history and review of bioequivalence standards, see Midha and McKay (2009). Other studies discussing bioequivalence includes Al-Jazairi et al. (2008).

Despite this distrust, particular features of the health care system can encourage generic substitution. In a meta-analysis conducted by Dunne and Dunne (2015), a key factor in accepting generic substitution is a patient's trust in his physician. Additionally, factors such as formulary structure do affect drug choice. Huskamp et al. (2003) examine a change in formulary structure for two companies and find that enrollees do respond to financial incentives to choose certain drugs. Similarly, Shrank et al. (2007) find that while pharmacy benefit design and pharmacy type do not affect the likelihood of initiating generic medication, enrollment in tiered pharmacy benefit plans and use of mail-order pharmacies do increase this likelihood.

The generic substitution of NTI drugs fits uneasily within this larger context. There are real concerns regarding the therapeutic equivalence of generic NTI drugs to their brand-name version. Much has been written regarding the clinical equivalence of generic and brand-name versions of NTI drugs. Kesselheim et al. (2008) review some of these studies for cardiovascular disease and do not find evidence that brand-name drugs are superior to their generic versions. In contrast, Blix et al. (2010) find that NTI drugs are associated with more drug-related problems (e.g., non-optimal doses, drug interactions, and need for monitoring) than non-NTI drugs. Others have argued that more care should be taken with NTI drugs, such as prohibiting autosubstitution and adjusting bioequivalence standards (Hottinger and Liang 2012).[29] Physicians' and patients' perspectives on generic substitution of NTI drugs seem similarly mixed. In a study of antiepileptic drugs, which are often considered NTI drugs, Berg et al. (2008) find that 66% of

[29] There is a lot of research on the appropriateness of antiepileptic drugs, as many of the drugs within the class are associated with a narrower therapeutic index. For more information, see Heaney and Sander (2007). Similarly, the appropriateness of bioequivalents has been discussed with respect to antiepileptics (Heaney and Sander 2007) and drugs treating mental illness (Fairfax-Columbo and DeMatteo 2015).

physicians and 34% of patients surveyed felt that the occurrence of seizures was linked to generic substitution.[30]

This Chapter is not concerned with the actual therapeutic equivalence of generic NTI drugs but rather with the perception of risk associated with NTI use. It is sufficient that either patients, or agents making decisions on behalf of patients, are uncertain of or negatively perceive the effects of substitution within NTI drugs. However, it is notable that although the FDA originally claimed that generic substitution of NTI drugs pose no additional risk, in 2010, the FDA Advisory Committee for Pharmaceutical Science decided that the current bioequivalence standards are not sufficiently strict for NTI drugs (Yu 2011). FDA scientists now suggest a "four-way, crossover, full-replicated study design" which will compare differences in means and within-subject variation, the latter of which had not been addressed previously.

The prior literature on NTI drugs mostly concerns determinants of consuming generic versions. Chao et al. (2002) use 1996–1998 MEPS data to examine the likelihood of consuming generic NTI drugs and find that NTI drugs are more likely to be prescribed as brand-name drugs than non-NTI drugs. They then calculate comparable discount rates for NTI and non-NTI drugs and predict the savings possible if further generic substitution were implemented with NTI drugs. In particular, Chao et al. (2002) find that discount rates were positively associated with use of drugs, more for NTI drugs than non-NTI drugs. They find no difference in switch rates between NTI and non-NTI drugs. Similarly, Gagne et al. (2013) analyze predictors of choosing generic NTI drugs for elderly Medicare participants, focusing on demographics and prior generic use.

[30] Bhosle et al. (2005) finds that patients filling prescriptions were unaware of the term "narrow therapeutic index" and found physicians' assurances more reassuring than pharmacists'.

The price consequences of the lack of therapeutic equivalence have also been studied previously from a generic entry perspective. Nabin et al. (2012) develop a model attributing the "generic competition paradox,"[31] the phenomenon in which brand-name companies raise their prices following generic entry, to lack of therapeutic equivalence. They estimate their model using Canadian data from 2000–2011 and find that drug classes with lower therapeutic equivalence (antiepileptics) were more likely to have higher post-patent prices than those with higher therapeutic equivalence (cardiovascular drugs).

This Chapter contributes to this literature by looking at sensitivity to NTI risk in both price and consumption behavior. Unlike Nabin et al. (2012), this Chapter examines NTI drugs from a hedonic price perspective—it looks at the compensating differential for the additional risk in NTI drugs. Like Chao et al. (2002), it looks at the gap between brand and generic prices between NTI and non-NTI drugs; however, it additionally explores whether there is a risk penalty associated with NTI drugs, despite the multiple payers and divergent incentives prevalent in the drug market. Finally, it also looks at consumption behavior to see whether consumers behave in a way that minimizes their risk. The next section provides a simple model of the general drug choice decision within the current context.

[31] The generic competition paradox is an empirical observation that has puzzled scholars. Some explanations for the increase in brand-name price following generic entry is market segmentation based on brand loyalty and insurance coverage. Nabin et al. (2012) instead attribute this to therapeutic inequivalence; consumers do not see generic versions as perfect substitutes and prefer the brand version. This concept is only tangentially related to this Chapter, as this Chapter focuses on consumer/agent valuations of a risky drug, not in estimating supply and demand functions. Additionally, while this Chapter does control for competition, this Chapter does not attempt to model behavior directly following initial generic entry. Finally, this Chapter allows for therapeutic inequivalence between generic versions of NTI drugs, while Nabin et al. (2012) seems to concentrate on the difference between brand-name and generic versions.

III. Model

This Chapter examines whether the perceived higher risks associated with NTI drugs result in a price penalty for NTI drugs. In general, it is assumed that consumers are willing to pay less for riskier products. However, it is unclear whether this preference will result in a price penalty in the context of pharmaceutical drug consumption.

Pharmaceutical drug consumption is a study in crossed incentives, asymmetric information, and limited attention. Expert physicians prescribe "appropriate" drugs to their uninformed patients. The patients experience any side effects or lack of efficacy associated with the drug. Third-party payers bear most of the financial burden; they use pharmacy benefit managers (PBMs) to negotiate drug prices with pharmaceutical manufacturers (Fox 2003)[32] and cover much of the up-front costs. Consumers generally only see a fraction of the "price" of the drug. While patients do pay for drugs, this is generally through premiums for their insurance coverage, rather than a direct payment associated with a particular purchase.

For these reasons, it is unclear whether the risks associated with NTI drugs will actually lead to price penalties. Specifically, third-party payers face the majority of the drug price but do not experience the adverse events of the drug. However, if third-party payers are cognizant of the risks and sensitive to the preferences of their consumers, [33] I expect third-party payers would demand a price penalty for risk. The study utilizes the following hedonic price model:

$$(1)\ \ln(P+1) = NTI'\beta_1 + \ Generic'\beta_2 + \ Generic^{*}NTI'\beta_3 + Z'\gamma + \varepsilon,$$

[32] Insofar as prices are based on cross-subsidization of other drugs, I cannot see this. However, this only produces biased results if cross-subsidization occurs between NTI and non-NTI drugs.
[33] Fox (2003) describes how PBMs work and set formularies. Fox notes that PBMs may treat NTI drugs as having no generic equivalents, decide on a drug-by-drug basis, or treat NTI drugs similarly to non-NTI drugs.

where P signifies the total cost of the drug or the portion of the cost that each third-party (or consumer) pays.[34] The vector Z includes data on market characteristics that determine the price of a drug, such as drug age and competition. *NTI* is an indicator variable for narrow therapeutic index drugs, *Generic* is an indicator for generic status, and *Generic*NTI* signifies when the drug is a generic NTI drug.

If consumers care about the extra risk associated with narrow therapeutic index drugs, I expect $\beta_1 < 0$. Here, the nature of the risk matters. There are two possible risks associated with NTI drugs, both stemming from the lack of substitutability between versions[35] of NTI drugs. The first type of risk might be called a *direct risk*: consumers might have different therapeutic reactions to different versions of the same NTI drug. The expected harm associated with a new version of the same NTI drug is nonzero, even conditional on the patient's experience of the current version. In a sense, the costs of switching are higher for NTI drugs. The second type of risk might be called an *indirect risk*: this is the risk arising from anticipated forced switching. A patient may perceive that he will eventually be forced to switch between versions of the NTI drug and thus face the "direct risk" associated with switching between NTI versions. Forced switching may occur due to drug shortages or to certain manufacturers exiting the market. If consumers (or their physicians) are forward-looking enough to anticipate such switching, NTI drugs entail an inherent risk, making the category as a whole less desirable. Thus, the "indirect

[34] The results are robust to only considering any nonzero prices (and not limiting observations to those expecting to be covered by a particular payer) and using $ln(price)$. These runs are listed in Table A-3 through Table A-6.

[35] For clarity, in this paper, a *drug class* (e.g., anticonvulsants) contains many *substances (e.g., carbamazepine)*. *Substance* categories contain both generic and brand-name versions of the drug (e.g., both carbamazepine and tegretol belong to the substance category carbamazepine), such that they correspond to the nonproprietary name. *Versions* of the drug involve unique product codes within each drug category. Versions are an even smaller category than brand-name/generic, as there are multiple generic versions generally available. For the sake of consistency, the drug hierarchy is as follows: drug class, substance, and drug version.

risk" is the anticipated future risk associated with choosing any NTI drug. Both the direct and indirect risks arise from underlying issues of substitution and are consistent with $\beta_1 < 0$.

The interaction term *Generic*NTI* is of particular interest. If consumers view generic drugs as simply riskier than brand-name NTI drugs, $\beta_3 < 0$. If, instead, the main risk associated with NTI drugs arises from substituting between versions, I expect switching to be costly. With costly switching, consumers would require a larger price difference in order to switch NTI versions, resulting in $\beta_3 > 0$.

To demonstrate this, consider the following example. Utility from consuming drug j can be formalized as

$$(2)\ U_{ij} = X_i'\beta + \gamma_1(V_i - P_j) + W_j'\delta + \gamma_2 I_j + \gamma_3 D_j + \varepsilon_i$$

where X_i is a vector of characteristics for patient i, V_i represents the patient's nonmedical consumption, and P_j is the price of drug j. W_j is a vector of drug version-specific variables, [36] while I_j and D_j represent drug risks. D_j is the direct risk and I_j is the indirect risk. Particularly, D_j is the expected harm caused by version j. If version j causes an adverse event, $D_j \neq 0$, and if it does not, $D_j = 0$. Since *I* and *D* are both risks and consumers generally prefer less risky drugs, $\gamma_2, \gamma_3 < 0$. A consumer should consume drug k when

$$(3)\ U_{ik} > U_{ij}\ for\ all\ j \cong k$$

Then, from equation (2), a patient will choose drug k when

$$(4)\ (P_j - P_k)\gamma_1 + (W_k - W_j)\delta + (I_k - I_j)\gamma_2 + (D_k - D_j)\gamma_3 > 0.$$

It is helpful to contrast the choice between two non-NTI drug versions and the choice between two NTI drug versions. Suppose drug j is a generic version of a non-NTI drug and drug

[36] If one of the versions of the drug is the brand version, W_j can contain aspects like brand loyalty.

k is another generic version of the same non-NTI drug. Since both drugs are generic versions, W_j should equal W_k.[37] Similarly, since both drugs are non-NTI, $I_j = I_k = 0$, and $D_j = D_k = 0$, a consumer will choose the cheaper version.

Instead, suppose drug j is a generic version of an NTI drug and drug k is another generic version of the same NTI drug. Since both drugs are generic, W_j should equal W_k. Since both versions are within the NTI category, $I_j = I_k \neq 0$; however, D_j does not necessarily equal D_k. If the consumer currently consumes drug j, she knows whether drug j results in an adverse reaction ($D_j \neq 0$) or not ($D_j = 0$). She does not know how she will react to drug k (whether $D_k = 0$ or $D_k \neq 0$). Thus, in order to switch to drug k,

(5) $(P_j - P_k)\gamma_1 > (D_j - D_k)\gamma_3$.

Since $\gamma_3 < 0$, if $D_j = 0$, and $D_k \neq 0$, a larger price gap is necessary to switch from drug j to k. The larger the expectation of experiencing adverse events from any given version, and the smaller the current chance of experiencing adverse events, the larger the "loyalty" to a particular generic NTI version. If this is true, I expect $\beta_3 > 0$.

The Chapter explores whether the results predicted by this conceptual model are consistent with market data. There are many reasons why they may not be, including the issues of intervening actors and asymmetric information inherent in the pharmaceutical drug context. The following section introduces the data used to evaluate this question.

IV. Data

This Chapter uses data from the Medical Expenditure Panel Survey (MEPS) 2008–2010.[38] MEPS is published by the U.S. Department of Health and Human Services and surveys

[37] W_j does not need to equal W_k for these results to hold. All that is necessary is that the relationship between W_j and W_k does not vary by NTI status.

families and individuals, their medical providers, and employers about medical expenditures. The Household Component of the MEPS is drawn from a nationally representative subsample of households participating in the prior year's National Health Interview Survey (NHIS) and provides information on demographic characteristics, health conditions, insurance, medical expenditures, and sources of payment. Participants are interviewed over a span of two full calendar years. The prescription drug expenditure data from MEPS is the sum of actual payments by various sources, rather than "charges," and does incorporate discounts.

While MEPS lists numerous different payers, this analysis focuses on the following payers: private insurance, Medicare, self-payment, and total payment. [39] Because of the increasing gap between listed charges and actual payments, MEPS reports the payment by each payer rather than the charge by each. While these payments incorporate discounts received, they do not incorporate rebates received by Medicaid or other buyers. Since rebates are a large part of Medicaid's pricing regime, I do not focus on the results for Medicaid payments, although they are listed for comparison.

Since there is no uniform FDA definition of a narrow therapeutic index drug, this Chapter uses the definition used by the North Carolina Board of Pharmacy as of January 27, 2009, in keeping with prior literature (Gagne et al. 2013).[40] The following NTI drugs are studied in this analysis: warfarin, theophylline, digoxin, carbamazepine, phenytoin, lithium, levothyroxine,

---

[38] The years 2008–2010 were chosen because Medicare Part D was implemented in 2006, making Medicare drug data likely to be a Part D plan. Additionally, in 2007, the rules used to identify outlier prices for prescription medication changed, allowing a larger range of drug prices.
[39] These are not the only payers reported by the MEPS, just the most interesting. MEPS includes Veterans Administration payments, workers' compensation, state and local, Tricare, and a series of "other" payment categories (including a set of residual payment categories ("other public" and "other private") meant to correct inconsistencies in the data). I do not consider these categories in the analysis except insofar as they contribute to the total sum of payments.
[40] The North Carolina Board of Pharmacy lists procainamide hydrochloride, but my data do not contain any observations involving this drug.

cyclosporine, tacrolimus, and ethosuximide. To preserve the proper comparison, this study retains all prescription drugs within the same Multum Therapeutic sub-class as each NTI drug.[41]

MEPS data are merged with data from the FDA by National Drug Code (NDC), in order to incorporate market status of each drug (brand-name drug or generic drug).[42] Using data from the NDC Database file,[43] I match approval information to MEPS data using the NDC code.[44]

Since price is sensitive to the number of competitors (Reiffen and Ward 2005; Frank and Salkever 1997), I include a measure of total competition. In order to get data on competition, I use information from the National Drug Code Directory. I look at unique observations of labelers and substance names and then count the number of labelers per substance name.

Drug "age" also contributes to drug price, so I measure the age of the drug by subtracting the year that the manufacturer started marketing the drug from the year of its usage.[45] For multiple entries on the same drug application, I use the earliest date to capture the date of the biggest innovation, not minor later adjustments.[46] I similarly control for the age of the substance

---

[41] The following Multum Lexicon Therapeutic sub-classes are retained: antiarrhythmic agents, anticonvulsants, anticoagulants, thyroid drugs, immunosuppressive agents, bronchodilators, and antipsychotics. Cyclosporine and tacrolimus are only considered as immunosuppressive agents. While cyclosporine and tacrolimus are used as an eye drop (ophthalmic preparations) and topical ointment (dermatological agents), respectively, I do not consider these drug classes, as the drugs are not consumed orally and the uses seem distinct from the rest of the classes. Adding these classes back in does not seem to make much difference, however.

[42] NDA authorized generics are coded as brand-name drugs, given that they are approved under an NDA. However, these account for only around 2% of the observations and recoding these as generics does not change the results.

[43] These data were downloaded in June 2015.

[44] The FDA NDC database can be found at http://www.fda.gov/Drugs/InformationOnDrugs/ucm142438.htm.

[45] The date that the marketing began came from the FDA NDC database. The results are robust to the inclusion of a squared age term as well, in case the effects are nonlinear. This variable is in terms of years.

[46] A small percentage (around 2–4 percent) of drugs have negative ages, implying that they were used before they were approved. Since this could be a function of a mistake in the NDC code or a part of an investigational drug, I drop these observations.

using the earliest marketing date listed for each substance (a category that encompasses multiple NDC codes, generic versions, and brand versions), to capture how novel the broad substance is.[47]

The MEPS Prescription Medication (PM) files are linked to the MEPS Full Year Consolidated Data (FY)[48] file, in order to incorporate information on demographics, such as education, race, and age.[49] Additionally, the FY file asks whether each person has a usual payer for prescription drugs and, if so, who the payer is. The following regressions on private insurance and Medicare contributions are run only on people asserting that they expect to be covered by a particular third-party-payer.[50] MEPS also provides information on the amount paid by each payer, the quantity of each drug, and the strength of each drug. Diagnosis codes are also indicated by ICD-9 codes and grouped into 19 diagnosis groups.[51]

---

[47] Each drug version is associated with an application number and a "substance" category. The substance category encompasses multiple drug versions. Usually multiple application numbers are nested within one substance; in some cases very similar substances are associated with one application number. Since this is relatively rare and usually involves highly related substances, I see the "age of the substance" as a broader measure of how novel the substance is, separate from how long a particular drug version has been on the market.

[48] Each observation on the Prescription Medication (PM) file is matched to a Full Year Consolidated Data File (though not vice versa). Many observations in the PM file can match to a single FY observation, since each person may buy multiple prescription medications over a year. Not all observations from FY match, since not all surveyed purchased a prescription medication. These files are linked by a unique identifier, DUPERSID.

[49] Age is measured at the end of each recorded year.

[50] ERISA plans are either noted as "private insurance" or "other payment." If an employer is associated with a private health plan, it is classified as a private payer. If the employer self-insures, MEPS might consider this coverage "other." However, if the self-insured plan is administered by a private insurance company, this might still be listed as "private insurance." If one is concerned that classification of employers' plans might skew the results, the Appendix has a list of tables that do not impose the expected coverage restriction (Table A-3 through Table A-6). Instead, any observation with non-zero payment for a payer is included. Thus, if a person does not expect private insurance payment but receives it for a given drug, that observation will be included in the private insurance price regressions. These results seem similar to the main results.

[51] Observations are additionally dropped if the following values are missing: NDC codes, drug quantity, race, and Hispanic status. Additionally, if the drug is not categorized as "NDA", "NDA authorized Generic", or "ANDA", it is dropped. Drugs with quantities less than five are dropped as well. Finally, drugs are dropped if they are labeled "over-the-counter."

All of the MEPS prices have been converted into 2010 prices.[52] The prices are divided by reported quantity so that they are prices per unit.[53] To ensure that prices are not skewed by inordinately small quantities (and to control for the possibility that they are systematically different or are promotional packages), I only include observations with quantities of 5 or more.[54] All strengths are reported in terms of milligrams.[55]

V. Results

A. Descriptive Statistics

Table 1 lists the mean price per unit based on NTI and generic status, as paid by each payer. These are the raw differences and do not control for other attributes that may affect price. Price per unit is clearly lower for generic drugs than for brand-name drugs. However, it also appears that NTI drugs are generally cheaper than non-NTI drugs. The first column lists the mean total price per unit. Brand-name NTI drugs have a mean price of $0.49 per unit while brand-name non-NTI drugs have a mean price of $5.99. In contrast, generic NTI drugs are priced at $0.28 while generic non-NTI drugs are $0.72.

Table 1 also demonstrates that the differences between generic and brand-name prices are smaller for NTI drugs than for non-NTI drugs. For example, the average total price per unit gap

[52] The CPI indices are from http://www.bls.gov/cpi/cpid10av.pdf.
[53] The unit reported by MEPS varies. However, the main results are robust to using only drugs sold in capsules, caplets, or tablets, shown in Table A-1–Table A-2. This alleviates concerns that types of units (and the quantity reported in each) drive these results.
[54] The main results are robust to ignoring this exclusion.
[55] For data observations that report the strength in grams, I assume this is an error and impute the active ingredient unit reported by the FDA NDC data. However, the risk penalties are robust to the exclusion of this imputation.

between brand-name drugs and generic drugs is $0.21 for NTI drugs and $5.27 for non-NTI drugs.[56] This pattern seems to hold for the other payers as well.

Table 1. Payment Per Unit, by NTI and Generic/Brand-Name Status[a]

	Total		Private		Medicare		Self Payment		Self Payment Only		Medicaid	
	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.
NTI Drug												
- Brand-Name	0.49	14,211	0.11	4,371	0.26	2,422	0.31	12,403	0.41	7,120	0.31	1,257
- Generic	0.28	21,354	0.05	5,745	0.14	4,004	0.17	18,182	0.21	10,366	0.20	2,310
Non-NTI Drug												
- Brand-Name	5.99	30,657	3.05	7,023	4.70	4,580	1.20	21,734	4.43	2,976	4.60	7,386
- Generic	0.72	26,785	0.35	6,357	0.45	4,698	0.23	19,986	0.41	7,336	0.47	5,650

[a] The above amounts are the means of payments for participants who report being covered by the relevant insurance. For self payment, the means are for those records reporting a non-zero self-payment.

B. Price Regression

The following section builds upon the descriptive statistics by examining whether the inherent riskiness in NTI drugs results in a price penalty. Table 2 lists the price penalties associated with NTI drugs, using a price equation with fixed effects for each individual.[57] Panel A lists results of price regressions while Panel B reports the results of *ln(p+1)* regressions. These regressions control for year, Multum therapeutic class, dosage form, pharmacy type, diagnosis code, substance age, drug age, competition, and dosage strength. Most variables included perform as expected. *Drug Age*, meant to capture the amount of time the particular drug has been on the market, has a negative and significant effect on price. Similarly, *Competition* is negative and significant. *Drug Strength* is generally positive, though sometimes insignificant. This is in line with expectations that stronger pills will be more expensive; however, there are situations in

[56] These figures are derived by subtracting the generic NTI price from the brand NTI price (0.49-0.28=0.21). A similar calculation is done for non-NTI drugs (5.99-0.72=5.27).
[57] I also run a weighted OLS regression, clustering by individual, which produce similar risk penalties. These results are available upon request.

which drug price is insensitive to the amount of the active ingredient, a phenomenon termed "flat pricing" discussed by Berndt (2002).

Column (1) shows the price penalties associated with NTI status for the total per-unit drug price. The NTI penalty is negative and significant at the one percent level, suggesting that there is a price penalty for NTI status for brand-name drugs. Generic status is also associated with a decrease in price, which is expected. The coefficient on the interaction term between NTI status and generic status (*NTI*Generic*) is positive and significant. This confirms the pattern observed in Table 1: the gap between brand-name and generic prices is smaller for NTI drugs than non-NTI drugs. This result was previewed in the descriptive statistics but seems robust to the inclusion of other relevant factors such as drug age.

Table 2. Price Sensitivity by Payer, Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-5.744***	-3.716***	-3.442***	-0.648***	-3.199***	-2.750
	(0.964)	(1.075)	(0.662)	(0.152)	(0.690)	(2.011)
Generic Status = 1	-3.326***	-1.502***	-2.938***	-0.445***	-2.718***	-2.383***
	(0.296)	(0.338)	(0.377)	(0.063)	(0.471)	(0.483)
NTI*Generic	3.593***	1.821***	3.073***	0.349***	2.622***	2.327***
	(0.412)	(0.440)	(0.426)	(0.078)	(0.520)	(0.882)
Drug Age	-0.070***	-0.047***	-0.032***	-0.007**	-0.017	-0.066**
	(0.015)	(0.017)	(0.011)	(0.003)	(0.013)	(0.029)
Competition	-0.046***	-0.038***	-0.022**	-0.004**	-0.039***	-0.036***
	(0.007)	(0.010)	(0.011)	(0.002)	(0.009)	(0.012)
Drug Strength/1000	-0.006	0.741***	0.014***	0.011	0.056***	-0.009
	(0.016)	(0.008)	(0.003)	(0.013)	(0.017)	(0.024)
R-squared	0.681	0.473	0.579	0.103	0.767	0.356
Panel B: ln(p+1) Regressions						
NTI Status = 1	-0.788***	-0.525***	-0.678***	-0.144***	-0.694***	-0.204*
	(0.051)	(0.089)	(0.117)	(0.022)	(0.070)	(0.121)
Generic Status = 1	-0.681***	-0.379***	-0.672***	-0.148***	-0.638***	-0.420***
	(0.033)	(0.057)	(0.076)	(0.015)	(0.066)	(0.062)
NTI*Generic	0.574***	0.362***	0.616***	0.072***	0.530***	0.350***
	(0.038)	(0.066)	(0.083)	(0.017)	(0.070)	(0.085)
Drug Age	-0.010***	-0.005***	-0.008***	-0.001**	-0.004**	-0.009***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Competition	-0.009***	-0.006***	-0.004**	-0.002***	-0.011***	-0.007***
	(0.001)	(0.002)	(0.002)	(0.000)	(0.002)	(0.002)
Drug Strength/1000	0.003	0.065***	0.004***	0.001	0.011***	0.002
	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	93,007	23,496	15,704	93,007	27,798	16,603
R-squared	0.725	0.384	0.589	0.150	0.708	0.324
Number of id	9,209	3,072	1,636	9,209	4,487	1,760

Additional variables included but not reported are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year. Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

This positive interaction term is possibly due to the lack of substitutability between versions of NTI drugs, as predicted in Section III. Given how small changes in dosage can lead to adverse events in NTI drugs, switching between brand-name drugs and generic drugs or between generic versions has a higher likelihood of adverse events. Thus, switching is costly, and generic versions of NTI drugs are no longer perfect substitutes for one another.

Given the current fractured system, in which different payers contribute to drug payments, I am interested whether third-party payers are sensitive to the NTI risk. Columns (2), (3), and (6) report the results for the price paid by private insurance, Medicare, and Medicaid, respectively. The same results observed in Column (1) are mirrored for private insurance contributions and Medicare contributions.[58] While the results for Medicaid are similar as well, the NTI penalty is insignificant in most specifications. Again, this might be due to the unobserved error introduced by prevalent rebates.

As noted above, columns (2), (3) and (6) only include people who report usually being covered by the relevant payer for prescription medication are included in these regressions. In case this restriction is unnecessarily stringent, Table A-3–Table A-6 in the Appendix include any observation with nonzero payment in each category. Thus, an observation will be included in the private contribution regression if private insurance contributes a nonzero amount, even if the person does not claim that private insurance is their expected payer. Since there are no zero prices in these regressions, Panel B of each runs *ln(p)* rather than *ln(p+1)*. The results reported in these tables are qualitatively similar to the main results.[59]

---

[58] The price penalties for Medicare contributions for generic NTI drugs seem smaller than private penalties. In Table A-1-Table A-2, the generic penalty is no longer negative for Medicare contributions.

[59] The only difference in these results is that log price results for private contribution lose significance in Table A-4, Table A-5, and Table A-6 and the interaction term is negative in Table A-

Column (4) studies any contribution by the patient; this includes copayments with insurance as well as situations in which the patient bears the entire burden of the purchase. Column (4) reports smaller effects than other columns, largely because it includes all observations, many of which have zero contributions. This is because Column (4) groups copayments with self-payments only, and copayments might be different than the latter circumstances. Copayments are often uniform or vary by tier, and are largely determined by the third-party payer. To separate out this effect, Column (5) only studies observations in which there is no other contributor other than the patient and where there are nonzero payments by the patient; this captures the burden on the consumer when they are the only payer for the drug. These results are qualitatively similar to those Columns (1), (2), and (3).

Column (6) lists the results for Medicaid reimbursements: these are qualitatively similar to those of the other payers, though the NTI penalty is often smaller.[60] Again, these results should be interpreted cautiously, as Medicaid incorporates a lot of rebates that are unobserved in my data. In terms of relative sensitivity, it seems like payments from patients who receive no other compensation and Medicare contributions have the largest NTI penalty for brand-name drugs. However, it is not clear if this difference is significant.

As noted above, the MEPS data do not take into account rebates. Insofar as rebates are a big part of drug pricing, experts say that they can apply to brand-name drugs, often those with one or more close substitutes.[61] I cannot observe rebates; however, I attempt to predict how

---

5, and Table A-6 (interaction term is negative and insignificant). However, the regular price regression has the same sign and significance as in Table 2-Table 3.

[60] Notably, in the weighted OLS regressions the Medicaid penalty seem closer to the magnitude of the other payers. Additionally, sometimes the percent difference between generic NTI drugs and generic non-NTI drugs is not negative.

[61] This information is from a conversation with an expert from the CBO, as well as a CBO report (CBO 2010).

unobserved rebates could affect my results. Rebates could affect my results in three possible ways: First, if the NTI market is more segmented, such that different versions of NTI drugs are not as interchangeable, then NTI brand-name manufacturers might have less need to offer rebates because there is less competition between versions. This would mean that the effective brand-name price paid by each party for non-NTI drugs is lower than reported, while the price for brand-name NTI and all generic drugs would remain the same. Second, brand-name drug manufacturers might be more desperate to start consumers on their drug initially, given the demonstrated inertia once on a version. This may cause them to offer rebates as much or more frequently than non-NTI manufacturers. Third, it is possible that this source of error is random, such that brand-name manufacturers of NTI drugs are just as likely to offer rebates as brand-name manufacturers of non-NTI drugs.

In the latter two scenarios, my results are unaffected or, if anything, understated. If instead the first scenario is true, the price penalty might be overstated by unobserved rebates. As a robustness check, I lower per unit payments to brand-name, non-NTI drugs by 20%. I do this to be safe, since the mean rebate for brand-name drugs is approximately 15–16%.[62] The results are robust to this change and are reported in Table 3.[63] This suggests that the NTI price penalty and the smaller gap between NTI brand-name and generic drugs are not solely an artifact of unobserved rebates.

[62] This information is from a conversation with an expert from the CBO. A CBO report also estimates that manufacturer rebates to plans averaged about 14 percent of brand-name prescription drug spending (CBO 2010).
[63] Rebates are also included in Table A-2,Table A-4, and Table A-6.

Table 3. Price Sensitivity by Payer, Fixed Effects, Adjustment for Rebates.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-4.369***	-2.902***	-2.596***	-0.470***	-2.462***	-1.948
	(0.766)	(0.856)	(0.539)	(0.120)	(0.546)	(1.597)
Generic Status = 1	-2.467***	-1.120***	-2.228***	-0.326***	-2.067***	-1.746***
	(0.237)	(0.268)	(0.303)	(0.050)	(0.375)	(0.384)
NTI*Generic	2.636***	1.362***	2.307***	0.229***	1.959***	1.677**
	(0.328)	(0.350)	(0.343)	(0.062)	(0.414)	(0.697)
Drug Age	-0.058***	-0.039***	-0.027***	-0.005**	-0.014	-0.055**
	(0.012)	(0.014)	(0.009)	(0.002)	(0.011)	(0.023)
Competition	-0.037***	-0.031***	-0.017**	-0.003**	-0.031***	-0.030***
	(0.005)	(0.008)	(0.008)	(0.001)	(0.007)	(0.009)
Drug Strength/1000	-0.004	0.593***	0.012***	0.009	0.049***	-0.006
	(0.013)	(0.007)	(0.003)	(0.011)	(0.014)	(0.019)
	0.676	0.470	0.572	0.101	0.764	0.353
Panel B: ln(p+1) Regressions						
NTI Status = 1	-0.650***	-0.454***	-0.567***	-0.101***	-0.578***	-0.137
	(0.049)	(0.083)	(0.111)	(0.021)	(0.066)	(0.114)
Generic Status = 1	-0.549***	-0.309***	-0.563***	-0.112***	-0.524***	-0.336***
	(0.031)	(0.052)	(0.070)	(0.013)	(0.062)	(0.058)
NTI*Generic	0.441***	0.290***	0.504***	0.036**	0.416***	0.269***
	(0.036)	(0.061)	(0.078)	(0.016)	(0.066)	(0.079)
Drug Age	-0.009***	-0.005***	-0.007***	-0.001**	-0.003**	-0.009***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Competition	-0.008***	-0.006***	-0.004**	-0.001***	-0.010***	-0.006***
	(0.001)	(0.002)	(0.002)	(0.000)	(0.002)	(0.002)
Drug Strength/1000	0.003	0.061***	0.003***	0.001	0.011***	0.002
	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	93,007	23,496	15,704	93,007	27,798	16,603
R-squared	0.698	0.377	0.575	0.134	0.677	0.313
Number of id	9,209	3,072	1,636	9,209	4,487	1,760

Additional variables included but not reported are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year . Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

The focus of this Chapter is on the existence of a penalty; different specifications produce various penalty magnitudes. In any specification, however, the magnitudes of these estimates seem economically significant. From Table A-6, the percentage difference between the total price of a brand-name NTI drug and a brand-name non-NTI drug is roughly –58%. This high percentage difference for brand-name drugs may be driven by outlier brand-name, non-NTI drugs. The percentage difference in total price between a generic NTI drug and a generic non-NTI drug is –31%. The corresponding results for total payment in Table A-2 suggest a –51% percentage difference between brand-name NTI drug and a brand-name non-NTI drug and a –12% difference between a generic NTI drug and a generic non-NTI drug.[64]

These results are distinguishable from those found in Nabin et al. (2012), which suggest that lower therapeutic equivalence drug classes have higher post-patent prices than classes with higher therapeutic equivalence. There may be several reasons for this. First, my data is from the US, as opposed to Canada, which has a very different medical care system. Second, while I do control for competition, my study does not concentrate on price changes in the period around patent expiration, as theirs does. Additionally, instead of classifying entire classes as having low therapeutic equivalence, this Chapter singles out drugs within classes that have been considered as having a narrow therapeutic index and compares them to other drugs within the same class. My results are similarly distinct from those found by Chao et al. (2002), who suggest that NTI generic drugs have similar discounts as generic non-NTI drugs. This difference might be due to the use of newer data, as Chao et al. look at data from 1996–1998. With the continuing emphasis on NTI drugs, the awareness of the risks might have increased over time. Additionally, Chao et

[64] I compute the percent difference for ln(p+1) as follows: $\frac{[e^{(c+\beta_1)}-1]-[e^{c}-1]}{[e^{c}-1]}$ for the brand-name NTI penalty and $\frac{[e^{(c+\beta_1+\beta_2+\beta_3)}-1]-[e^{c+\beta_2}-1]}{[e^{c+\beta_2}-1]}$ for the generic NTI penalty.

al. (2002) consider 23 NTI drugs, and it is unclear which these were. This study uses the North Carolina Board of Pharmacy definition, as following previous work (Gagne et al. 2013); expanding the definition might produce different results, especially as the perception of NTI risk becomes more attenuated.

Finally, this analysis does not take into account any differences in marginal costs of production. However, Berndt (2002) notes that drug pricing reflects marginal value, not marginal production cost. Thus, it is unlikely that differences in marginal costs are driving the result.

C. Substitutability Amongst Versions

In order to claim that the consumers are sensitive to costly switching, it is helpful to look at consumer behavior in addition to price. This section examines whether consumers are loyal to their particular version of NTI drug, whether it be a generic or brand-name version.

In this analysis, each observation is an individual–substance–purchase round combination indicating the mixture of versions of a substance consumed by an individual in a given purchase round. Each unique product code[65] is treated as a different version of each drug. For example, each unique product code associated with the drug digoxin is treated as a separate "version." If a consumer buys multiple different versions of the drug within a given purchase round, he will be considered to have consumed a "mixed" bundle of that drug. Only patients with multiple purchases in a round are included in this analysis, since it would be infeasible to consume a "mixed" bundle with only one purchase.

While MEPS provides therapeutic class categories, there might be a concern that this classification may not be stable over a period of time. In order to accommodate such fluctuations, I consider drugs considered to be in each of the seven subclasses in 2008 and follow

[65] Product code is the first two segments of the NDC code, distinguishing based on manufacturer and product (but not package).

them over time. This ensures that I see a stable group of drugs. Additionally, since each drug can be assigned up to three therapeutic classes, this is an overly broad classification and makes sure that drugs are compared within each class they possibly belong to.[66]

Table 4 shows a linear probability model[67] predicting the likelihood with which a particular drug is consumed in a mixed bundle. Column (1) only includes NTI status. *NTI* is negative but insignificant, which suggests that if the drug is a NTI drug, it is insignificantly less likely to be consumed in a mixed bundle. However, by including the lagged mix variable in Column (2), the story becomes more complex. Consumers of NTI drugs are insignificantly less likely to consume a mixed bundle ($NTI < 0$) if the consumer did not mix previously; however, there is considerable evidence of path dependence: if a patient consumed the drug in a mixed bundle in the prior period, he is more likely to consume it in a mixed bundle in the current period (*Previous Mixed Period* > 0) than if he previously did not mix. However, the interaction term *NTI*Previous Mixed Period* is negative and significant, showing that consumers that previously mixed NTI drugs were less likely to currently mix than consumers who previously mixed non-NTI drugs. This suggests that NTI status undermines some of the path dependence. This implies that while consumers may experiment with NTI drugs, they are less likely to continue to consume them in a mixed bundle. The predicted probabilities of being in each state, calculated from the probit model in Table A-7 are listed in Table A-8.

[66] This does mean that a drug purchase may be considered more than once if it is associated with more than one of the chosen therapeutic classes.
[67] The linear probability model should estimate the approximation to the marginal effect of the conditional expectation. However, as a robustness check, a probit model is displayed in Table A-7.

Table 4. Probability of Mixing in a Given Period, Linear Probability Model.

	(1)	(2)
NTI	-8.750e-05	-9.774e-03
	(1.037e-02)	(1.181e-02)
Previous Mixed Period		1.695e-01***
		(1.824e-02)
Mixed Previous Period*NTI		-4.675e-02*
		(2.728e-02)
Observations	22,467	14,524
R-squared	0.040	0.065
Additional variables included but not reported are age, missing age, race, sex, education, and the number of drugs consumed in each individual-substance-period combination. Also included are indicators for Multum therapeutic class, and insurance coverage. Standard errors are clustered by individual. Each observation is an individual-substance-period combination. Each version is defined as a unique product code. Only periods in which more than one drug is consumed are included. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.		

The above analysis only categorizes consumption bundles as mixed or not. In Table 5, this analysis is taken further. The consumption of each drug is now separated into 4 states: no mixture of versions, mix of both generic and brand-name versions, a mix of brand-name versions only, or a mix of generic versions only.[68] Linear probability models are run regarding the likelihood of being in each stage. For example, Column (1) shows the likelihood of not mixing versions, Column (2) the likelihood of mixing brand-name and generic versions, Column (3) the likelihood of mixing only within brand-name versions,[69] and Column (4) the likelihood of mixing only within generic versions.

[68] I only consider substances that have more than one version on the market: for Table 4 and Column (1) of Table 5, this means requiring more than one version being on the market. For Table 5 Column (2), I only exclude substances that have no generic versions or no brand versions since this necessarily mixes brand and generic versions. Column (3) only uses substances with more than one brand versions and column (4), only substances with more than one generic version.

[69] The "mixed brand" category encompasses consuming different brands of the same substance, the same brand made by different labelers, or the same brand by the same labeler that is given a separate product code. Since this is a large range of differences, this Chapter suggests that the chief comparison should be between unmixed bundles and mixing brand and generic versions.

Table 5. Version Loyalty: Linear Probability Model.

	(1)	(2)	(3)	(4)
	No Mix	Mixed Brand-Generic	Mixed Brand	Mixed Generic
NTI = 1	0.013	0.001	-0.008*	-0.017*
	(0.012)	(0.009)	(0.004)	(0.010)
Lagged Consumption Bundles				
Previous Mixed Brand-Generic	-0.243***	0.236***	-0.020***	-0.003
	(0.036)	(0.035)	(0.003)	(0.020)
Previous Mixed Brand	-0.108***	0.078	0.148***	-0.073***
	(0.035)	(0.053)	(0.035)	(0.009)
Previous Mixed Generic	-0.155***	-0.032***	-0.015***	0.174***
	(0.025)	(0.008)	(0.004)	(0.024)
Lagged Consumption Bundles–NTI Interactions				
NTI*Previous Mixed Brand-Generic	0.148***	-0.129***	-0.006	0.013
	(0.044)	(0.043)	(0.004)	(0.023)
NTI*Previous Mixed Brand	0.050	-0.110**	-0.042	0.052***
	(0.053)	(0.055)	(0.049)	(0.010)
NTI*Previous Mixed Generic	-0.036	-0.014	-0.005	0.073
	(0.048)	(0.013)	(0.005)	(0.049)
Observations	14,524	10,662	12,873	11,868
R-squared	0.067	0.042	0.042	0.081

Each column is a separate linear probability model. Additional variables included but not reported are age, race, missing age, sex, education, and the number of drugs consumed in each individual-substance-period combination. Also included are indicators for Multum therapeutic class, and insurance coverage. Standard errors are clustered by individual. Each observation is an individual-substance-period combination. Each version is defined as a unique product code. Only periods in which more than one drug is consumed are included. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Each column has different observations: for Column (1) only substances more than one version are included. For Column (2), I exclude substances that have no generic versions or no brand versions since this necessarily mixes brand and generic versions. Column (3) only uses substances with more than one brand versions and column (4), only substances with more than one generic version.

It is clear that path-dependence is very important in this context. The omitted stage category is *Previous Unmixed*. Relative to this, the lagged consumption bundles in Column (1) are significantly negative (*Previous Mixed Brand-Generic, Previous Mixed Brand, Previous Mixed Generic* < 0), indicating that the likelihood of consuming an unmixed bundle in the current period increases if a person previously consumed an unmixed bundle. Similarly, in Column (2), *Previous Mixed Brand-Generic* is significant and positive, indicating that previously consuming a mixed brand-generic bundle increases the likelihood of currently consuming a mixed brand-generic bundle, relative to previously consuming an unmixed bundle. A similar

effect can be seen for *Previous Mixed Brand* in Column (3) and *Previous Mixed Generic* in Column (4).

Table 5 confirms the pattern observed in Table 4: learning undermines path dependence for NTI drugs. The interactions between lagged consumption bundle and NTI status in Column (1) demonstrate that NTI users exhibit learning. Patients who previously consumed NTI drugs in mixed brand-generic bundles (*NTI*Previous Mixed Brand-Generic*>0) are more likely to consume unmixed bundles in the current period than those whose consumed non-NTI drugs in mixed brand-generic bundles in the prior period. This effect is insignificant for those who consumed NTI drugs in mixed brand bundles or in mixed generic bundles. [70]

The interaction terms in Column (2) are similarly informative: if a patient had consumed a mixed brand-generic bundle of an NTI drug in the prior period, they are less likely to consume a mixed brand-generic bundle in the current period (*NTI*Previous Mixed Brand-Generic* <0) than if they previously consumed a non-NTI drug in a mixed brand-generic bundle. The same is true for those who consume previous mixed brand bundles (*NTI*Previous Mixed Brand*<0). For both Columns (3) and (4), *NTI* is negative and significant, indicating that patients who previously consumed unmixed NTI bundles were significantly less likely to consume mixed brand or mixed generics bundles than those who previously consumed non-NTI drugs in unmixed bundles. The interaction terms in Column (3) are negative and insignificant, indicating that they are not significantly less likely to consume mixed brand bundles than those previously consuming unmixed NTI bundles. Since those consuming unmixed NTI bundles are significantly less likely to consume a mixed brand bundle, this means that for Column (3), path-dependence is undercut for all previous NTI consumption bundles. The interaction terms in Column (4) are a bit more

[70] The interaction for *NTI*Previous Mixed Generic* is negative but insignificant.

puzzling; the interaction terms are positive and sometimes significant, indicating that relative to previously consuming unmixed bundles of NTI drugs, consumers are significantly more likely to consume a mixed bundle of generic NTI drugs if they previously consumed a mixed bundle of brand NTI drugs or a mixed bundle of generic NTI drugs. In order to illustrate the differences in predicted probabilities of being in each state, Table A-10 lists the predicted probabilities in each stage.[71]

In sum, while NTI status does not necessarily reduce the likelihood of mixing versions, the results seem largely consistent with the idea of learning. In general, those who mix NTI drugs are less likely to continue mixing than those taking non-NTI drugs. One possible explanation for this is that NTI users might struggle to find a version that works for them. Once they find one, they continue to use it. If this is the case, this supports the idea that each version of an NTI drug has a "loyalty" usually only found with brand-name drugs, explaining the positive interaction in Section I.B.

Finally, some might be concerned that consumption bundles per purchase round is not the right way to analyze this data. If patients consume an unmixed bundle within each period but consume a different unmixed bundle across periods, this might be weaker evidence of loyalty. This might not be a problem, since each purchase round spans multiple months, creating a relatively long period to capture purchases. However, if this is a concern, I perform the following robustness check. I examine all the observations in which both the current and previous consumption bundles were unmixed. I run a linear probability model measuring the likelihood that the current and previous consumption bundles contain two different versions of the drug.

[71] These probabilities are calculated using the probit models displayed in Table A-9.

The results are listed in Table 6. Patients consuming NTI drugs were significantly less likely to switch to a different version than non-NTI drugs.

Table 6. Likelihood of Switching Between Bundles, Linear Probability Model.

	(1)
NTI	-0.049***
	(0.013)
Anticonvulsants	0.000
	(0.013)
Anticoagulants	0.103***
	(0.021)
Thyroid Drugs	0.071***
	(0.016)
Immunosuppressive Agents	-0.127***
	(0.018)
Bronchodilators	-0.096***
	(0.013)
Antipsychotics	-0.031**
	(0.015)
Observations	11,809
R-squared	0.022

Each column is a separate linear probability model. Additional variables included but not reported are age, missing age, sex, education, and the number of drugs consumed in each individual-substance-period combination. Also included are indicators for Multum therapeutic class, and insurance coverage. Standard errors are clustered by individual. Each observation is an individual-substance-period combination in which both the current and previous consumption bundle is unmixed. Each version is defined as a unique product code. Only periods in which more than one drug is consumed are included. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

There are important limitations of these results. First, as noted above, each participant is interviewed for 5 rounds, in which their prescription purchases are monitored. Some participants only purchase prescriptions in a particular class in nonconsecutive purchase rounds. For example, a participant may purchase drugs in rounds 3 and 5 of their panel. For the purpose of this section, I treat purchase round 3 as the "lagged" period for purchase round 5, instead of saying that there were no drug purchases in the lagged period. I do this to preserve information

about prior purchases and to allow for arbitrariness in the division of purchase rounds. If instead, a participant purchases drugs in the same class in rounds 3, 4, and 5 but only consumes a particular drug in rounds 3 and 5, the lagged consumption of that drug for round 5 would be missing, as it is less likely that a consumer was merely not in the market for drugs. The link between the two periods seems too tenuous, as a separate purchase was made in the interim period. Additionally, since the panel set is thus not attached to years (purchase rounds span years in unpredictable ways), this analysis does not account for any uniform time trend that might shift path dependence by consumer.[72]

Another possible concern is that some of the lack of mixing and switching might be due to PBM deals to keep certain drug versions on the formulary. If this is true, perhaps the most informative bundle comparison is the mixture of generic and brand versions: there is usually a brand or generic option for any formulary, while the formulary may only cover particular brand or generic versions.

However, even if freedom to mix is constrained by insurance formularies, the results do suggest mixing behavior differs between NTI and non-NTI drugs. One concern might be that the freedom to mix is more limited for NTI drugs than non-NTI drugs: then the observed difference is merely a function of having fewer NTI options on a formulary. I have no reason to believe this is true—in fact it might be possible that more versions of NTI drugs are available on a formulary. One scholar suggests that formularies might not consider NTI drugs versions as substitutes for

[72] Similarly, the mixing status is determined with respect to an FDA classification of "substances." Insofar as this substance classification is too broad, such that multiple drugs within the substance must be taken in conjunction with one another, this might introduce error.

one another. [73] If this means that they keep more versions of drugs on their formularies, then the possible switching for NTI drugs is actually higher than non-NTI drugs, understating this result.

If the freedom to mix is equally constrained by insurance formularies for NTI and non-NTI drugs, then the differences in behavior imply that the loyalty effect is merely seen at the PBM level, not at the consumer level. If PBMS and third-party payers are aware of the problems of switching within NTI versions, they might be less willing to switch versions on the formulary.

In sum, these results suggest that observed consumption behavior is consistent with expected sensitivities to NTI risk. These results support the results from the price model by corroborating the predicted loyalty to drug version for NTI drugs.

VI. Conclusion

The consumption of pharmaceuticals is characterized by disconnected incentives: physicians prescribe drugs, third-party payers bear most of the costs, and patients face all of the risk of a given drug. In such a context, does the market reflect a lower willingness to pay for riskier drugs? NTI drugs provide a case study of a perceived risky drug that remained unregulated by the government. This Chapter examines whether the risks associated with NTI drugs result in a price penalty, and if so, which payers were sensitive to this risk.

This Chapter suggests that there is sensitivity to risk despite lack of government action. Although there are slightly different price sensitivities to risk by payer, most were responsive to substitution risks. Since the risk involved is triggered when consumers switched between versions of the NTI, generic versions of the drug are less likely to be considered substitutes for brand-name versions or for other generic versions. When switching is costly, consumers become

[73] Fox (2003) describes how PBMs work. Fox notes that PBMs may treat NTI drugs as having no generic equivalents, decide on a drug-by-drug basis, or treat NTI drugs similarly to non-NTI drugs.

loyal to versions of the drug. This story is consistent with the observed smaller gap between brand and generic prices for NTI drugs than for non-NTI drugs. Additionally, the analysis of consumption bundles by NTI status is consistent with the idea that NTI consumers experiment with versions and then remain loyal to a particular version. These findings also reveal that the nature of the risk matters. The fact that switching is costly both makes the entire category of drugs less preferable and also affects decisions between versions within the category.

It is unclear how applicable these results are to new products with less publicized risks. The FDA Advisory Committee for Pharmaceutical Science and Clinical Pharmacology only decided to support more stringent guidelines in 2010, so the majority of purchases were done during a time in which there was no regulatory guidance on the risks. However, NTI drugs had been around for a long time and had a big body of research discussing its possible risks. Even though the verity of the risk was still undecided, it is possible that this time made third-party payers more aware of the possible risk; it is unclear whether this would be true for drugs with newer, less publicized potential risks, such as generic NCBD products. Alternatively, perhaps the ambiguity surrounding the risks of such new generic drugs would be sufficient to trigger the same sensitivity. This is best explored by future work.

The takeaway is relatively optimistic. The response to risk in both the price context and in the analysis of consumption bundles is consistent with consumers and third-party payers being sensitive to risk. The sensitivity of third-party payers is particularly intriguing: in Section I.B I observe a NTI price penalty even for third-party payers' contributions. Similarly, if the results in Section I.C are influenced by formulary decisions, this would imply that third-party payers are sensitive to consumers' preferences. This sensitivity may be for competition reasons: insurance

companies want to keep consumers from switching to other insurance companies.[74] Alternatively, PBMS utilize Pharmacy and Therapeutics Committees composed of pharmacists and doctors (Cohen 2000). These members are aware of such risks and may take this into account while establishing their formularies. Either way, these results seem to suggest that despite the agency issues in the pharmaceutical consumption market, the observed outcomes seem consistent with consumers' preferences against risk.

VII. References

Al-Jazairi, Abdulrazaq S., Sakra Blhareth, Iyad S. Eqtefan, & Saleh A. Al-Suwayeh. 2008. "Brand and Generic Medications: Are They Interchangeable?" *Annals of Saudi Medicine, 28*(1), 33.

Berg, Michel J., Robert A. Gross, Lisa S. Haskins, Wendy M. Zingaro, & Kenneth J. Tomaszewski. 2008. "Generic Substitution in the Treatment of Epilepsy: Patient and Physician Perceptions." *Epilepsy & Behavior, 13*(4), 693–699.

Berndt, Ernst R. 2002. "Pharmaceuticals in US Health Care: Determinants of Quantity and Price." *Journal of Economic Perspectives*, 45–66.

Bhosle, Monali, Suji S Sansgiry, & Rajesh Balkrishnan. 2005. "Consumer Perspectives Regarding Use of Generic Equivalents for Medications with Narrow Therapeutic Index." *Poster Session*, *available at* http://www.academyhealth.org/files/2005/studentposters.pdf.

Blix, Hege S., Kirsten K. Viktil, Tron A. Moger, & Aasmund Reikvam. 2010. "Drugs with Narrow Therapeutic Index as Indicators in the Risk Management of Hospitalised Patients." *Pharm Pr Granada, 8*, 50–55.

[74] This might not be true for Medicaid patients, who likely do not have access to other third-party payers.

Chao, J., S. D. Taylor, P. L. McKercher, & D. M. Kirking. 2002. "Generic Narrow Therapeutic Index Drug Use: 1996–1998. " *Value in Health, 5*(3), 139.

Cohen, Joshua P. 2000. "PBMs and the Medicare Prescription Drug Benefit." *Food & Drug Law Journal,* 55*,* 311.

Congressional Budget Office. 2010. *Effects of Using Generic Drugs on Medicare's Prescription Drug Spending*, *available at* https://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/118xx/doc11838/09-15-prescriptiondrugs.pdf.

Crawford, Pamela, Morgan Feely, Alan Guberman, and Gunter Kramer. 2006. "Are There Potential Problems with Generic Substitution of Antiepileptic Drugs?: A Review of Issues." *Seizure*, *15*(3), 165–176.

Duh, Mei Sheng, Kevin E. Cahill, Pierre Emmanuel Paradis, Pierre Y. Cremieux, and Paul E. Greenberg. 2009. "The Economic Implications of Generic Substitution of Antiepileptic Drugs: A Review of Recent Evidence." *Expert Opinion on Pharmacotherapy*, *10*(14), 2317–2328.

Dunne, Suzanne S., and Colum P. Dunne. 2015. "What Do People Really Think of Generic Medicines? A Systematic Review and Critical Appraisal of Literature on Stakeholder Perceptions of Generic Drugs." *BMC Medicine 13*(1), 173.

Fairfax-Columbo, Jaymes V., and David DeMatteo. 2015. "Are Bioequivalents Really Equal: Genetic Substitution in the Context of Mental Illness." *Ind. Health L. Rev., 12*, 281.

Fischer, Michael A., and Jerry Avorn. 2003. "Economic Consequences of Underuse of Generic Drugs: Evidence from Medicaid and Implications for Prescription Drug Benefit Plans." *Health Services Research, 38*(4), 1051–1064.

Fox, Peter D. 2003. “Prescription Drug Benefits: Cost Management Issues for Medicare.“ *Health Care Financing Review, 25*(2), 7.

Frank, Richard G., and David S. Salkever. 1997. “Generic Entry and the Pricing of Pharmaceuticals.” *Journal of Economics & Management Strategy, 6*(1), 75–90.

Gagne, Joshua J., Jennifer M. Polinski, Aaron S. Kesselheim, Niteesh K. Choudhry, David Hutchins, Olga S. Matlin, Angela Tong, and William H. Shrank. 2013. “Patterns and Predictors of Generic Narrow Therapeutic Index Drug Use Among Older Adults.” *Journal of the American Geriatrics Society, 61*(9), 1586–1591.

Gothe, H., I. Schall, K. Saverno, M. Mitrovic, A. Luzak, D. Brixner, and U. Siebert. 2015. “The Impact of Generic Substitution on Health and Economic Outcomes: A Systematic Review.” *Applied Health Economics and Health Policy, 13*(1), 21–33.

Gottlieb, Scott. 2014. “FDA's Looming Decision on a Generic to Teva's Copaxone Reveals Drug Approval Woes.” *Forbes*, July 2. *available at* http://www.forbes.com/sites/scottgottlieb/2014/07/07/fdas-looming-decision-on-generic-copaxone-from-teva-reveals-drug-approval-woes/.

Haas, Jennifer S., Kathryn A. Phillips, Eric P. Gerstenberger, & Andrew C. Seger. 2005. “Potential Savings from Substituting Generic Drugs for Brand-Name Drugs: Medical Expenditure Panel Survey, 1997–2000.” *Annals of Internal Medicine*, *142*(11), 891–897.

Heaney, Dominic C., and Josemir W. Sander. 2007. “Antiepileptic Drugs: Generic Versus Branded Treatments.” *The Lancet Neurology, 6*(5), 465–468.

Hottinger, Michelle, and Bryan A. Liang. 2012. “Deficiencies of the FDA in Evaluating Generic Formulations: Addressing Narrow Therapeutic Index Drugs.” *American Journal of Law and Medicine, 38*(4), 667–689.

Huskamp, Haiden A., Patricia A. Deverka, Arnold M. Epstein, Robert S. Epstein, Kimberly A. McGuigan, and Richard G. Frank. 2003. "The Effect of Incentive-Based Formularies on Prescription-Drug Utilization and Spending." *New England Journal of Medicine, 349*(23), 2224–2232.

Junod, Suzanne White. 2014. *FDA and Clinical Trials: A Short History*, FOOD AND DRUG ADMINISTRATION, *available at* http://www.fda.gov/AboutFDA/WhatWeDo/History/Overviews/ucm304485.htm#_edn52

Kesselheim, Aaron S., Alexander S. Misono, Joy L. Lee, Margaret R. Stedman, M. Alan Brookhart, Niteesh K. Choudhry, and William H. Shrank. 2008. "Clinical Equivalence of Generic and Brand-Name Drugs used in Cardiovascular Disease: A Systematic Review and Meta-Analysis." *JAMA, 300*(21), 2514–2526.

Midha, Kamal K., and Gordon McKay. 2009. "Bioequivalence; Its History, Practice, and Future." *The AAPS Journal, 11*(4), 664–670.

Nabin, Munirul Haque, Vijay Mohan, Aaron Nicholas, and Pasquale M. Sgro. 2012. "Therapeutic Equivalence and the Generic Competition Paradox." *The BE Journal of Economic Analysis & Policy* 12(1).

Omojasola, Anthony, Mike Hernandez, Sujit Sansgiry, & Lovell Jones. 2012. "Perception of Generic Prescription Drugs and Utilization of Generic Drug Discount Programs." *Ethnicity & Disease, 22*(4), 479.

Reiffen, David, and Michael R. Ward. 2005. "Generic Drug Industry Dynamics." *Review of Economics and Statistics, 87*(1), 37–49.

Shrank, William H., Niteesh K. Choudhry, Jessica Agnew-Blais, Alex D. Federman, Joshua N. Liberman, Jun Liu, Aaron S. Kesselheim, M. Alan Brookhart, and Michael A. Fischer.

2010. “State Generic Substitution Laws can Lower Drug Outlays under Medicaid.” *Health Affairs, 29*(7), 1383–1390.

Shrank, William H., Margaret Stedman, Susan L. Ettner, Dee DeLapp, June Dirstine, M. Alan Brookhart, Michael A. Fischer, Jerry Avorn, and Steven M. Asch. 2007. “Patient, Physician, Pharmacy, and Pharmacy Benefit Design Factors Related to Generic Medication Use.” *Journal of General Internal Medicine, 22*(9), 1298–1304.

Singh, Amrita, Nicole M. Maisch, and Maha Saad. 2014. “A Closer Look at Generic Interchangeability in Narrow Therapeutic Index Drugs.” *U.S. Pharmacist, 39*(6), Generic suppl. 8–12, *available at* http://www.uspharmacist.com/content/s/312/c/48990/.

Yu, L. X., W. Jiang, X. Zhang, R. Lionberger, F. Makhlouf, D. J. Schuirmann, L. Muldowney, B. Davit, D. Conner, and J. Woodcock. 2015. “Novel Bioequivalence Approach for Narrow Therapeutic Index Drugs.” *Clinical Pharmacology & Therapeutics 97*(3): 286–291.

Yu, Lawrence X. 2011. “Quality and Bioequivalence Standards for Narrow Therapeutic Index Drugs.” *FDA Center for Drug Evaluation and Research*, *available at* http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/AbbreviatedNewDrugApplicationANDAGenerics/UCM292676.pdf#sthash.OJmCF9B4.dpuf.

VIII. Appendix

Table A-1. Price Sensitivity by Payer, Fixed Effects, Tablets.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-3.013***	-1.847***	-2.409***	-0.369***	-2.885***	-1.048
	(0.313)	(0.493)	(0.574)	(0.079)	(0.357)	(0.849)
Generic Status = 1	-2.999***	-1.500***	-2.591***	-0.393***	-2.612***	-2.458***
	(0.250)	(0.354)	(0.455)	(0.058)	(0.378)	(0.465)
NTI*Generic	2.968***	1.496***	2.658***	0.290***	2.536***	2.418***
	(0.272)	(0.385)	(0.482)	(0.063)	(0.397)	(0.532)
Drug Age	-0.050***	-0.021**	-0.016	-0.004*	-0.022***	-0.079***
	(0.007)	(0.009)	(0.010)	(0.002)	(0.007)	(0.019)
Competition	-0.049***	-0.029**	-0.022**	-0.004**	-0.035***	-0.037**
	(0.008)	(0.012)	(0.010)	(0.002)	(0.010)	(0.015)
Drug Strength/1000	2.606***	1.228	2.356***	-0.035	0.142	1.419
	(0.525)	(0.791)	(0.901)	(0.144)	(0.545)	(1.051)
R-squared	0.554	0.380	0.553	0.051	0.587	0.262
Panel B: ln(p+1) Regressions						
NTI Status = 1	-0.634***	-0.491***	-0.432***	-0.121***	-0.718***	-0.201
	(0.059)	(0.106)	(0.115)	(0.024)	(0.073)	(0.164)
Generic Status = 1	-0.661***	-0.421***	-0.539***	-0.143***	-0.650***	-0.486***
	(0.041)	(0.075)	(0.078)	(0.016)	(0.073)	(0.082)
NTI*Generic	0.559***	0.383***	0.478***	0.070***	0.555***	0.422***
	(0.046)	(0.084)	(0.085)	(0.018)	(0.077)	(0.100)
Drug Age	-0.009***	-0.005**	-0.003	-0.001*	-0.006***	-0.011***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.004)
Competition	-0.013***	-0.009***	-0.008***	-0.002***	-0.011***	-0.008***
	(0.001)	(0.003)	(0.002)	(0.000)	(0.002)	(0.002)
Drug Strength/1000	0.595***	0.435***	0.626***	-0.039	0.074	0.383**
	(0.078)	(0.160)	(0.172)	(0.038)	(0.105)	(0.176)
Observations	71,431	17,865	12,666	71,431	25,028	11,888
R-squared	0.695	0.391	0.593	0.087	0.650	0.300
Number of id	6,990	2,327	1,422	6,990	3,870	1,212

Additional variables included but not reported are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A-2. Price Sensitivity by Payer, Adjustment for Rebates, (Tablets only), Fixed Effects.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-2.167***	-1.402***	-1.715***	-0.254***	-2.225***	-0.483
	(0.264)	(0.395)	(0.476)	(0.065)	(0.288)	(0.755)
Generic Status = 1	-2.182***	-1.107***	-1.918***	-0.283***	-1.986***	-1.740***
	(0.203)	(0.283)	(0.365)	(0.047)	(0.304)	(0.382)
NTI*Generic	2.116***	1.089***	1.940***	0.182***	1.896***	1.686***
	(0.221)	(0.308)	(0.386)	(0.051)	(0.320)	(0.434)
Drug Age	-0.041***	-0.017**	-0.013	-0.003**	-0.018***	-0.063***
	(0.006)	(0.007)	(0.008)	(0.001)	(0.006)	(0.016)
Competition	-0.041***	-0.024**	-0.018**	-0.003**	-0.029***	-0.030***
	(0.006)	(0.010)	(0.008)	(0.001)	(0.008)	(0.012)
Drug Strength/1000	2.176***	1.050	2.017***	-0.040	0.181	1.066
	(0.426)	(0.650)	(0.730)	(0.116)	(0.434)	(0.875)
	0.537	0.366	0.546	0.047	0.577	0.254
Panel B: ln(p+1) Regressions						
NTI Status = 1	-0.505***	-0.417***	-0.337***	-0.082***	-0.605***	-0.119
	(0.057)	(0.099)	(0.111)	(0.023)	(0.069)	(0.157)
Generic Status = 1	-0.530***	-0.346***	-0.439***	-0.107***	-0.537***	-0.387***
	(0.039)	(0.069)	(0.073)	(0.015)	(0.069)	(0.077)
NTI*Generic	0.427***	0.307***	0.375***	0.035**	0.442***	0.324***
	(0.044)	(0.077)	(0.080)	(0.017)	(0.073)	(0.094)
Drug Age	-0.009***	-0.005**	-0.003	-0.001**	-0.005***	-0.010***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)
Competition	-0.012***	-0.008***	-0.007***	-0.002***	-0.011***	-0.007***
	(0.001)	(0.002)	(0.002)	(0.000)	(0.002)	(0.002)
Drug Strength/1000	0.581***	0.429***	0.620***	-0.040	0.092	0.360**
	(0.075)	(0.152)	(0.165)	(0.035)	(0.101)	(0.165)
Observations	71,431	17,865	12,666	71,431	25,028	11,888
R-squared	0.663	0.375	0.579	0.073	0.614	0.284
Number of id	6,990	2,327	1,422	6,990	3,870	1,212

Additional variables included but not are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year. Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A-3. Price Sensitivity by Payer, Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-5.761***	-7.750***	-3.127***	-0.837***	-3.199***	-10.132**
	(0.965)	(2.254)	(0.532)	(0.206)	(0.690)	(4.072)
Generic Status = 1	-3.336***	-2.676***	-3.245***	-0.663***	-2.718***	-4.317***
	(0.297)	(0.492)	(0.382)	(0.094)	(0.471)	(0.711)
NTI*Generic	3.599***	2.981***	3.137***	0.548***	2.622***	5.888***
	(0.412)	(0.705)	(0.419)	(0.114)	(0.520)	(1.426)
Drug Age	-0.070***	-0.089***	-0.049***	-0.010**	-0.017	-0.130***
	(0.015)	(0.026)	(0.011)	(0.004)	(0.013)	(0.046)
Competition	-0.046***	-0.036***	-0.024**	-0.004	-0.039***	-0.066***
	(0.007)	(0.013)	(0.012)	(0.002)	(0.009)	(0.014)
Drug Strength/1000	-0.006	-0.543***	0.002	0.014	0.056***	0.085***
	(0.016)	(0.110)	(0.002)	(0.016)	(0.017)	(0.028)
R-squared	0.681	0.729	0.648	0.156	0.767	0.717
Panel B: ln(P) Regressions						
NTI Status = 1	-1.427***	-1.460***	-1.496***	-0.776***	-1.555***	-1.114***
	(0.097)	(0.291)	(0.202)	(0.094)	(0.167)	(0.258)
Generic Status = 1	-1.298***	-1.179***	-1.474***	-0.871***	-1.370***	-1.273***
	(0.062)	(0.132)	(0.127)	(0.062)	(0.143)	(0.117)
NTI*Generic	0.872***	0.435*	1.078***	0.504***	0.944***	0.926***
	(0.075)	(0.223)	(0.153)	(0.077)	(0.161)	(0.207)
Drug Age	-0.020***	-0.019***	-0.025***	-0.012***	-0.012***	-0.024***
	(0.002)	(0.007)	(0.005)	(0.002)	(0.004)	(0.007)
Competition	-0.017***	-0.011***	-0.013***	-0.013***	-0.024***	-0.018***
	(0.002)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)
Drug Strength/1000	0.006	0.002	0.001	-0.000	0.019***	0.036***
	(0.006)	(0.011)	(0.001)	(0.001)	(0.005)	(0.006)
Observations	92,958	17,503	24,478	72,305	27,798	15,008
R-squared	0.661	0.553	0.649	0.268	0.547	0.702
Number of id	9,207	2,843	2,480	7,905	4,487	1,868

Additional variables included but not are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year. Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. For consistency's sake, Column (4) only includes observations with nonzero self-pay contributions (unlike Tables 2-5).

Table A-4. Price Sensitivity by Payer, Fixed Effects, Adjustment for Rebates.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-4.383***	-5.976***	-2.330***	-0.609***	-2.462***	-7.694**
	(0.766)	(1.803)	(0.435)	(0.162)	(0.546)	(3.232)
Generic Status = 1	-2.474***	-1.959***	-2.428***	-0.487***	-2.067***	-3.198***
	(0.237)	(0.393)	(0.308)	(0.074)	(0.375)	(0.566)
NTI*Generic	2.640***	2.148***	2.304***	0.372***	1.959***	4.412***
	(0.328)	(0.565)	(0.338)	(0.091)	(0.414)	(1.128)
Drug Age	-0.058***	-0.073***	-0.042***	-0.009***	-0.014	-0.106***
	(0.012)	(0.021)	(0.009)	(0.003)	(0.011)	(0.036)
Competition	-0.037***	-0.028***	-0.019**	-0.003	-0.031***	-0.053***
	(0.005)	(0.010)	(0.009)	(0.002)	(0.007)	(0.011)
Drug Strength/1000	-0.004	-0.434***	0.001	0.011	0.049***	0.071***
	(0.013)	(0.088)	(0.002)	(0.013)	(0.014)	(0.022)
	0.676	0.722	0.641	0.154	0.764	0.714
Panel B: ln(P) Regressions						
NTI Status = 1	-1.204***	-1.237***	-1.273***	-0.552***	-1.332***	-0.891***
	(0.097)	(0.291)	(0.202)	(0.094)	(0.167)	(0.258)
Generic Status = 1	-1.075***	-0.956***	-1.251***	-0.647***	-1.147***	-1.050***
	(0.062)	(0.132)	(0.127)	(0.062)	(0.143)	(0.117)
NTI*Generic	0.649***	0.212	0.854***	0.281***	0.721***	0.703***
	(0.075)	(0.223)	(0.153)	(0.077)	(0.161)	(0.207)
Drug Age	-0.020***	-0.019***	-0.025***	-0.012***	-0.012***	-0.024***
	(0.002)	(0.007)	(0.005)	(0.002)	(0.004)	(0.007)
Competition	-0.017***	-0.011***	-0.013***	-0.013***	-0.024***	-0.018***
	(0.002)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)
Drug Strength/1000	0.006	0.002	0.001	-0.000	0.019***	0.036***
	(0.006)	(0.011)	(0.001)	(0.001)	(0.005)	(0.006)
Observations	92,958	17,503	24,478	72,305	27,798	15,008
R-squared	0.630	0.525	0.618	0.232	0.513	0.674
Number of id	9,207	2,843	2,480	7,905	4,487	1,868

Additional variables included but not are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year. Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. For consistency's sake, Column (4) only includes observations with nonzero self-pay contributions (unlike Tables 2-5).

Table A-5. Price Sensitivity by Payer, Fixed Effects, Tablets.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-3.010***	-1.714**	-2.837***	-0.554***	-2.885***	-3.761***
	(0.313)	(0.668)	(0.555)	(0.105)	(0.357)	(1.069)
Generic Status = 1	-2.998***	-1.816***	-3.541***	-0.618***	-2.612***	-3.939***
	(0.250)	(0.295)	(0.518)	(0.090)	(0.378)	(0.546)
NTI*Generic	2.967***	1.632***	3.540***	0.516***	2.536***	4.297***
	(0.272)	(0.342)	(0.535)	(0.099)	(0.397)	(0.720)
Drug Age	-0.050***	-0.019*	-0.078***	-0.009***	-0.022***	-0.121***
	(0.007)	(0.012)	(0.013)	(0.002)	(0.007)	(0.027)
Competition	-0.049***	-0.027*	-0.036**	-0.005*	-0.035***	-0.063***
	(0.008)	(0.014)	(0.014)	(0.003)	(0.010)	(0.017)
Drug Strength/1000	2.613***	1.695	3.873***	-0.016	0.142	1.777
	(0.525)	(1.111)	(1.179)	(0.212)	(0.545)	(1.307)
R-squared	0.554	0.557	0.560	0.089	0.587	0.508
Panel B: ln(P) Regressions						
NTI Status = 1	-1.083***	-0.523	-1.009***	-0.477***	-1.525***	-0.582
	(0.127)	(0.450)	(0.267)	(0.115)	(0.194)	(0.398)
Generic Status = 1	-1.125***	-0.794***	-1.265***	-0.699***	-1.289***	-0.965***
	(0.077)	(0.168)	(0.160)	(0.072)	(0.169)	(0.164)
NTI*Generic	0.713***	-0.005	0.943***	0.345***	0.876***	0.568**
	(0.092)	(0.258)	(0.179)	(0.086)	(0.189)	(0.255)
Drug Age	-0.014***	0.005	-0.026***	-0.008***	-0.013***	-0.013
	(0.003)	(0.010)	(0.007)	(0.003)	(0.005)	(0.010)
Competition	-0.027***	-0.030***	-0.025***	-0.016***	-0.027***	-0.027***
	(0.002)	(0.005)	(0.004)	(0.002)	(0.005)	(0.004)
Drug Strength/1000	1.337***	1.523***	2.182***	-0.423**	0.601**	0.977***
	(0.143)	(0.371)	(0.319)	(0.182)	(0.280)	(0.300)
Observations	71,403	11,813	18,180	56,228	25,028	10,497
R-squared	0.611	0.490	0.617	0.167	0.445	0.665
Number of id	6,989	1,939	2,050	6,209	3,870	1,148

Additional variables included but not are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year. Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. For consistency's sake, Column (4) only includes observations with nonzero self-pay contributions (unlike Tables 2-5).

Table A-6. Price Sensitivity by Payer, Adjustment for Rebates, (Tablets only), Fixed Effects.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Payment	Private Contribution	Medicare Contribution	Self Contribution	Self Only Contribution	Medicaid Contribution
Panel A: Price Regressions						
NTI Status = 1	-2.165***	-1.056*	-2.076***	-0.390***	-2.225***	-2.455***
	(0.264)	(0.563)	(0.463)	(0.085)	(0.288)	(0.947)
Generic Status = 1	-2.181***	-1.208***	-2.647***	-0.449***	-1.986***	-2.795***
	(0.203)	(0.243)	(0.419)	(0.072)	(0.304)	(0.448)
NTI*Generic	2.115***	0.996***	2.619***	0.346***	1.896***	3.043***
	(0.221)	(0.287)	(0.433)	(0.079)	(0.320)	(0.595)
Drug Age	-0.041***	-0.014	-0.066***	-0.007***	-0.018***	-0.096***
	(0.006)	(0.010)	(0.011)	(0.002)	(0.006)	(0.022)
Competition	-0.041***	-0.022**	-0.029***	-0.004*	-0.029***	-0.051***
	(0.006)	(0.011)	(0.011)	(0.002)	(0.008)	(0.014)
Drug Strength/1000	2.183***	1.474	3.233***	-0.031	0.181	1.324
	(0.426)	(0.939)	(0.939)	(0.170)	(0.434)	(1.079)
	0.537	0.525	0.548	0.083	0.577	0.493
Panel B: ln(P) Regressions						
NTI Status = 1	-0.860***	-0.300	-0.786***	-0.254**	-1.302***	-0.359
	(0.127)	(0.450)	(0.267)	(0.115)	(0.194)	(0.398)
Generic Status = 1	-0.902***	-0.571***	-1.042***	-0.476***	-1.066***	-0.742***
	(0.077)	(0.168)	(0.160)	(0.072)	(0.169)	(0.164)
NTI*Generic	0.489***	-0.229	0.720***	0.122	0.653***	0.345
	(0.092)	(0.258)	(0.179)	(0.086)	(0.189)	(0.255)
Drug Age	-0.014***	0.005	-0.026***	-0.008***	-0.013***	-0.013
	(0.003)	(0.010)	(0.007)	(0.003)	(0.005)	(0.010)
Competition	-0.027***	-0.030***	-0.025***	-0.016***	-0.027***	-0.027***
	(0.002)	(0.005)	(0.004)	(0.002)	(0.005)	(0.004)
Drug Strength/1000	1.337***	1.523***	2.182***	-0.423**	0.601**	0.977***
	(0.143)	(0.371)	(0.319)	(0.182)	(0.280)	(0.300)
Observations	71,403	11,813	18,180	56,228	25,028	10,497
R-squared	0.579	0.463	0.587	0.134	0.413	0.635
Number of id	6,989	1,939	2,050	6,209	3,870	1,148

Additional variables included but not are substance age, and indicators for Multum therapeutic classes, diagnosis code, mail pharmacies, missing strength, dosage form names, and year. Robust standard errors are reported. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. For consistency's sake, Column (4) only includes observations with nonzero self-pay contributions (unlike Tables 2-5).

Table A-7. Probability of Mixing in a Given Period, Probit.

	(1)	(2)
NTI	-0.005	-0.054
	(0.055)	(0.063)
Previous Mixed Period		0.636***
		(0.059)
Mixed Previous Period*NTI		-0.161*
		(0.090)
Observations	22,467	14,524
Additional variables included but not reported are age, missing age, race, sex, education, and the number of drugs consumed in each individual-substance-period combination. Also included are indicators for Multum therapeutic class, and insurance coverage. Standard errors are clustered by individual. Each observation is an individual-substance-period combination. Each version is defined as a unique product code. Only periods in which more than one drug is consumed are included. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.		

Table A-8 Predicted Probabilities of Mixing, from Table A-13

	(1)	(2)
Non-NTI	0.103***	
	(0.00435)	
NTI	0.103***	
	(0.00615)	
Non-NTI, Previous No Mix		0.107***
		(0.00546)
NTI, Previous No Mix		0.0982***
		(0.00653)
Non-NTI, Previous Mix		0.262***
		(0.0181)
NTI, Previous Mix		0.200***
		(0.0194)
Observations	22,467	14,524
The predicted probabilities in this table can be subtracted to approximate the marginal effects reported in Table A-13. For example, [(NTI, Previous Mix) – (non-NTI, Previous Mix)] –[(NTI, Previous No Mix) – (non-NTI, Previous No Mix)]= -.0532, which is similar to the coefficient on Previous Mixed Period*NTI in Table 6, column (2).		

Table A-9. Version Loyalty: Probit.

	(1)	(2)	(3)	(4)
	No Mix	**Mixed Brand-Generic**	**Mixed Brand**	**Mixed Generic**
NTI = 1	0.065	0.085	-0.429**	-0.129*
	(0.063)	(0.084)	(0.184)	(0.077)
Lagged Consumption Bundles				
Previous Mixed Brand-Generic	-0.842***	1.072***		-0.014
	(0.101)	(0.113)		(0.133)
Previous Mixed Brand	-0.486***	0.553**	0.986***	
	(0.133)	(0.267)	(0.146)	
Previous Mixed Generic	-0.571***	-0.472***	-0.606*	0.694***
	(0.078)	(0.170)	(0.354)	(0.080)
Lagged Consumption Bundles–NTI Interactions				
NTI*Previous Mixed Brand-Generic	0.433***	-0.471***		0.164
	(0.139)	(0.153)		(0.202)
NTI*Previous Mixed Brand	0.227	-0.921***	-0.205	
	(0.205)	(0.338)	(0.214)	
NTI*Previous Mixed Generic	-0.078	-0.268		0.327**
	(0.143)	(0.329)		(0.156)
Observations	14,524	10,662	12,131	11,662

Each column is a separate linear probability model. Additional variables included but not reported are age, missing age, sex, education, and the number of drugs consumed in each individual-substance-period combination. Also included are indicators for Multum therapeutic class, and insurance coverage. Standard errors are clustered by individual. Each observation is an individual-substance-period combination. Each version is defined as a unique product code. Only periods in which more than one drug is consumed are included. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Each column has different observations: for Column (1) only substances more than one version are included. For Column (2), I exclude substances that have no generic versions or no brand versions since this necessarily mixes brand and generic versions. Column (3) only uses substances with more than one brand versions and column (4), only substances with more than one generic version.

Table A-10 Predicted Probabilities from Table 15.

Predicted Probabilities	**Unmixed**	**Mixed Brand-Generic**	**Mixed Brand**	**Mixed Generic**
Non-NTI # Previous Unmixed	0.892***	0.0464***	0.0342***	0.0642***
	(0.00550)	(0.00466)	(0.00799)	(0.00382)
Non-NTI # Previous Mixed Brand-Generic	0.669***	0.262***		0.0626***
	(0.0344)	(0.0352)		(0.0151)
Non-NTI # Previous Mixed Brand	0.781***	0.126**	0.174***	
	(0.0372)	(0.0544)	(0.0375)	
Non-NTI # Previous Mixed Generic	0.757***	0.0163**	0.00907	0.189***
	(0.0229)	(0.00668)	(0.00810)	(0.0190)
NTI # Previous Unmixed	0.903***	0.0550***	0.0137***	0.0506***
	(0.00647)	(0.00496)	(0.00219)	(0.00561)
NTI # Previous Mixed Brand-Generic	0.819***	0.155***		0.0667***
	(0.0244)	(0.0235)		(0.0183)
NTI # Previous Mixed Brand	0.854***	0.0253**	0.0660***	
	(0.0356)	(0.0120)	(0.0212)	
NTI # Previous Mixed Generic	0.753***	0.0103		0.242***
	(0.0346)	(0.00747)		(0.0352)
The predicted probabilities after the probit regression in Table A-15 correspond to the marginal effects listed in the linear probability models. For example, the difference in likelihood of consumed an unmixed bundle is 0.011 between NTI#Previous Unmixed and non-NTI#Previous Unmixed (.903-.892). This corresponds to the coefficient on NTI in Table 5, Column(1). Similarly, NTI#Previous Mixed Brand-Generic – non-NTI# Previous Mixed Brand-Generic = 0.15, similar to the sum of the coefficients on NTI*Previous Mixed Brand-Generic and NTI (0.148+0. 013) in Table 5, Column (1).				

CHAPTER 3

RELINQUISHMENT OF INAPPROPRIATE OFF-LABEL USES: THE EFFECT OF THE FALSE CLAIMS ACT[75]

I. Introduction

The Food and Drug Administration (FDA) does not approve every prescription drug use currently taken by patients. The FDA does not prohibit physicians from prescribing drugs for unapproved uses, a practice called prescribing "off-label." Off-label prescriptions are highly prevalent: almost 50% of cardiac therapy and anticonvulsant prescriptions, and over 30% of anti-asthmatic, allergy therapy, and psychiatric therapy prescriptions, are for off-label uses (Radley et al. 2006). Physicians have considerable autonomy in prescribing drugs off label, as they are free to use their expertise to adopt and relinquish off-label treatments.

Since the legal regime governing off-label drug use affords physicians so much freedom, it relies on physicians to accurately evaluate new information about off-label innovations, adopt beneficial innovations, and relinquish inappropriate ones. For the purposes of this Chapter, an off-label use is "inappropriate" if it is ineffective, unsafe, or both. The risk that an off-label use is unsafe or ineffective, however, is often ambiguous. Since companies do not submit these uses for FDA approval, there is often little published evidence supporting such uses. In a survey of nationally representative drug data, Radley et al. (2006) estimate that only 27% of off-label uses studied were supported by strong scientific evidence.

Prescribing drugs with uncertain risks is not inherently bad; restricting access only to drugs with precisely known risks can harm patients by depriving them of potentially helpful

[75] This Chapter is an extension and revision of a previous working paper under the citation Learning About Ineffectiveness: Physicians' Prescription Decisions Regarding Off-Label Drug Uses (October 20, 2014). Available at SSRN: http://ssrn.com/abstract=2469191 or http://dx.doi.org/10.2139/ssrn.2469191. Parts of unpublished prior work were also incorporated into the draft.

treatments (Abbott and Ayres 2014a; Abbott and Ayres 2014b; Viscusi and Zeckhauser 2015). Ambiguity is also not insurmountable—physicians can update their beliefs about ambiguously risky treatments based on new patient outcomes or scientific information as it becomes available. Appropriate updating, however, depends on physicians using reliable information.

Whether physicians can distinguish between reliable and unreliable information is a real concern. First, physicians receive a lot of information daily, making it is difficult to critically analyze each new piece of information. Second, physicians' reliance on one-on-one interactions with pharmaceutical representatives may outweigh their reliance on scientific studies. The latter point is particularly troubling because pharmaceutical companies have an incentive to encourage off-label prescriptions: off-label sales are highly profitable because pharmaceutical companies do not incur the cost of supplemental FDA approval and are able to reach a larger market.

Given the overwhelming amount of information physicians receive, along with their receptiveness to information from pharmaceutical representatives, do physicians learn when to adopt or relinquish inappropriate off-label treatments? If so, what sources of information are physicians most responsive to?

The importance of these questions is two-fold. First, insufficient relinquishment of inappropriate uses results in monetary waste: insurers end up paying for ineffective treatments and passing those costs to their clients. Second, insufficient relinquishment increases the likelihood that patients receive suboptimal medical care. If physicians fail to relinquish inappropriate treatments, patients lose the opportunity to try more appropriate treatments.

The government has tried to penalize pharmaceutical companies for promoting off-label uses of their drugs through the use of False Claims Act (FCA) lawsuits. The FCA prohibits fraudulent submission of claims to the government for reimbursement (31 U.S.C. §§ 3729, 3730;

Eichel 2011). While the FCA was originally passed in 1863 as a way to prevent profiteering during the Civil War (Eichel 2011), it has recently been used to curb off-label promotion.

This Chapter discusses the possibility that the FCA can serve as a source of information for physicians and third-party payers regarding the appropriateness of an off-label use. This can be done in one of two ways: First, the FCA can bring to attention public, but obscure, scientific information on inappropriate off-label uses. Second, the FCA suit can publicize internal information regarding a pharmaceutical company's fraudulent conduct. The FCA has a whistleblower provision, which awards whistleblowers a portion of the resulting fine. This provides incentives for employees to publicize internal information that might not have otherwise come to light. Thus, the False Claims Act process could act as a source of information for physicians and third-party payers by advertising the inappropriateness of the use.

Despite the potential for the FCA to operate as a source of information, the particulars of its execution can undermine this role. First, if the FCA is used against both inappropriate and appropriate off-label uses, the informational value of the suit might be mitigated. Second, if the FCA requirements for recovering against pharmaceutical companies are too easily satisfied, there might be heterogeneity in response of payers to news of FCA suit. While private payers might incorporate any informational content into their reimbursement decisions, Medicare and Medicaid might not be as responsive. This may be because the FCA provides a source of reimbursement for their loss. Alternatively, Medicare and Medicaid may be less able to monitor or to refuse to reimburse certain uses. While recovery through the FCA addresses the issue of monetary waste that the government experiences, it does not alleviate the medical opportunity cost that a patient could have received better treatment earlier.

This Chapter examines whether the FCA practically functions as a source of information about the appropriateness of an off-label use and measures the effect of FCA settlement on relinquishment of particular off-label uses. Part II gives some background on the regulation of off-label uses and how the False Claims Act is used in this context, and Part III discusses the theoretical model underlying this Chapter's analysis. Part IV empirically examines the issue by studying prescriptions of anticonvulsants and antipsychotics from 2005 to 2010. It finds evidence of relinquishment of off-label uses after FCA settlement. Part V supplements this analysis with a case study of neurontin, one of the first off-label promotion FCA cases to settle. It studies relinquishment in response to scientific and legal information shocks and finds that there is considerable heterogeneity by payer in response to these information shocks, although the response seems to correspond to the FCA suit more than to scientific studies. Part VI concludes.

II. The Current Landscape of Off-Label Drug Use

The regulation of off-label uses of drugs is as important for patient care as it is complex. This section defines and discusses the issues surrounding off-label usage and then discusses the mechanism the government primarily uses to regulate it, the False Claims Act.

A. The Importance of Appropriate Off-Label Usage

The FDA studies the safety and effectiveness of prescription drugs for specific diseases ("indications"), dosages, and populations through several phases of controlled study (FDA 2012). Not all drug uses go through this process. The practice of prescribing a drug for indications, dosages, and populations for which it has not undergone FDA approval is called prescribing "off label" (Wittich et al. 2012). Appropriate off-label drug use requires that physicians incorporate new scientific evidence about the safety and effectiveness of drugs into their prescription decisions. The FDA does not regulate physicians' off-label drug prescriptions

but instead places restrictions on manufacturers' advertisements of off-label uses (FDA 2009; FDA 2014). Pharmaceutical manufacturers that violate these restrictions often face False Claims Act suits, which can result in large monetary penalties. Examples of such offenders include Warner Lambert, which illegally marketed neurontin for the treatment of bipolar disorder (DOJ 2004), and Eli Lilly, which marketed zyprexa for the treatment of dementia (Fisk, Lopatto, and Feeley 2012).

Physicians are allowed to prescribe drugs for off label uses. Prescription practices are limited by medical malpractice liability: physicians are liable for any deviations from the "acceptable and prevailing standard of practice" in prescribing drugs off label, just as they are for on-label drugs (Riley & Basilius 2007, p.27). Physicians prescribing drugs off label are not presumed to be negligent; in fact, several off-label drugs uses are so prevalent that they are considered the standard of care (for example, aspirin for coronary disease prophylaxis) (Wittich et al. 2012).

Given that off-label drug uses do not go through FDA-mandated systematic scrutiny for the off-label use, some are concerned that these uses are not medically appropriate. As noted earlier, Radley et al. (2006) find that only 27% of off-label uses are supported by strong scientific evidence.[76] The lack of publicly available scientific evidence does not necessarily mean that off-label treatments are inappropriate. Suggestive scientific evidence can provide support for off-label treatments. If an off-label drug belongs to the same class as a drug already approved for the indication, physicians may expect the drug to perform similarly (Stafford 2008).

[76] Radley et al. (2006, p.1022) clarify that "An indication was considered to be scientifically supported if, according to DRUGDEX, its effectiveness has been shown in controlled trials or observed in clinical settings. All other indications that lacked FDA approval or that did not meet the criteria for having scientific support were considered to be off-label with little or no scientific support."

Similarly, if two indications have similar symptoms, physicians may expect a treatment for one indication to relieve symptoms in the other indication. Such evidence is merely suggestive, however; without systematic study, physicians do not have any assurance of the appropriateness of a drug for an off-label use.

Even if such systematic studies were available, it is unclear whether physicians would find them and whether they would be the predominant information source. There is a large literature on physician information needs and their ability to find relevant information. Covell et al. (1985) survey physicians regarding their information needs and find that physicians cited insufficient time as the most frequently reported barrier to finding necessary information. Ely et al. (2005) also find that insufficient time was a reason physicians only looked for answers to about 55% of the questions that they had. In a survey of physicians regarding their information needs, Williamson et al. (1989) find that physicians have difficulty locating appropriate studies to resolve questions of treatment choice. Even if physicians locate the study, 87% of the polled physicians assess study validity by comparing the results to their own experiences, rather than by evaluating the study's methodology.[77] A recent study found more optimistic results. Kesselheim et al. (2012) give 503 physicians three journal abstracts each and ask them to evaluate the methodological rigor of each. They also asked for the likelihood that, based on the evidence, physicians would prescribe a given drug. 269 physicians responded, and most seemed able to assess the relative rigor of the studies. Moreover, physicians seemed to discount studies that received funding by pharmaceutical companies. There may be some concern that this is not a representative finding, however, given the low response rate.

[77] Johnson (2007) provides an in-depth discussion of the issues regarding physician learning.

"Pharmaceutical detailing," where pharmaceutical representatives personally promote a drug to a physician, is another source of information for physicians' treatment decisions (Chan et al. 2013; Chintagunta, Goettler, and Kim 2012). This line of literature sees pharmaceutical detailing as a legitimate avenue of information dissemination, even though legal promotion of off-label uses is limited, and fraud is frequent. Pharmaceutical companies are frequently fined for distributing false or misleading evidence about off-label uses of their drugs (DOJ 2004; Fisk, Lopatto, and Feeley 2012).[78] It is unclear whether or not physicians anticipate this difference in information value.

If physicians do encounter reliable information on off-label drug use, the process by which they stop prescribing an inappropriate use is called "relinquishment." Previous literature suggests that there is substantial heterogeneity in relinquishment. In a study of an unsafe drug in the British market, Mapes (1977) finds that after journals published news of the drug's adverse effects, physicians did not reduce their prescriptions uniformly. Physicians were less likely to relinquish if the physician started practicing while the drug was still considered a good treatment or if the physician attended fewer post-graduate medical courses. Additionally, physicians who are more likely to consider a patient's social surroundings and environment continued to prescribe the drug. In contrast, Majumdar et al. (2001) finds that relinquishment is not different between generalists and specialists.

This study extends this relinquishment literature to off-label uses. Off-label drugs may not experience the same pattern of relinquishment as on-label drugs because the FDA does not

[78] While pharmaceutical representatives were originally unable to promote off-label uses legally, the FDA has created "safe harbors" of activities that would not be considered illegal off-label promotion (FDA 2009; FDA 2014). Given the current First Amendment litigation over off-label promotion as protected commercial free speech (Robertson 2014; Greene 2014; Philip 2014), the realm of legal off-label promotion is undefined now, but presumably larger.

issue industry-wide warnings or notices for off-label uses of drugs. This might suggest that another visible, quasi-regulatory notice, such as litigation, is necessary. This Chapter discusses whether such a signal seems to be effective and under what circumstances such a signal might even be desirable.

B. Regulating Off-Label Uses through the False Claims Act

Given the potential for information regarding off-label uses to be unreliable, the government has sought to reduce the freedom with which pharmaceutical companies can promote off-label uses. The government first tried to make off-label promotion illegal per se under the Food, Drug, and Cosmetic Act (FDCA), labeling it a misbranding violation.[79] This pathway has recently been attacked on First Amendment grounds. In a recent case, *United States v. Caronia*,[80] the Second Circuit ruled that truthful off-label promotion was protected as commercial free speech. Critics have questioned whether the court was warranted in granting such protection without requiring pharmaceutical companies to prove that their speech was not false or misleading (Robertson 2014); others acknowledge the potential pitfalls of having a broad truthfulness definition but suggest that requiring a significantly narrower definition indirectly through speech constraints might be too burdensome, especially subject to criminal sanctions (Philip 2014). Others suggest that the government should specifically allege that off-label promotion is false or misleading (Greene 2014). *Caronia* was not the final word on First Amendment protections of off-label promotion; in 2015, the Southern District of New York

79 21 U.S.C. § 331 (a)–(c) (2012) ("(a) The introduction or delivery for introduction into interstate commerce of any food, drug, device, tobacco product, or cosmetic that is adulterated or misbranded. (b) The adulteration or misbranding of any food, drug, device, tobacco product, or cosmetic in interstate commerce. (c) The receipt in interstate commerce of any food, drug, device, tobacco product, or cosmetic that is adulterated or misbranded, and the delivery or proffered delivery thereof for pay or otherwise.").

80 United States v. Caronia, 703 F.3d 149 (2d Cir. 2012).

faced a similar case (Gibbons 2015). The FDA tried to distinguish *Caronia* by saying that "it does not read *Caronia* to preclude a misbranding action where the acts to promote off-label use consist solely of truthful and non-misleading speech, provided that the evidence also shows that the drug had been introduced into interstate commerce and that the FDA had not approved it as safe and effective for the off-label use."[81] The District Court rejected this argument, further suggesting that truthful speech alone cannot be used as a predicate action for misbranding. Much has been written about the validity of First Amendment challenges in this context; the state of this law is still in flux, but the future use of misbranding to limit off-label promotion is uncertain. For these reasons, this Chapter will focus on the second, more lucrative way of policing off-label uses: the False Claims Act.

The False Claims Act (FCA) was enacted during the Civil War in order to enable the government to recover losses from fraud; recently, however, it has been used to punish promotion of off-label uses (31 U.S.C. §§ 3729, 3730; Blair 2010). FCA liability is triggered when someone "knowingly presents, or causes to be presented, a false or fraudulent claim for payment or approval."[82] The FCA is particularly unique in its qui tam provision, which allows a whistleblower with knowledge of fraud[83] to bring suit on behalf of the government (Eichel 2011). The whistleblower, in turn, receives a percentage of the resulting penalty. The qui tam provision results in two benefits: First, it reduces the resources the government has to spend on litigation, since the government receives a cut of the award regardless of whether or not it

[81] Amarin Pharma, Inc. v. U.S. Food & Drug Admin., No. 15 CIV. 3588 PAE, 2015 WL 4720039, at *22 (S.D.N.Y. Aug. 7, 2015).
[82] 31 U.S.C. § 3729.
[83] 31 U.S.C. § 3730.

chooses to intervene,[84] Second, whistleblowers, especially employees or insiders, are able to provide internal information about fraud (Eichel 2011).

Recently, the FCA has been used to sanction pharmaceutical companies for off-label promotion. The government seeks to recover for prescriptions submitted to Medicaid and Medicare for reimbursement when the pharmaceutical company promotes off-label uses (Blair 2010). FCA suits are highly lucrative and increasingly prevalent;[85] however, the theory of liability in the off-label promotion context is not intuitive. There are at least two possible theories of liability: First, a pharmaceutical company can be liable for making a *false* claim about the safety and effectiveness of an off-label use, which induces the physician to submit the claim for reimbursement from the government (Eichel 2011). Second, the *Franklin v. Parke Davis* court suggested that *truthful* information about an off-label use can also be the basis for FCA liability as long as the off-label promotion induces claims that are ineligible for reimbursement to be submitted to the government (Eichel 2011).[86]

This second theory is problematic, especially as it possibly triggers the same First Amendment issues as misbranding. If restricting truthful off-label promotion violates the First Amendment in the misbranding context, does it violate the First Amendment in the context of the False Claims Act?[87] The government has contended that these two areas are distinct. Rogoff, Mayell, and Ramer (2014) note that the government filed a Statement of Interest in *U.S. ex rel.*

---

[84] 31 U.S.C. § 3730 b(1), (d). Notably, there may be some difference between cases in which the government chooses to intervene or not. The differential effect of such cases might be interesting to explore in the future.

[85] From 2004 to 2010, twenty-one off-label marketing cases have allowed the government to recover $7.9 billion in criminal fines and civil settlement (Blair 2010).

[86] U.S. ex rel. Franklin v. Parke-Davis, Div. of Warner-Lambert Co., No. CIV.A. 96-11651PBS, 2003 WL 22048255, at *2 (D. Mass. Aug. 22, 2003).

[87] Greene and Noah (2014) also questions whether the *Caronia* decision will affect FCA prosecutions, as "whistleblowers have repeatedly pointed to off-label promotion as a basis for triggering prosecution even where the FDA later approved some of these uses."

Cestra v. Cephalon, Inc., distinguishing it based on the fact that the FCA "prohibits any conduct that causes the submission of false claims to the government. . ." Rogoff, Mavell, and Ramer note that "[a]ccording to the government, even if that conduct is carried out through truthful speech — the same speech that Caronia holds may be constitutionally protected under the FDCA — FCA liability could still attach." It is unclear how persuasive this argument actually is.

Notwithstanding the First Amendment concerns with the FCA, there are other issues with using the FCA to curb inappropriate off-label drug use. The FCA is a very broad statute, made even broader by the enactment of the Fraud Enforcement and Recovery Act (FERA) (Eichel 2011). After the Supreme Court tried to limit the scope of the FCA through *Allison Engine v. Sanders*,[88] Congress passed FERA to expand the scope of the FCA. In particular, FERA removed the requirement that the government "establish a direct link between a false statement and the eventual government payment of the claim" (Eichel 2011 p. 428). FERA also clarified the materiality element of the FCA, defining materiality as "having a natural tendency to influence, or be capable of influencing, the payment or receipt of money or property."[89] While this did establish that materiality is required for an FCA claim, the "natural tendency" standard has been criticized for being too vague and too broad (Hoffman 2009).

Given this background, it is possible that FCA may not actually serve as a source of information for physicians and third-party payers. First, given the *Franklin v. Parke Davis* interpretation of FCA requirements, an FCA claim may be able to be brought based on truthful

[88] Allison Engine Co. v. United States ex rel. Sanders, 553 U.S. 662 (2008).
[89] 31 U.S.C. § 3729 b(4).

statements. If this is true, the FCA can be equally enforced against appropriate and inappropriate off-label uses, making the signal value of the FCA suit less informative.[90]

Even if the FCA claim was predicated on a false claim, the legal standards for bringing the claim are important. Given the expansion of the FCA, the vague definition of materiality, and the high stakes for continuing litigation, manufacturers may settle cases in which the off-label uses are not actually medically inappropriate. This also would reduce the informational value of the suit. Similarly, the legal standards may determine public third-party payers' incentives to prospectively screen out inappropriate off-label uses for reimbursement. Because of the relative ease with which an FCA claim can be brought—and generally settled—the government may be less sensitive to distributing information on off-label appropriateness because they know they can recoup their monetary losses through suit. On the other hand, the legal standards of the FCA may be a result of the Medicare's lack of flexibility to monitor drug usage. Medicaid and Medicare may also be statutorily less able to revise their reimbursement policies than private payers.

The rest of the Chapter empirically tests whether settlement of FCA suits actually leads to relinquishment of off-label uses. The next Section outlines the theoretical model underlying the role of the FCA as a source of information.

III. Theoretical Model

The following analyses rely on a very basic drug choice model. In the basic supply side treatment choice model (Chandra et al. 2011), physicians make prescription choices based on the

[90] Presumably, physicians would catch on to this weakening signal if off-label uses that they know are well-established or successful become subject to the FCA. Pharmaceutical companies could also build this narrative by focusing on technical violations of the FCA and suggesting that the government is simply greedy.

patient's perceived utility and any net benefit to themselves. Assuming that the physician is a perfect agent of the patient, utility[91] can be formalized as

(1) $U_{kpi} = B_{ik}(\sigma_i, I) + (Y_i - P_k) + \theta_p$

for treatment choice *k* by physician *p* for patient *i*. B_{ik} denotes expected patient benefit and is a function of patient characteristics, σ_i, and the information the physician receives, *I*. θ_p is a physician-specific random error.

A physician compares (1) for all drugs. She prescribes drug j when $U_j > U_k$ for all k, choosing drug j when

(2) $(B_{ij} - B_{ik}) + [(Y_i - P_j) - (Y_i - P_k)] > 0.$

Physicians estimate B_{ik}, the patient's benefit, which is a function of patient characteristics, σ, and available information, I. Y_i is the individual's income and P_j is the price of drug j. I is a function of information shocks and physician learning patterns. The Bayesian learning process, *I*, can be formalized as

$I_k = \frac{\gamma\rho + \xi s(m_k)}{\gamma + \xi}$ where ρ is the prior belief of the drug's appropriateness and $s(m_k)$ is the risk implied by new information given by information shock m_k. The weights on the priors and new information are denoted by γ and ξ, respectively.[92]

The Chapter hypothesizes that the FCA functions as a source of information for physicians and payers, essentially as m_k. As stated previously, the FCA can provide new

[91] Chandra et al. (2011) assume that the physician has perfect knowledge. I add the *I* element to symbolize learning by the physician.
[92] Viscusi (1997) discusses this Bayesian learning process in the context of ambiguity.

information in one of two ways: the FCA suit can publicize already-existing but obscure scientific information or can uncover new information through the whistleblower provisions. Specifically, the Chapter hypothesizes that after settlement of an FCA claim, the number of inappropriate off-label uses should decline.

Second, as noted in Section 0, if different payers are incentivized to be sensitive to such sources of information to varying degrees, the Chapter hypothesizes that relative relinquishment of inappropriate off-label uses should differ by payer. Payers can influence treatment decisions in two ways: 1) they can refuse to reimburse a treatment (or as a less extreme option, can place the drug on a more expensive tier in their formulary) or 2) they can influence treatment by making "preferred" treatment guidelines. The former can affect the treatment decision by making the treatment marginally more expensive and less attractive. The former can also serve as a signal from the payer that they do not value a particular use. The latter allows the payer to be another source of information for physicians, such that it can persuade physicians that a use is inappropriate through research conducted by its PBM.

The rest of the Chapter tests these hypotheses. Section IV examines the average effect of relinquishment over numerous FCA settlements. Section V examines the FCA suit over neurontin as a case study to understand the intricacies of the relinquishment process in a particular circumstance.

IV. Average Effect of FCA Suits on Relinquishment

As discussed in Section 0, the FCA has the potential to be an information source for physicians and third-party payers in order to relinquish inappropriate off-label uses. This potential signal, however, may be dulled by the broad execution of the statute. This section examines the average relinquishment effect caused by FCA settlements over multiple suits. This

has the advantage of not overly relying on the timeline of one drug; instead, it estimates the average effect of FCA settlement on relinquishment for a particular subset of drugs. This Section finds that there is a significant decrease in off-label prescriptions associated with the settlement of FCA claims.

A. Data

This section uses data from the Medical Expenditure Panel Survey (MEPS) 2005–2010.[93] The U.S. Department of Health and Human Services publishes MEPS, drawing from a nationally representative subsample of households participating in the prior year's National Health Interview Survey (NHIS). MEPS surveys families and individuals, their medical providers, and employers about medical expenditures and provides information on demographic characteristics, health conditions, insurance, medical expenditures, and sources of payment.

The MEPS data are linked to FDA data by National Drug Code (NDC), [94] in order to incorporate larger "substance" groups that incorporate generic and brand-name versions of the same drug. Using data from the Full Year Consolidated File subcomponent of MEPS, I determine whether the person buying the drug expects prescription drug coverage from a particular source, including Medicare, Medicaid, private insurance, Veteran's Affairs, state and local coverage, other, unknown, or no drug coverage.

In order to capture years in which no prescriptions were purchased, I merge the year–coverage–diagnosis code categories with a full-factorial matrix of year–coverage–off-label categories for each substance and impute categories with zero prescriptions. I drop observations

[93] The years 2005–2010 were chosen because MEPS provides a variable indicating "expected payer for prescription drug" during this period.
[94] This data was downloaded in June 2015. The FDA NDC database can be found at http://www.fda.gov/Drugs/InformationOnDrugs/ucm142438.htm.

that have no prescriptions in any substance–coverage–off-label status over any years. This leaves 2,898 substance–year–coverage–off-label status observations.

This study focuses on two classes of drugs: anticonvulsants and antipsychotics.[95] These classes were chosen because there were multiple anticonvulsants and antipsychotics subject to FCA suit for off-label promotion during the studied time. The following drugs were studied as those subject to FCA suit during this period (with the date of settlement in parentheses): abilify (2007), gabitril (2008),[96] zyprexa (2009), geodon (2009), and lyrica (2009).[97] While MEPS provides therapeutic class categories, these may not be stable over a long period of time. In order to accommodate such fluctuations, I follow a group of drugs considered anticonvulsants or antipsychotics in 2008 over time. This ensures that I see a stable group of anticonvulsants and antipsychotics agents.

Finally, relinquishment should spur relinquishment of off-label uses. In order to measure this, I have to determine which uses are off-label. I do this by using the drug compendium DRUGDEX. DRUGDEX lists the uses of the drug that are approved by the FDA. Each listing was then matched to an ICD-9 code.[98] Then, the ICD-9 codes provided by MEPS for each prescription purchase are linked to the matched ICD-9 code. Any purchase whose ICD-9 codes do not match the ones implied by the approved uses is considered an off-label treatment for the purposes of the study.

---

[95] These classes were delineated according to Multum Lexicon Class.
[96] Gabatril was not found in my data.
[97] A few drugs experienced a settlement in 2010 (seroquel, topamax, trileptal, and zonegran); however, since the data stops at 2010, I cannot estimate a *Settlement* effect for these. These drugs were found by searching the Department of Justice archives under "False Claims Act" and "Off-label." I also looked at cases listed in Westlaw to ensure I found all the settled cases.
[98] This mapping is available upon request.

Settlement should only spur relinquishment for *inappropriate* off-label uses. An off-label use that is known to be appropriate should not be relinquished following settlement. This analysis cannot yet distinguish between appropriate and inappropriate uses. However, the analysis uses the current DRUGDEX database to code off-label uses, meaning that some of the current on-label uses may have been off-label at the time of prescription. This means that the uses coded as off-label are those that still had not received approval several years later, suggesting that the flagged off-label uses are marginally more likely to be inappropriate uses than if concurrent information was used.

B. Differences-in-Differences Analysis

This section measures the effect of an FCA claim settlement on the number of prescriptions of a drug. It does this by comparing the number of prescriptions for drugs that are subject to an FCA claim to those which are not, before and after settlement, a technique called difference-in-differences. Since FCA settlement date varies for each drug, I create an indicator variable, *Settlement*, which takes the value of one in the period after settlement for each drug subject to such suit (e.g., after 2007 for abilify; after 2008 for gabatril; and after 2009 for zyprexa, geodon, and lyrica). Specifically, for drugs that were subject to FCA settlement during the period, *Settlement* is zero before settlement and is one after settlement. For drugs that were not subject to FCA settlement during this period, *Settlement* is always zero. The following equation is estimated:

(3) $\#Prescriptions_{dpwy} = \gamma_1 Settlement_{dpwy} + \gamma_2 OffLabel + \gamma_3 Settlement_{dy} \times OffLabel + \varphi_y + \omega_d + \varepsilon.$

The dependent variable is the number of prescriptions, by year y, payer p, off-label status w, and drug d. *Settlement* is defined above and is considered the difference-in-differences estimator.

$OffLabel$ is an indicator variable for whether the use was considered an off-label use. Year indicator variables are included in φ_y, controlling for periods before and after FCA settlement, and drug indicator variables are included in ω_d.

The coefficient γ_1 captures the effect of *Settlement* on on-label prescriptions. I hypothesize that this should be zero, as the FCA settlement provides no new information on already-approved uses. The coefficient of interest is γ_3, which should capture the effect of *Settlement* on off-label prescriptions relative to on-label prescriptions. I expect $\gamma_3 < 0$, which would indicate that the relinquishment effect from FCA settlement is stronger for off-label prescriptions than for on-label prescriptions.

Table 1 reports the results of the average effect of FCA settlement on relinquishment. Columns (1) and (2) stratify by drug, year, off-label status, and payer. Column (1) uses the number of prescriptions while Column (2) uses the ln(prescription +1).[99] The coefficient on *OffLabel* is negative, indicating that as a whole, off-label uses are less prevalent than on-label uses. *Settlement* signifies the drop in on-label prescriptions of drugs subject to FCA suits after settlement. This coefficient is positive and significant. This coefficient was originally hypothesized to be zero; however the significance might be explained by companies anticipating losing revenue on off-label uses and increasing detailing efforts for on-label uses.

[99] Ln(Prescription + 1) is used because some periods/observations have zero prescriptions. These periods are important to preserve to accurately understand the fluctuations in prescriptions.

Table 1. Number of Prescriptions by Payer and Year, 2005–2010: OLS Regressions.

	Differential Treatment Effect		Triple Differences	
	Prescriptions	ln(Prescriptions+1)	Prescriptions	ln(Prescriptions+1)
Settlement	2.471e+05***	8.893e-01*	2.583e+05***	1.063e+00**
	(8.319e+04)	(5.096e-01)	(8.400e+04)	(5.314e-01)
OffLabel	-2.283e+05***	-1.357e+00***	-2.004e+05***	-1.793e+00***
	(1.648e+04)	(1.715e-01)	(3.602e+04)	(4.322e-01)
Settlement* OffLabel	-3.391e+05***	-1.173e+00**	-3.654e+05***	-1.559e+00***
	(8.767e+04)	(4.736e-01)	(8.941e+04)	(5.872e-01)
Indicators for Substance	Yes	Yes	Yes	Yes
Indicators for Year	Yes	Yes	Yes	Yes
OffLabel-FCA Interactions			Yes	Yes
OffLabel-Year Interactions			Yes	Yes
Observations	2,898	2,898	2,898	2,898
R-squared	0.406	0.446	0.407	0.447

Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Variables included but not shown are indicators for off-label status, Multum Therapeutic class, year, and substance. For Columns (3) and (4), FCA is an indicator variable that takes a value of one for any drug that was subject to FCA suit during this period. Controls for coverage type are also included but not shown.

The interaction term *Settlement*OffLabel* represents the effect of settlement on off-label prescriptions, relative to the effect of settlement on on-label prescriptions. This effect is significantly negative, indicating that settlement reduces off-label prescriptions relative to its effect on on-label uses. Additionally, the magnitude on *Settlement*Off-Label* exceeds that of *Settlement*, suggesting that *Settlement* has a negative effect on off-label prescriptions (for column (1), a net effect of around -92,000 prescriptions). The average number of prescriptions per period is 252,747 prescriptions, indicating that the magnitude of this net effect is economically significant. The results in Column (2) are similar.

Columns (3) and (4) extend the main model into a triple difference analysis by including interactions between *OffLabel* and substance and between *OffLabel* and year. To preserve degrees of freedom, instead of calculating an interaction between off-label status and each substance, I create a dummy variable, *FCA*, which takes the value of one for any drug that was

ever subject to an FCA suit during this period. *FCA* is then interacted with off-label status. The results are robust to these inclusions.

In sum, the overall results suggest that for anticonvulsants and antipsychotics that underwent FCA settlements between 2005–2010, there is a significant average relinquishment of off-label uses. Since this analysis estimates the relinquishment effect over multiple drugs, the effect is unlikely to be dependent on one idiosyncratic drug experience.

While these results do lend credence to the idea that FCA claims serve as a source of information, a couple of questions remain. As noted earlier, this analysis does not take into consideration differences in appropriate and inappropriate off-label uses. If the off-label use is medically appropriate and supported by evidence (but just not subjected to FDA approval), FCA suits should not lead to relinquishment of that use. It is possible, however, that if a use is medically appropriate but there is not a formal scientific study to document safety or efficacy, the FCA suit may be sufficient to cast doubt on all ambiguous off-label uses. In this case, the uniform relinquishment might be more reasonable.

This analysis additionally only considers the effect of settlement on relinquishment, not any other visible landmarks in the FCA case or any scientific information shocks. It is possible that other landmarks, such as press coverage or unsealing of the case could be earlier visible information shocks. Similarly, new scientific publications might be expected to spur relinquishment. This analysis does not currently take into account differential reactions by payer. Additionally, some of the previous literature on relinquishment suggests that relinquishment may happen at different times for different actors (Mapes 1977). This analysis does not currently allow for this heterogeneity.

The next section addresses some of these concerns by choosing a particularly focal FCA claim. It attempts to answer some of these questions by studying relinquishment of an ineffective use based on both scientific and legal information shocks. It additionally allows for heterogeneous responses by different third party payers.

V. Neurontin: A Case Study

The previous section estimated an average effect of relinquishment over multiple cases. While this does lend more confidence to the idea that the estimated effect is driven by FCA settlement, it makes understanding heterogeneity in relinquishment more difficult. This section fills that gap by allowing for not only legal information shocks but also scientific sources of learning. It also allows relinquishment behavior to vary based on payer type. This section examines a focal off-label promotion case, that of neurontin. Neurontin was the subject of a False Claims Act (FCA) lawsuit, which settled in 2004. Neurontin was marketed for many off-label uses, one of the most important being for bipolar disorder. Unlike the off-label uses in the prior section, in which appropriateness is not measured, neurontin was not found to be effective for the treatment of bipolar disorder in a study conducted by its own manufacturer, Warner Lambert (Pande et al. 2000). This section studies prescriptions of neurontin for the treatment of bipolar disorder to cleanly measure relinquishment of an "inappropriate" off-label treatment. A few studies have studied neurontin prescription patterns for publicly-funded programs (Kesselheim 2011; Fullerton 2010); this Chapter expands this analysis to allow for heterogeneity by payer.

Using data from the National Ambulatory Medical Care Survey (NAMCS), this Section measures the responsiveness of physicians to news of neurontin's ineffectiveness in treating bipolar disorder. The results suggest that physicians do not uniformly relinquish the drug after

scholarly news of its ineffectiveness. Instead, the Chapter finds that a patient's payment method affects the likelihood of whether the patient is prescribed the disfavored drug, with patients with private insurance being less likely to receive neurontin after 2002. These results suggest that prescriptions may not be as sensitive to scientific data as society might like and that some payers relinquish treatment earlier than others.

A. Testing the Relinquishment Hypotheses

Patient-level records from the National Ambulatory Medical Care Survey (NAMCS) from 1998–2008 are used to estimate the determinants of bipolar treatment choices. Each year, the NAMCS sample includes around 3,000 physicians and samples the eligible physicians' patient records. The CDC surveys nonfederally employed physician offices engaged in "office-based, patient care."

Dependent Variable: Choice of Drug. Physicians have several options for bipolar disorder treatment. In the 1970s, lithium was discovered to be an effective mood stabilizer (NIMH 2013). Anticonvulsant drugs are also generally effective mood stabilizers. In particular, depakote was FDA-approved for the treatment of bipolar in 1995. Several other anticonvulsant drugs used off label for bipolar were used as mood stabilizers and later approved by the FDA for the treatment of bipolar disorder. Lamictal was approved for bipolar disorder on June 20, 2003 (FDA 2010b), equetro was approved in December 10, 2004 (FDA 2004), and stavzor on July 29, 2008 (FDA 2008).

"Lithium"=1 if lithium is prescribed as one of the medications for the visit. Similarly, "approved anticonvulsant"=1 if depakote, tegretol, lamictal, equetro, stavzor, epitol, or depakene[100] (or their generic versions) are prescribed during the visit, the reason for the visit is

[100] Epitol is has the same generic as tegretol and depakene has the same generic as stavzor.

not related to convulsions, and none of the diagnoses are for epilepsy. These exclusions ensure that these drugs are prescribed for the bipolar disorder, not a concurrent epilepsy problem. The "neurontin" variable is constructed similarly.

The number of prescriptions for each of these drugs are summarized in the upper half of Table 2. Since these are simply indicator variables, each mean represents the percent of the sample that received each drug. Each patient could be prescribed multiple drugs, and "other" drugs are not listed here. Thus, the percentages should not add to 1. Neurontin seems to be one of the least prevalent drugs, with the exception of carbamazepine. Lithium is the most prevalent.

Table 2. Descriptive Statistics, 1998-2008

Variable	Mean	Std. Dev.
Dependent Variables		
Neurontin	0.066	0.249
Lithium	0.201	0.401
Lamictal	0.135	0.342
Valproic	0.212	0.409
Carbamazepine	0.040	0.195
Payer:		
Private	0.455	0.498
Medicare	0.133	0.340
Medicaid	0.146	0.353
Self	0.170	0.376
Other	0.096	0.295

Information Shocks. There are 3 major information shocks in the neurontin scandal, some scientific and some legal. These are summarized in Table 3. Figure 1 plots the prescriptions of neurontin, lithium, and approved anticonvulsants (as a percent of the total neurontin, lithium, and approved anticonvulsant prescriptions per year) over time, with dashed lines marking each information shock. The first shock involved a journal article regarding neurontin's ineffectiveness for bipolar disorder, which was published in 2000 by Warner-Lambert itself. In 1998, Warner-Lambert conducted a study, which revealed that neurontin was

not effective for bipolar disorder, but the company did not publish the study until 2000 (Lenzer 2004; Pande et al. 2000) [hereafter, the "Pande study"]. Since the study was conducted by the manufacturer and still reported a negative finding, it should have a large, negative effect on physicians' prescriptions of neurontin. Moreover, this study was more rigorous than previous studies—it was one of the first randomized controlled trials ("RCT") to be conducted (Williams, et al. 2009). Another randomized controlled study was published later in 2000 (Frye, M.A. et al. 2000) [hereafter "the Frye" study]. While several randomized controlled studies were subsequently published, one in 2002 (Obrocea et al. 2002) and one in 2006 (Vieta et al 2006), this study should be most probative because of its novelty, its authorship, and scientific rigor.

Table 3. Information Shock Hypothesized to Lead to Relinquishment.

Information Shock	Date
Warner-Lambert study is published, finding that a placebo outperforms neurontin in treatment of bipolar disorder.	2000
News of the False Claims Act becomes publicized.	2002
Pfizer settles False Claims Act.	2004

Litigation provides the next 2 information shocks. The media began to report on the suit in 2002.[101] Several NPR pieces and other media outlets carried this news (Prakash 2002a, Prakash 2002b, Purse 2012). During litigation, various internal documents showed Warner-Lambert's efforts at promoting neurontin despite no evidence of effectiveness (DOJ 2004).[102]

[101] The lawsuit was not filed at this time but had been kept under seal until 1999. U.S. ex rel. Franklin v. Parke-Davis, Div. of Warner- Lambert Co., 147 F. Supp. 2d 39, 46 (D. Mass. 2001). The first opinion did not come out until mid-2001 but the news began reporting on the suit in 2002. In October 2002, the District of Massachusetts allowed the medial to see nonprivileged documents produced in discovery. U.S. ex rel. Franklin v. Parke-Davis, 210 F.R.D. 257 (D. Mass. 2002).

[102] This litigation was accompanied by another suit, a class action suit, filed against Pfizer in 2002 alleging that the company engaged in off-label promotion and sham patent litigation to retain market exclusivity (Longstreth 2014). Though this is a separate suit, the allegation of off-label promotion is the same as the allegation in the False Claims Act suit and the patent litigation should not affect physician decisions if price does not actually change.

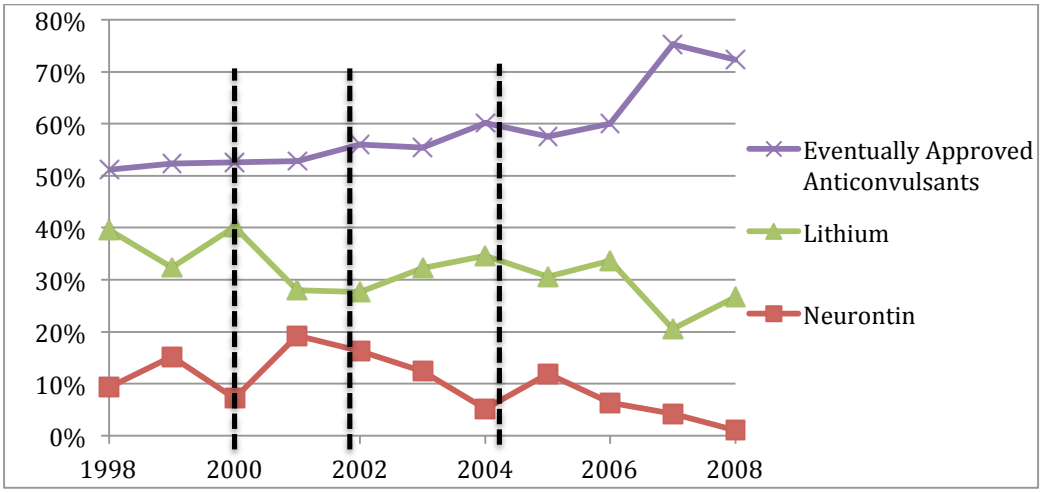

Figure 1. Percent Prescriptions by Year

Notes: The data were obtained from the NAMCS. The percent is the number of prescriptions for each drug divided by the sum of neurontin, lithium, and eventually approved anticonvulsants for the year. The information shocks are denoted by dashed lines: the Pfizer article was published in 2000, news of the litigation broke in 2002, and the litigation settled in 2004.

The final information shock occurred in 2004, when the litigation settled. Pfizer paid $430 million to settle its criminal charges for illegal and fraudulent promotion of off-label uses of neurontin.[103] This is a very visible landmark, as neurontin was one of the first FCA cases based on off-label promotion.

These information "shocks" are not without complications. Since they are mostly identified by year, I list possible concurrent events that might influence relinquishment: One

[103] Another information shock regarding neurontin in general occurred in 2005. On April 22, 2005, Pfizer and the FDA issued a voluntary recall from the manufacturer for 40,000 bottles of capsules distributed in October and November of 2004 because an error in production resulted in empty or partially filled capsules (FDA 2005). Since this date is so close to the litigation date and since this is a national alert, I cannot disentangle the effect of the litigation and the recall. It is possible that physicians prescribing neurontin for bipolar disorder were sensitive to this information shock. One could argue that the recall shook physicians' trust in Warner-Lambert so that they decrease their prescriptions. However, physicians are more likely to distrust a manufacturer more for its fraudulent promotion than for a mechanical error. The former would seem to breed more lasting fears and actually change prescribing behavior long term. Additionally, such recalls are fairly prevalent and physicians should not be too sensitive to them. In 2005 alone, there were 108 safety alerts of drugs and therapeutic biological products (FDA 2010a).

concurrent event that complicates the analysis is that, as previously noted, a third RCT was also published in 2002, which found that a placebo outperformed neurontin in treating bipolar disorder. I am unable to control for this, since it also occurred in 2002. It is possible that this was the extra information necessary to catch physicians' attention. However, this seems unlikely for several reasons: first, the 2002 RCT does not seem to be cited by review studies – of the seven review studies published between 2003–2005 (after the third RCT and before the fourth RCT was published), each cited the two RCTs published in 2000, but not the RCT published in 2002.[104] (Williams, et al 2009, Table 1, Figure 2). This suggests that the study was not very influential at all and likely did not drive the results. Second, the third study was no more negative than the previous two studies. While noting that neurontin performed no better than the placebo, the study noted that it was most effective in young people and people with lower baseline weight. Second, nothing about its authorship or novelty should have struck physicians as more probative than the Pfizer study. Third, this was not the only other study published confirming Pfizer's study (a second RCT was published two months after Pfizer's study confirming the results (Frye et al 2000)).

The second concurrent event is that the American Psychiatric Association issued a new practice guideline for patients with bipolar disorder in 2002 (Hirschfeld et al. 2002), which incorporated the 2000 negative study by Warner-Lambert (Fullerton 2010) as well as the Frye study. Thus, it is unclear whether a drop in 2002 for psychiatrists would be due to the litigation or the revised bulletin. To account for this, only patients seeing psychiatrists (around 80% of the sample) are considered, since they are equally likely to have seen the guidelines in 2002.

[104] One study lists the second RCT in its citations but only discusses the Pande and Frye studies in the paper.

Third-Party Payer Behavior. Indicators for payment type included Medicare, Medicaid,[105] private insurance, "other" or no insurance,[106] and Medicaid.[107] The residual category is self-payment. The bottom panel of Table 2 lists the summary statistics for payer type. The first issue is that the expected payer for the visit may be different from the person who pays for the prescription, which I do not observe. For a couple of these categories, there is a high likelihood that these two payers will be the same, particularly private insurance, Medicaid, and self-payment.[108] For Medicare, however, this is different: there may be a difference in the expected payment for prescription drugs and the expected payment for the visit. Medicare patients did not have Medicare coverage for outpatient drugs until Medicare Part D[109] was enacted in 2006. Thus,

[105] NAMCS documentation lists this as Medicaid for 1998–2000 and Medicaid/SCHIP from 2001–2008. I do not think this actually indicates an expanded coverage group, as there is no mention of changing the coding in the 2001 documentation and the description of the "other" category remains the same between 2000–2001.

[106] This category is included as a control but not discussed.

[107] The process by which expected payer is coded in NAMCS data changes in 2005. Previously, they collected "primary" expected source of payment; in 2005 they collected multiple sources and imposed the following hierarchy: Medicaid, Medicare, private, worker's compensation, self pay, no charge, other, and unknown. In 2007 they reversed this hierarchy, making Medicare dominant over Medicaid. A series of robustness checks group Medicare and Medicaid together as one category (the omitted one) in order to account for this change across years. The results are qualitatively similar. If there is a concern that the relinquishment in 2002 for private payers was driven by the imposition of the hierarchy in general, running the model on 1998-2004 produces the same *Post2002* results. Additionally, insofar as the change in 2004 categorized Medicare above private in contrast to previous coding procedure, an additional robustness check recoded the hierarchy with private first. Whenever "paypriv", a dummy indicating that private insurance was expected to pay, the recode indicates that payment was from private insurance, even if the hierarchy would have listed Medicare/Medicaid. Additionally, Medicare was recoded as Medicaid to preserve the hierarchy established in 2005. The results using this recoding seem qualitatively unchanged.

[108] Of course, this will not always be true but I do not have any reason to believe the error will vary by payer.

[109] Medicare Part D covers prescription drugs that are approved by one of three compendia in order to be reimbursable. The three compendia include American Hospital Formulary Service – Drug Information (AHFS-DI), United States Pharmacopeia – National Formulary (USP-NF), and DRUGDEX (CMA 2010). DRUGDEX, the most inclusive of the compendia, was approved as an official compendium in 1997 (Armstrong 2003).

these patients may have paid for their drugs in a number of ways: self-payment, supplemental private coverage, or dual coverage under Medicaid. If the drug was administered inpatient, it may have been covered by another Medicare Part, though this is less likely. The Medicare category is retained in the analysis, but caution should be taken in interpreting these coefficients, and the Chapter does not focus on these results. However, this takes into account possible treatment differences when the patient's total visit is characterized predominantly as being covered by Medicare.

The variables of interest are the sensitivities of each payer to these information shocks. As discussed in Section III, expected payer can influence treatment in two possible ways: a payer can refuse to cover a particular treatment or can use persuasive measures to spread information about a use's appropriateness. Each of these measures are discussed below.

First, a payer can refuse to cover a particular treatment. . Refusal to reimburse results in two effects: First, if a physician maximizes expected patient benefit, changes in reimbursement make a treatment relatively more expensive for a patient. This should make physicians marginally less likely to prescribe the drug. Private payers seem to cover neurontin's off-label uses to varying degrees. Public payers are more complicated: *Franklin v. Parke Davis* court struggled with whether government programs actually allowed reimbursement of neurontin's off-label uses: while neurontin was not supported by a medical compendium, Parke Davis argued that a majority of state Medicaid programs allowed coverage of non-compendium off-label uses. The government, in turn, argued that Medicaid was confined to uses listed in the designated compendia.[110] The court did not resolve this issue, but noted that if a state Medicaid program did

[110] This arises from an interpretation of the Medicaid statute (Greene 2005), which stated that "[a] State may exclude or otherwise restrict coverage of a covered outpatient drug if the prescribed use is not for a medically accepted indication (as defined in subsection (k)(6) of this

cover neurontin, the state may not have been able to recover under the FCA.[111] The second effect refusal to reimburse has is that a physician may see the change in reimbursement as a signal of whether the use is appropriate.

The second way a payer can influence treatment decisions is by implementing drug utilization reviews to examine how a drug is prescribed and to make suggestions to its physicians. Private payers often do this through their Pharmacy Benefit Managers (Fox 2003). The Medicaid statute also provides for a drug use review program in order to "educate physicians and pharmacists to identify and reduce the frequency of patterns of fraud, abuse, gross overuse, or inappropriate or medically unnecessary care . . . " (42 U.S.C. § 1396r-8(g)(1)(A)).

At least one private insurance company seemed to influence treatment not through refusing reimbursement but through the persuasive means. Kaiser campaigned to reduce its neurontin prescriptions after it was alerted to Pfizer's fraudulent conduct; however, Kaiser did not reduce its prescriptions by refusing to reimburse neurontin. [112] Instead it retained its open formulary, in which it would even reimburse prescriptions not on the formulary. Kaiser issued reports about preferred effective drugs through its Drug Information Service ("DIS"). Its physicians relied on these reports to such an extent that, although they were permitted to prescribe off-formulary, 95% of Kaiser physicians' prescriptions were on-formulary. Upon receiving news of neurontin's ineffectiveness, Kaiser's campaign to against neurontin

---

section." 42 U.S.C. § 1396r-8. Some suggest that this negative framing means that Medicaid can reimburse uses not in the designated compendia: American Hospital Formulary Service – Drug Information (AHFS-DI), United States Pharmacopeia – National Formulary (USP-NF), and DRUGDEX (e.g., CMA 2010; Martin 2004).

[111] U.S. ex rel. Franklin v. Parke-Davis, Div. of Warner-Lambert Co., No. CIV.A. 96-11651PBS, 2003 WL 22048255, at *3-4 (D. Mass. Aug. 22, 2003).

[112] In re Neurontin Mktg. & Sales Practices Litig., 712 F.3d 21, 29 (1st Cir. 2013).

prescriptions reduced new prescriptions by thirty-three percent.[113] This demonstrated screening process helped Kaiser win its Racketeer Influenced and Corrupt Organizations Act (RICO) claim against Pfizer, proving that it would not have reimbursed neurontin prescriptions but for the fraudulent information.

Given that payers can theoretically influence treatment in these two ways, an interesting comparison would be between third-party payers and self-payers. For this reason, the omitted payment category is "self-pay." Self-pay patients arguably are not influenced by payers through either mechanism: they are not subject to reimbursement changes and they presumably do not receive any persuasive literature. Treatment for self-pay patients may be influenced by persuasive techniques used by payers of patients with the same doctor. The study attempts to account for this by clustering errors by physician code and year.[114] Examining the other payers relative to self-pay patients provides an interesting comparison.

[113] The RICO court opinion states the following:
Neurontin prescriptions written by PMG physicians increased dramatically after September 1999 (the fraudulent marketing campaign began in 1997). This notable increase led some Kaiser regions to "examine their members' use of Neurontin" and make efforts to limit it. By the spring of 2002, the Northern California PMG had barred Pfizer drug representatives from detailing its physicians regarding Neurontin, and the same PMG's Drug Utilization Group ("DRUG") began a campaign to promote only the appropriate use of Neurontin, which other regional PMGs joined. In late 2002, Kaiser learned about Franklin's qui tam action and escalated its efforts to limit prescribing of Neurontin for neuropathic pain, bipolar disorder, migraine, and nociceptive pain. Kaiser shared materials about Neurontin produced by DRUG and the Southern California PMG's Drug Utilization Action Team ("DUAT") with all regional PMGs. The district court found that though Neurontin use continued to increase nationally, Kaiser's efforts to limit its use "result[ed] in a 33–34% decrease in new starts of Neurontin." In re Neurontin Mktg. & Sales Practices Litig., 712 F.3d 21, 31–32 (1st Cir.) (internal citations omitted).

[114] Physician codes are only available yearly (there is no way to match across years).

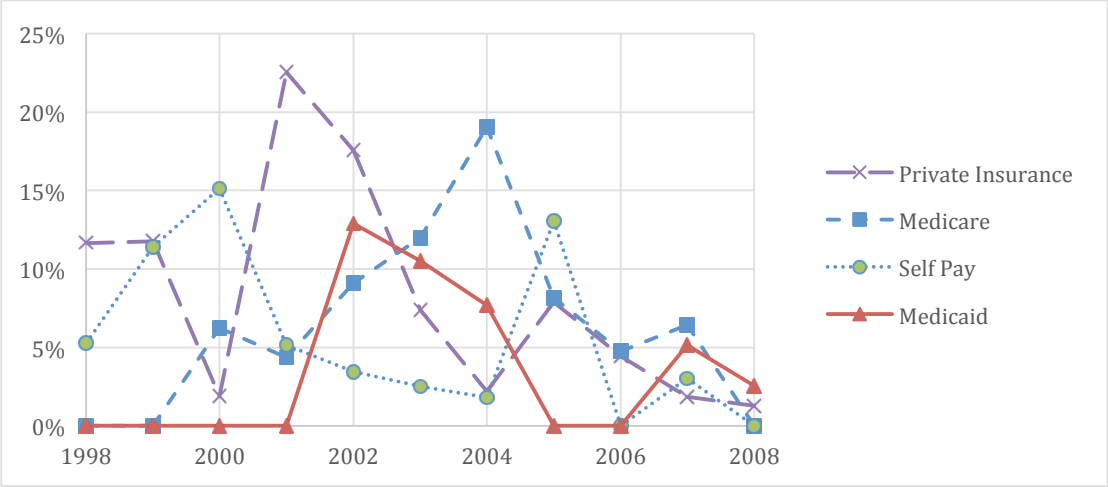

Figure 2 Percent of Neurontin Prescriptions by Payer and Year

Figure 2 plots the percent of neurontin prescriptions in the sample by payer and year. Figure 2 shows that there is considerable heterogeneity in adoption/relinquishment by payer. Self-pay patients seem to increase their prescriptions until 2000, after which they decline until 2005. Private insurance patients experience a big increase in prescriptions after 2000 and then a sharp decline after 2002. Medicaid prescriptions increase after 2001 and gradually decline after 2002. Medicare patient prescriptions continue increasing until 2004, after which they decline.

Since state Medicaid claims were the basis for the FCA suit, another important comparison is how patients with Medicaid were treated relative to patients with private insurance. Medicaid was seeking to be reimbursed using the FCA, while private insurance companies had to seek compensation elsewhere. Private insurance companies have had a harder time recovering their reimbursements and might apply more pressure on physicians to police prescriptions more carefully. Private third-party payers bring suit under the Racketeer Influenced and Corrupt Organizations Act ("RICO") to recover for fraudulent off-label promotions.[115] RICO prohibits anyone associated with an "enterprise" "to conduct or participate, directly or indirectly,

[115] 18 U.S.C. §§ 1961–1968 (2012).

in the conduct of such enterprise's affairs through a pattern of racketeering activity or collection of unlawful debt."[116] Third-party payers have used this statute to allege that pharmaceutical companies engaged in an enterprise through a pattern of racketeering (usually mail fraud) that caused injury to the third-party payer (Cooney et al. 2010). These attempts are rarely met with success. Kaiser's success can be attributed, at least in part, to its proactive relinquishment of neurontin, such as barring pharmaceutical representatives from detailing their physicians, and launching their own campaigns for appropriate drug promotion.[117] For FCA claims, no such proof is necessary as the burden is on pharmaceutical companies to know what uses are nonreimbursable.

This section tests whether there is heterogeneity in relinquishment by payer. Importantly, this Chapter looks at relative relinquishment, the comparative change in a payers' prescriptions in response to an information shock relative to the corresponding change for self-pay payers. Relative relinquishment does not necessarily mean that nominal prescriptions declined or that the total nominal decline is significant. However, I believe this is an informative measure, as comparing behavior relative to self-pay patients provides a baseline of consumer behavior in the absence of third-party intervention. This section tests for heterogeneity in relinquishment through two analyses: a linear probability model for the likelihood of prescribing neurontin and a triple-difference analysis comparing neurontin prescriptions to prescriptions of lithium and approved anticonvulsants.

[116] 18 U.S.C. § 1962 (c).
[117] In re Neurontin Mktg. & Sales Practices Litig., 712 F.3d 21, 31 (1st Cir. 2013).

B. Linear Probability Model

The likelihood of being prescribed a particular drug is estimated using linear probability models.[118] Specifically, these linear probability models measure the propensity of physicians to prescribe approved anticonvulsants, lithium, or neurontin. The data are not longitudinal, and the model treats each observation as a separate draw. All patient records that list bipolar disorder as one of the three possible diagnoses for the visit are included.

The decision to prescribe each drug might not be independent from one another; drugs might be substitutes for one another and the factors associated with choosing one might be correlated with another drug choice. To account for this dependence, the linear probability models are run together in a seemingly unrelated regression. The seemingly unrelated regression calculates the probability of several discrete choices and allows for the errors for all choices to be correlated (ρ_{ki}). Since this Chapter focuses on the likelihood of prescribing neurontin, however, only that linear probability model is reported. The basic model is in equation (4).

(4) $$Treatment_k = X'\beta_1 + J'\beta_2 + I'\beta_3 + Z'\beta_4 + Z*I'\beta_5 + \varepsilon_i ,$$

where *X* is a vector of patient characteristics that measure differences in medical benefits based on physical differences, *J* contains information like journal articles on neurontin, *I* is a vector of the aforementioned information-shock time periods, *Z* is a vector of payment characteristics, and *Z*I* is a series of interaction terms between time shocks and payment characteristics.

A number of patient level controls are used, including sex and age. Since the data do not have detailed information about patient health, smoking status is used as a proxy for patient

[118] To confirm the appropriateness of the linear probability model, I run a probit model and use inteff to estimate the marginal effects of *Post2002*Private*. The effect is similar in magnitude, providing support for the linear probability model.

health.[119] Patient-level controls also include measures of bipolar severity. Diagnosis codes found in NAMCS are used as a measure for this.[120] Bipolar severity should be positively correlated with the prescription of a nontraditional treatment, such as neurontin or other anticonvulsants, because people with severe bipolar might not respond to traditional treatment. Similarly, the presence of comorbidities such as psychotic behavior presents complexity for which physicians might seek innovative treatments. To measure comorbidities, an indicator variable measures a patient displays psychotic symptoms. A patient is categorized as having psychotic symptoms if the diagnosis code indicates psychotic behavior.

To capture the information spreading outside of the information shocks, I include a cumulative measure of the number of review studies that either made positive or negative conclusions about neurontin's effectiveness for bipolar disorder. This measure is from Williams et al. (2009), which documents review studies that evaluated the use of neurontin for the treatment of bipolar disorder.

Finally**,** previous research emphasizes the role of pharmaceutical detailing on physician learning. Unfortunately, NAMCS does not include a measure for detailing patterns. To crudely account for pharmaceutical detailing would result in omitted variable bias, so as a crude indicator, the four region controls provided by NAMCS are included. These region controls attempt to account for different pharmaceutical representative territories, different CME programs, and region-specific prescription idiosyncrasies. Table 4 only includes the results for

[119] Since smoking is correlated with heart disease, stroke, and various cancers (CDC 2012), it can be used as a rough measure of patient general health. Current smoking status is indicated for years 1994–1996 and 2001–2010 for NAMCS. For the missing years, smoking status is assigned if the record indicated that the patient received counseling for smoking cessation. The results are robust to the exclusion of this measure.

[120] There are two diagnosis codes designated as severe: one indicating that the diagnosis is severe with psychotic behavior and the other indicating that the diagnosis is severe without psychotic behavior. I use both diagnosis codes for my severity measure.

the likelihood of prescribing neurontin, as this section will focus on these results. Columns (1) and (2) display the results for the seemingly unrelated regressions. The correlation between each pair of errors is ρ_{hj}. The null hypothesis is ρ_{hj}=0 and that the errors are independent of one another. A chi-square test tests whether ρ_{hj} is statistically different than 0. If the null hypothesis is rejected for at least one of the error pairs, the standard errors under the seemingly unrelated regression model are more appropriate. The seemingly unrelated regression rejects the null hypothesis that ρ=0, and the standard errors under this model are more appropriate than in the separate models. In the total prescription model, the error correlations are negative, suggesting that these drugs substitute for each other.[121] Columns (3) and (4) run weighted linear probability models for comparison. These do not account for dependency in the errors of the system of equations. However, separate linear probability models allow for weighted regressions and robust standard errors. Comparing Columns (1)–(2) to (3)–(4) shows that the results are similar and generally robust to either assumption.

[121] However, neurontin and carbamezapine's errors are positively correlated.

Table 4. Likelihood of Prescribing Neurontin, Seemingly Unrelated Regressions and Linear Probability Models, 1998-2008.

	Seemingly Unrelated Regression		Separate Linear Probability Models	
Patient age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Male patient	-0.032***	-0.033***	-0.027**	-0.027**
	(0.011)	(0.011)	(0.012)	(0.012)
Severe bipolar diagnosis	0.073**	0.071**	0.108*	0.105
	(0.031)	(0.031)	(0.065)	(0.065)
Use tobacco	-0.009	-0.009	0.000	-0.000
	(0.015)	(0.015)	(0.019)	(0.019)
Psychotic behavior	0.005	0.007	-0.009	-0.006
	(0.044)	(0.044)	(0.096)	(0.095)
Medicare	-0.095**	-0.094**	-0.116***	-0.113***
	(0.044)	(0.044)	(0.038)	(0.038)
Medicaid	-0.106**	-0.099*	-0.111***	-0.103**
	(0.053)	(0.053)	(0.041)	(0.043)
Private Insurance	-0.024	-0.023	-0.032	-0.035
	(0.032)	(0.032)	(0.047)	(0.046)
Post2000	-0.068*	-0.016	-0.066	0.015
	(0.039)	(0.047)	(0.040)	(0.058)
Post2000*Payment				
Medicare	0.129**	0.130**	0.135**	0.135**
	(0.062)	(0.062)	(0.056)	(0.057)
Medicaid	0.187***	0.181***	0.166***	0.158***
	(0.067)	(0.067)	(0.055)	(0.056)
Private Insurance	0.192***	0.191***	0.182***	0.183***
	(0.047)	(0.047)	(0.060)	(0.058)
Post2002	-0.017	0.078	-0.005	0.108*
	(0.038)	(0.051)	(0.038)	(0.057)
Post2002*Payment				
Medicare	0.092	0.090	0.085	0.083
	(0.061)	(0.061)	(0.077)	(0.078)
Medicaid	-0.053	-0.054	-0.018	-0.019
	(0.055)	(0.055)	(0.070)	(0.069)
Private Insurance	-0.146***	-0.145***	-0.152***	-0.151***
	(0.045)	(0.045)	(0.051)	(0.051)

Post2004	0.005	0.078*	-0.030	0.036
	(0.034)	(0.040)	(0.036)	(0.044)
Post2004*Payment				
Medicare	-0.105**	-0.114**	-0.089	-0.099
	(0.053)	(0.053)	(0.068)	(0.068)
Medicaid	-0.047	-0.050	-0.037	-0.039
	(0.047)	(0.047)	(0.062)	(0.062)
Private Insurance	-0.009	-0.015	0.032	0.026
	(0.039)	(0.039)	(0.041)	(0.041)
Positive Studies		0.022		0.001
		(0.014)		(0.020)
Negative Studies		-0.028***		-0.025***
		(0.009)		(0.009)
Constant	0.139***	0.052	0.125***	0.149
	(0.031)	(0.079)	(0.041)	(0.116)
Observations	2,142	2,142	2,142	2,142
R-squared	0.065	0.070	0.061	0.066

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Variables included but not shown include indicators for region and the main effects/interactions for "other payment". Columns (1) and (2) allow errors across multiple linear probability models (not shown here) to be dependent. Columns (3) and (4) involve separate linear probability models that are weighted by patient weight and include standard errors clustered by physician code and year.

Columns (1) and (3) only include information shocks, while Columns (2) and (4) include a cumulative measure of positive and negative review articles on neurontin for the treatment of bipolar disorder. The patient demographics are significant. Males are less likely to receive neurontin. Patients with a "severe" bipolar diagnosis are more likely to receive neurontin than those without such designation, suggesting that neurontin might have been more of a "last-resort" treatment. The payment indicator variables are also interesting: relative to self-pay patients, patients with Medicare and Medicaid were less likely to receive neurontin. As predicted, "negative" studies, studies implying that neurontin would be inappropriate for the treatment of bipolar disorder, reduce the likelihood that neurontin is prescribed.

The information shock interactions are the main variables of interest. The main effect of the information shock can be interpreted as the incremental change from the previous information shock for self-pay patients. *Post2002* should be interpreted as the change in likelihood for self-pay patients after 2002, relative to after 2000. The information shock–payment interactions should be interpreted as the incremental change from the information shock for a particular payer, relative to the corresponding change for self-pay patients. For example, *Post2002*Private Insurance* should be interpreted as the relative change for private insurance patients after 2002 (relative to after 2000) compared to the corresponding change for self-pay patients.

Given that the negative scientific study was published in 2000, *Post2000* and its subsequent interactions were hypothesized to be negative, as the negative scientific study on neurontin should have reduced the number of neurontin prescriptions for bipolar disorder. This variable is only negative for self-pay patients (*Post2000*<0), and even then only significantly in Column (1). *Post2000*Private Insurance*, *Post2000*Medicaid, and Post2000*Medicare* are positive and significant, suggesting that the publication of the article did not correspond to decreased likelihood of being prescribed neurontin for private insurance, Medicaid, or Medicare patients, relative to the changes in likelihood for self-pay patients. Instead, relative to self-pay patients, the above patients were more likely to receive neurontin after this time. A possible reason for this relative increase is found in Kaiser's RICO lawsuit against Pfizer. Kaiser claims that Pfizer's "misrepresentations and omissions during the development of drug monographs" led them to remove any restrictions on the prescription of neurontin in late 1999.[122]

[122] In re Neurontin Mktg. & Sales Practices Litig., 712 F.3d 21, 29 (1st Cir. 2013).

Post2002 is largely insignificant. Relative to this, the likelihood that private insurance patients are prescribed neurontin drops around 14–15 percentage points. This is consistent with private insurance company Kaiser's claims that they intensified their campaign to decrease the prescriptions of neurontin after hearing about the FCA litigation.[123] The likelihood of prescription for Medicaid patients also declines, but not significantly. The difference between *Post2002*Medicaid* and *Post2002*Private* is significant.[124] This differential response is intriguing and suggests that anticipated payment is a strong influence on a physician's prescription patterns.

Post2004 main terms and interactions for Medicaid and private insurance are largely insignificant. Patients who are classified as Medicare patients tend to have a reduced likelihood of receiving neurontin after 2004 (*Post2004*Medicare*<0). The reason for this finding is unclear: Medicare was not paying for prescription drugs during this time. It is possible that this was driven by a Medicare-specific treatment plan, if physicians are just prone to treat Medicare physicians differently.

An alternate way of analyzing relinquishment is not to look at incremental changes but to examine mutually exclusive period intervals. As a robustness check, I run a specification in which the information shocks are considered as intervals in time: before the Warner Lambert study (*1998-2000*), after the Warner Lambert study and before news of the FCA (*2001–2002*), after the news of the FCA and before settlement (*2003–2004*), and after settlement (*2005–2008*).

---

[123] *See supra* note 113.

[124] For Table 4, Column (4), the difference between *Post2002*Medicaid* and *Post2002*Private* has an F-statistic of 3.10 (prob >F=0.0786). Similarly, in column (2), the F statistics is 3.96 (prob > F = 0.0466). Unweighted separate, regressions with robust standard errors produce an F of 2.62 (prob > F =0 .1059) and an F statistic of 2.05 (prob > F =0.1528) with clustered errors. Unweighted separate regressions, however, are not my preferred specification.

Table 5. The Likelihood of Prescribing Neurontin, Separate OLS regression, Intervals.

	(1)	(2)
Patient age	0.000	0.000
	(0.000)	(0.000)
Male patient	-0.027**	-0.027**
	(0.012)	(0.012)
Severe bipolar diagnosis	0.108*	0.105
	(0.065)	(0.065)
Use tobacco	0.000	-0.000
	(0.019)	(0.019)
Psychotic behavior	-0.009	-0.006
	(0.096)	(0.095)
Medicare	-0.116***	-0.113***
	(0.038)	(0.038)
Medicaid	-0.111***	-0.103**
	(0.041)	(0.043)
Private Insurance	-0.032	-0.035
	(0.047)	(0.046)
2001-2002	-0.066	0.015
	(0.040)	(0.058)
2001-2002 *Payment		
Medicare	0.135**	0.135**
	(0.056)	(0.057)
Medicaid	0.166***	0.158***
	(0.055)	(0.056)
Private Insurance	0.182***	0.183***
	(0.060)	(0.058)
2003-2004	-0.071	0.122
	(0.048)	(0.096)
2003-2004*Payment		
Medicare	0.220***	0.218***
	(0.074)	(0.074)
Medicaid	0.148**	0.140*
	(0.072)	(0.073)
Private Insurance	0.030	0.032
	(0.060)	(0.059)

2005-2008	-0.100***	0.158
	(0.038)	(0.105)
2005-2008*Payment		
Medicare	0.131***	0.119***
	(0.042)	(0.043)
Medicaid	0.111**	0.100**
	(0.047)	(0.050)
Private Insurance	0.062	0.059
	(0.051)	(0.050)
Positive Studies		0.001
		(0.020)
Negative Studies		-0.025***
		(0.009)
Constant	0.125***	0.149
	(0.041)	(0.116)
Observations	2,142	2,142
R-squared	0.061	0.066

Standard errors are clustered by physician code and year. *** p<0.01, ** p<0.05, * p<0.1. Variables included but not shown include indicators for region and the main effects/interactions for "other payment." In this table, periods of time are separated into intervals. The omitted time interval is *1998-2000*.

In the following analysis, *1998-2000* is omitted, along with its interactions with payers. Thus, all the coefficients reported in Table 5 are interpreted with respect to *1998-2000* (before the Warner Lambert study), such that the effect of the following periods is with respect to the omitted category. *2001–2002* should be interpreted as the change in likelihood for self-pay patients between 2001 to 2002, relative to the self-pay patients in *1998-2000*, and *2003–2004* should be interpreted as the change in likelihood for self-pay patients between 2003 to 2004, also relative to the self-pay patients in *1998-2000*. The payment interactions should be interpreted as the difference between patients with the given payment method and self-pay patients in the given interval, relative to the corresponding difference in *1998-2000*.

The results in Table 5 reflect similar findings as those in Table 4. From Column (2), relative to *1998-2000*, the difference in likelihoods of receiving neurontin after 2000 between private insurance and self-pay patients increases 18.3 percentage points (*2001-2002*Private*).

For Medicaid, this difference increases by 15.8 percentage points and by 13.5 percentage points for Medicare patients. The payment interactions for *2003-2004* and *2005-2008* are positive, however, indicating that relative to the difference between self-pay patients and patients with private insurance, Medicaid, or Medicare in *1998-2000*, the difference in future periods are larger. However, comparing the coefficients on the interaction terms between *2001-2002* and *2003-2004* shows that this relative difference decreases after news of the suit hit. Thus, regardless of whether information shocks are studies as incremental changes or mutually exclusive intervals, heterogeneity in payer relinquishment is apparent.

C. Triple-Differences

The second way to analyze the relinquishment of neurontin is to compare neurontin prescriptions to prescriptions of approved anticonvulsants or lithium. A triple difference analysis helps to do this by comparing the response of a treated group (neurontin) to that of an untreated group (approved anticonvulsants or lithium) before and after treatment (information shocks). The triple difference allows the response of the treated group to vary by payer. To construct the triple differences analysis, I only keep observations where one of the diagnoses relates to bipolar disorder and where one of the prescriptions was lithium, an eventually approved anticonvulsant, or neurontin. For the reasons stated earlier, only patients who visited a psychiatrist are retained. Each prescription is weighted by the patient weight, to reflect a national estimate of the prescription, and the weighted sum of prescriptions is collapsed by month, year, region, and payer type.[125] The weighted purchases of each drug are described in Table 6.

[125] As in the analysis in Section I.B, I merge the month-year–payer–region categories with a full-factorial matrix of month-year–payer–region categories for each substance and impute categories with zero prescriptions. I drop observations that have no prescriptions in any substance–payer–region category over any time period. This leaves 7,788 substance-month-year–payer–region observations.

Table 6. Number of Purchases[a], 1998-2008

	Mean	Std. Dev.
Drug		
Neurontin	1004.15	6729.39
Lithium	3053.04	10626.54
Approved Anticonvulsants	5564.75	17307.17

[a] These values are determined by multiplying each purchase by the corresponding patient weight.

The following equation is estimated:

$$(5) Prescription_{myp} = Drug'\beta + Payer'\gamma + Information\ Shock'\delta + Drug * Payer'\vartheta + Drug * Information\ Shock'\theta + Payer * Information\ Shock'\alpha + Drug * Payer * Information\ Shock'\tau,$$

where *Drug* is a vector of drug choices: neurontin, lithium, or approved anticonvulsants. The omitted drug category is lithium in Column (1)–(2) of Table 7 and approved anticonvulsants in Column (3)–(4). The omitted category provides a baseline, such that the number of neurontin prescriptions is measured relative to the omitted category. *Payer* is a vector including Medicare, Medicaid, private insurance, "other" or no insurance, and self-payment. The residual category is self-payment. *Information Shock* is a vector of the aforementioned events in 2000, 2002, and 2004.

Table 7 lists the results. These results tell a similar story as the linear probability model results. Column (1)–(2) compares neurontin prescriptions to those of approved anticonvulsants, evaluating trends in neurontin with the trends of lithium as the baseline. Column (3)–(4) compares neurontin prescriptions to lithium prescriptions. Both are provided because it is unclear which provides the better comparison group. Lithium is an established treatment while the approved anticonvulsants are the same drug class as neurontin.

Table 7. Number of Prescriptions, by Month, Year, and Coverage: Difference in differences.

Neurontin	-1,872.35	-62.42	-2,172.25**	-362.31
	(1,164.42)	(2,551.93)	(1,036.47)	(2,583.10)
Medicare	-126.29	-126.29	-1,548.03	-1,548.03
	(1,297.99)	(1,294.11)	(1,123.03)	(1,122.16)
Medicaid	-1,800.81	-1,800.81	-2,973.97***	-2,973.97***
	(1,101.85)	(1,100.70)	(832.26)	(831.91)
Private Insurance	5,427.94***	5,427.94***	2,350.24*	2,350.24*
	(1,775.21)	(1,774.84)	(1,321.54)	(1,321.63)
Neurontin*Payment				
Medicare	-1,088.48	-1,088.48	333.26	333.26
	(1,475.57)	(1,473.05)	(1,317.24)	(1,317.52)
Medicaid	479.53	479.53	1,652.69	1,652.69
	(1,298.71)	(1,298.52)	(1,077.51)	(1,078.22)
Private Insurance	-5,150.93**	-5,150.93**	-2,073.23	-2,073.23
	(2,031.33)	(2,030.57)	(1,646.02)	(1,645.54)
Post2000	1,075.61	835.39	-1,120.50	-1,360.72
	(2,064.38)	(2,180.14)	(1,076.40)	(1,380.90)
Neurontin*Post2000	-2,017.63	-387.77	178.47	1,808.33
	(2,214.34)	(2,437.41)	(1,341.75)	(1,759.21)
Payment*Post2000				
Medicare	-1,343.51	-1,343.51	1,582.23	1,582.23
	(2,391.16)	(2,389.58)	(1,526.67)	(1,526.88)
Medicaid	773.09	773.09	3,217.35*	3,217.35*
	(2,372.81)	(2,372.72)	(1,683.95)	(1,685.23)
Private Insurance	2,846.93	2,846.93	1,099.83	1,099.83
	(3,720.05)	(3,720.77)	(2,133.98)	(2,136.29)
Neurontin*Payment*Post2000				
Medicare	2,860.35	2,860.35	-65.39	-65.39
	(2,557.30)	(2,556.28)	(1,775.62)	(1,776.48)
Medicaid	1,220.85	1,220.85	-1,223.41	-1,223.41
	(2,541.19)	(2,541.10)	(1,913.93)	(1,915.04)
Private Insurance	2,507.23	2,507.23	4,254.33	4,254.33
	(4,363.93)	(4,363.57)	(3,123.93)	(3,124.14)
Post2002	-335.15	-11.52	-645.68	-322.06
	(2,191.08)	(2,546.89)	(928.31)	(1,500.56)

Neurontin*Post2002	383.62	2,112.53	694.15	2,423.06
	(2,315.78)	(2,793.01)	(1,193.20)	(1,888.38)
Payment*Post2002				
Medicare	-396.44	-396.44	1,316.39	1,316.39
	(2,445.65)	(2,446.36)	(1,730.86)	(1,731.90)
Medicaid	550.19	550.19	93.16	93.16
	(2,722.01)	(2,722.69)	(1,903.46)	(1,904.94)
Private Insurance	2,766.06	2,766.06	4,771.09*	4,771.09*
	(4,426.26)	(4,427.54)	(2,830.76)	(2,832.94)
Neurontin*Payment*Post2002				
Medicare	1,179.85	1,179.85	-532.97	-532.97
	(2,707.64)	(2,708.33)	(2,084.69)	(2,085.62)
Medicaid	-258.90	-258.90	198.13	198.13
	(3,043.04)	(3,043.41)	(2,339.63)	(2,340.54)
Private Insurance	-6,810.21	-6,810.21	-8,815.24**	-8,815.24**
	(4,983.84)	(4,984.35)	(3,641.43)	(3,642.28)
Post2004	-1,002.35	-412.63	75.91	665.64
	(1,375.59)	(1,600.00)	(726.78)	(1,107.82)
Neurontin*Post 2004	490.65	1,213.53	-587.62	135.27
	(1,514.13)	(1,829.15)	(963.61)	(1,418.82)
Payment*Post2004				
Medicare	5,864.04***	5,864.04***	284.25	284.25
	(1,976.96)	(1,977.94)	(1,663.99)	(1,663.85)
Medicaid	1,941.66	1,941.66	-307.32	-307.32
	(2,079.86)	(2,081.27)	(1,422.31)	(1,423.11)
Private Insurance	4,934.59	4,934.59	-3,447.75	-3,447.75
	(4,001.71)	(4,003.28)	(2,593.36)	(2,593.56)
Neurontin*Payment*Post2004				
Medicare	-6,283.01***	-6,283.01***	-703.21	-703.21
	(2,234.23)	(2,234.79)	(1,962.72)	(1,962.25)
Medicaid	-2,452.08	-2,452.08	-203.10	-203.10
	(2,427.91)	(2,429.68)	(1,895.23)	(1,896.55)
Private Insurance	-4,534.76	-4,534.76	3,847.58	3,847.58
	(4,187.64)	(4,188.81)	(2,871.96)	(2,871.66)
Positive Studies		491.93		491.93
		(393.35)		(393.35)
Neurontin*Positive Studies		-285.15		-285.15
		(508.04)		(508.04)
Negative Studies		-235.89		-235.89
		(343.45)		(343.45)

Neurontin* Negative Studies		-289.15		-289.15
		(420.50)		(420.50)
Omitted category	**Approved Anticonvulsants**	**Approved Anticonvulsants**	**Lithium**	**Lithium**
Observations	7,788	7,788	7,788	7,788
R-squared	0.09	0.09	0.09	0.09
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Variables included but not listed are indicator variables for region, main effects and interactions for "other payment" and all the main effects and interactions of the non-baseline category (lithium for Column (1)-(2) and approved anticonvulsants for Column (3)-(4)).				

Consistent with previous results, no decline is apparent after the original publication of the negative Pfizer study in 2000. Few of the triple interactions with *Post2000* are significant. For 2002 interactions, the private payer triple interactions in Columns (3)–(4) are significant. *Neurontin*Private* Post2002*<0 compares the difference in incremental changes[126] between neurontin prescriptions and lithium prescriptions for private insurance patients, relative to the corresponding difference in self-pay patients. It essentially signifies that the change in neurontin prescriptions relative to the change in lithium prescriptions after 2002 is significantly more negative for private patients than for self-pay patients. This triple interaction is also negative and but not significant in Columns (1)–(2), relative to approved anticonvulsants.

As in the linear probability model, the relinquishment of neurontin for Medicare patients seems to only take place after 2004, as *Neurontin*Medicare Post2004* <0, although this coefficient again should be interpreted with caution. Other triple interactions with *Neurontin*Post2004* are negative and mostly insignificant, relative to self-pay patients.

In sum, this model seems to support the results from the linear probability model: no relative relinquishment for private, Medicare, and Medicaid is observed after the 2000 study. Instead, relative relinquishment happens later. The change in likelihood of receiving neurontin is

[126] The incremental change is the change in prescriptions relative to 2000-2002.

significantly more negative for private insurance patients after 2002, relative to changes in the likelihood for self-pay patients.

Clearly this identification strategy makes causal inference challenging. What can be concluded, however, is that any relative decrease in neurontin prescriptions does not occur directly after publication of the 2000 study for patients with third-party payers. It is possible that relinquishment is caused by the scientific study but takes time to be adopted. However, the trend observed seems more consistent with delayed relinquishment than gradual relinquishment, since prescriptions seem to increase after 2000. I do not know if reimbursement changed for various Medicaid programs during this time, but Kaiser's story of removing restrictions on neurontin suggests that the availability of unrestricted reimbursement might have caused physicians to continue prescribing neurontin. For self-pay payers, in contrast, it is possible that the negative study was sufficient to cast doubt on the appropriateness of the drug, such that physicians did not want to impose an expensive drug on their patients without insurance. This might suggest that physicians focus less on medical opportunity cost of a drug if the monetary cost is low.

For third-party payers, the results suggest that relinquishment is spurred more by litigation than the 2000 study. Attributing the decline after 2002 to litigation is not without challenges: it is possible that physicians catering to patients with private insurance were more likely to pay attention to the APA updated guidelines publishing the previous negative studies about neurontin. However, Kaiser does at least claim that the lawsuit caused them to escalate their efforts to limit prescriptions of neurontin.[127] The relinquishment observed after 2002 occurs for private insurance and Medicaid patients, with a significantly stronger effect for private insurance patients.

[127] *See supra* note 113.

One concern might be that these results are actually driven by the differing costs of neurontin over time. Neurontin was set to go off patent around 2001 (Neafsey 2005), which might make the drug cheaper for patients.[128] The increase in prescriptions after 2000 might be caused by physicians simply finding neurontin sufficiently cheap to prescribe, even if they knew about the study. However, this argument is problematic for two reasons. First, this does not explain the decline in neurontin prescriptions after 2002. If physicians knew about the drug ineffectiveness in 2000 and simply prescribed neurontin because of the lower price, the 2002 litigation would produce no new information and should not affect the prescription decision (since the price would remain at off-patent levels between 2000 and 2002, there should be no decline in 2002). Second, it is unclear whether the price of neurontin actually dropped during this period. Before the patent expired, Pfizer obtained a production patent to extend protection until 2014; though generic companies contested this patent,[129] it was upheld in 2007 (Neafsey 2005; Cox 2007).

Another proposed explanation is that the changes in prescriptions were due to underlying changes in pharmaceutical detailing effort. Although the data only include a crude measure of pharmaceutical detailing, anecdotal evidence suggests that this was not the case. In its announcement of settlement, the DOJ notes that the Pfizer settlement included agreeing to a corporate compliance program "which will ensure that the changes Pfizer Inc made after acquiring Warner-Lambert in June 2000, are effective in training and supervising its marketing and sales staff..." (DOJ 2004). This suggests that Pfizer might have tried to curb some detailing

[128] It is not clear that prices would have necessarily dropped post-patent. Post-patent price changes are dependent on many different variables not discussed in this Chapter.
[129] Some sources suggest that a generic launch was attempted in 2004 (Cox 2007), which would have resulted in lower-priced neurontin; however, it is unclear how large or successful this launch was. Additionally, this launch should have increased the number of neurontin prescriptions after 2004 if cost is relevant, but there is little evidence of this.

after buying the company. Additionally, the DOJ asserts that the "charged conduct" occurred before Pfizer bought Warner-Lambert in 2000. Finally, the original FCA claim was unsealed in 1999, suggesting that Pfizer was under greater federal scrutiny during this period. Arguably then, pharmaceutical detailing should have declined around at least 1999 and 2000, which would not explain the continued increase through 2000 and the decline after 2002.

As noted in Section 0, there are two reasons that litigation could have spurred this relinquishment. Litigation may have emphasized and publicized already existing information on neurontin. Insurance companies have accused pharmaceutical companies of a particular form of publication bias in which negative studies are published in lower-circulating journals. Thus, even if physicians differentiate between strong and weak studies, they just might not be aware of the negative studies because they are in lower-circulating journals. If this is the case, however, litigation is an expensive and inefficient flag for negative studies.

The second potential reason is that litigation brought to light new internal information regarding fraud. If physicians were swayed by the new *internal* documents regarding neurontin's effectiveness, which came to light during the trial, this brings into question whether physicians can rely on the public pharmaceutical literature to adequately inform themselves of an off-label use's effectiveness. If internal documents are necessary to ascertain a drug's effectiveness, the assumption underlying off-label uses—that physicians will make appropriate decisions based on current public scientific literature—might be flawed.

The results also suggest that third-party payers exert considerable influence over whether physicians prescribe a particular drug. The results indicate that patients with private insurance became less likely to be prescribed neurontin after 2002, relative to the change in likelihood for self-pay patients. Patients with Medicaid similarly experienced relinquishment after 2002,

although this effect is significantly different than the effect for private insurance patients. In its suit against Pfizer, Kaiser claimed that after learning about the FCA lawsuit in 2002, it started an aggressive campaign to get its physicians to stop prescribing neurontin to treat bipolar disorder.

This analysis adds to the analysis of the Section IV by taking a more nuanced look at one focal FCA case. In accordance with Mapes's (1977) findings, this Section finds evidence of heterogeneity in relinquishment by payer. Moreover, the results imply that some types of information are more influential in spurring relinquishment.

VI. Conclusion: Consequences of the False Claims Act

This Chapter sought to examine the effect of the FCA on relinquishment with two analyses. Part IV provides a measure of average relinquishment over multiple FCA settlements. It cannot, however, measure the effect of alternate information sources or distinguish between inappropriate and appropriate off-label uses. If all off-label uses are relinquished, the FCA settlement might instead function more like a stigma than an information source. To fill this gap, Part V studies one inappropriate off-label use of one drug, and examines important deadlines in the drug's timeline besides FCA settlement; however, attributing the relinquishment in Part V to FCA suit is more difficult, given the many different events in one drug's timeline. Finally, it is unclear how generalizable the results of Part V are. Neurontin was the first big FCA case based on off-label promotion, after which many such suits were settled.

Despite these caveats, there are a couple of clear takeaways. First, there seems to be a significant effect of FCA settlements on relinquishment of off-label uses. Relinquishment seems to only occur for off-label uses of these drugs, which is understandable since these are the only uses for which the suit should be at all informative.

Second, Part V demonstrates that, at least in the case of neurontin, there are differences in relinquishment by expected payer. The likelihood that patients received neurontin decreased more after news of the FCA suit in 2002 for patients with private insurance, and to a lesser extent Medicaid, relative to the change for self-pay patients. These two conclusions will discussed separately below.

A. FCA as a Informational Learning Mechanism

Part IV finds an average relinquishment of off-label prescriptions after settlement of an FCA claim. Part V finds that relative to traditional sources of relinquishment, such as the publication of influential articles, the FCA seems to be a visible event that spurs relinquishment. Despite the strong evidence of neurontin ineffectiveness published in 2000, relinquishment—at least for third-party payers—only occurred afterward seemingly in response to the FCA suit and discovery. As previously discussed, there are at least two reasons for this phenomenon: First, the FCA can act as an industry-wide signal, bringing to light obscure scientific publications. Second, the FCA can bring new knowledge to light by exposing internal documents or suppressed studies.

In the first case, it is important to note that the FCA is not the ideal information source for off-label appropriateness. A lawsuit is an expensive signal for what should be purely scientific information. Ideally, some government entity would publish a drug digest synthesizing all studies done on each use and making a suggestion as to the appropriateness of the use. [130] However, given the high incidence of publication bias, such a government entity might not be sufficient to correct the problem. In the second case, new information is generated from the FCA trial through whistleblowers publicizing internal information. If studies are actually falsified or if

[130] There are private drug digests, but some wonder about whether the publications are truly unbiased (Armstrong 2003).

publication bias is too difficult to overcome, "smoking gun" evidence revealed through the FCA suit becomes particularly important.

B. Implications of Heterogeneity in Relinquishment by Payment

Part V finds evidence of different relative relinquishment patterns among patients with different payers, in the case of neurontin. This section discusses both the hypothesized reasons behind the current findings and possible future effects.

The presence of a third-party payer seems to make a difference in adoption and relinquishment patterns. From the raw data alone, it looks as though self-pay patients were less likely to receive neurontin after 2000, while Medicaid patients adopted neurontin and then relinquished it after 2002. Private patients were more liberal in their adoption; however, their relinquishment effect was also strong. Through the regression analysis, patients with private insurance were less likely to receive neurontin after 2002, relative to the change in likelihood experienced by self-pay patients. Medicaid patients also experienced a decline after 2002 relative to self-pay patients, although significantly different than the decline for private insurance patients.

Although both private and Medicaid patients face a reduced relative likelihood of being prescribed neurontin, the relinquishment for Medicaid is significantly different from private relinquishment. These results provide some evidence of heterogeneity within third-party payers and can have implications for the broader question of heterogeneity between public and private third-party payers. There may be two reasons for such heterogeneity: varying ability and different incentives[131] of each payer to relinquish early.

[131] Notably, the neurontin case was among the first off-label promotion cases, in which a lot of the FCA jurisprudence was developed; thus, it is unclear whether the differences in legal

The first concern may be that there is a significant difference in the ability of Medicaid/Medicare to refuse to reimburse a particular use and subsequently to prevent it from being prescribed, relative to private payers. The Medicaid statute does seem to allow for requiring prior authorization or even exclusion of "a covered outpatient drug if…the prescribed use is not for a medically accepted indication." [132] As soon as a compendium updates its recommendation based on the new negative rules (or immediately, if there was no preexisting compendia evidence and Medicaid was exercising its discretion in covering the use), Medicaid can refuse to reimburse. Moreover, through its drug utilization review,[133] it may be able to issue recommendations against such uses, as Kaiser did.

Medicare Part D might pose a different set of challenges. While Medicare Part D was not a big player in the neurontin case, going forward, Medicare Part D will play a larger role in FCA suits. Medicare Part D likely sets a more uniform reimbursement standard compared to Medicaid, relying strongly on uses listed in three compendia (CMA 2010). There are also additional requirements to cover all drugs falling within "protected classes," (Barker and Margulies 2014). Finally, depending on the level of communication between private insurance companies managing the Part D program and the Parts of Medicare outlining medical care, Medicare may not have sufficient tools to monitor appropriate usage.

The second possibility reason for heterogeneity between public and private payers is that the broad standards of the FCA reduce incentives for Medicare/Medicaid to exercise persuasive power to relinquish inappropriate treatments. The FCA standards are rather lenient and have

---

obligations were fully exploited in this case. Future work might examine whether these effects get stronger as more off-label FCA suits are brought.

[132] 42 U.S.C. § 1396r-8.

[133] While, according to the *Parke Davis* court, Medicaid does not gather information on indication in their reimbursement forms, which would make monitoring usage more difficult, this is theoretically something the government could require.

become increasingly so after FERA. Thus, to the extent that the federal government both has control over setting reimbursement guidelines and is the sole payer, I might expect that the government program might invest less time in disseminating information on inappropriate uses than a payer who has a harder time recovering through a fraud statute.

It is important to note that there are agency issues when it comes to Medicaid reimbursement. Although the *Parke Davis* court side-stepped the question of heterogeneity in reimbursement in state Medicaid courts (Greene 2005), this might be a concern for future cases (Martin 2004). If a state allows for the reimbursement of off-label non-compendia uses, they may not be able to recover for the prescription. They would have to prove that the particular use was still nonreimbursable in their state. In the neurontin case, this did not matter as at least eight states did not allow reimbursement, providing sufficient basis for the case to go forward; the issue only was relevant in terms of damages (and the case subsequently settled).[134] Medicaid programs, however, might take this uncertainty into account in policing their reimbursements. Insofar as state programs do reimburse uses not listed on the compendia, they have greater incentive to relinquish inappropriate ones as soon as they can, as recovering for these uses might be more difficult.

These results have interesting implications for future work. The data from Section IV and V are largely from the pre-FERA period. With the expansion of the scope of the FCA provided

[134] U.S. ex rel. Franklin v. Parke-Davis, Div. of Warner-Lambert Co., No. CIV.A. 96-11651PBS, 2003 WL 22048255, at *3 (D. Mass. Aug. 22, 2003). This type of rationale was recently used to defend against a motion to dismiss in U.S. ex rel. Booker v. Pfizer, Inc., 9 F. Supp. 3d 34 (D. Mass. 2014). *See also* U.S. ex rel. Banigan v. Organon USA Inc., 883 F. Supp. 2d 277, 294 (D. Mass. 2012) (noting that "Organon contends that if a state Medicaid program chooses to reimburse a claim for a drug prescribed for off-label use, then that claim is not "false or fraudulent," and liability cannot therefore attach for reimbursement. The court agrees.").

in FERA and Kaiser's RICO victory in 2013, it would be interesting to see whether the differences in payer response is accentuated in the most recent years. The fact that Kaiser won their RICO suit may prompt other private insurance companies to follow the steps that Kaiser took to recover. If FERA simultaneously causes public payers to be less likely to preemptively relinquish inappropriate off-label uses, the gap in payer response might widen in the future. Similarly, it will be interesting to see whether cases based on Medicare Part D reimbursements display the relinquishment behavior that this Chapter hypothesizes a purely federal payer would exhibit. Alternatively, Medicare patients' behavior may mimic that of private insurance patients given that Medicare Part D is managed by private insurance companies. The results of this Chapter indicate that this line of inquiry is important and should be further pursued.

C. Conclusion

This Chapter sought to determine whether False Claim Act suits can function as a source of information for physicians and third party payers regarding the appropriateness of off-label uses. The importance of quick relinquishment of inappropriate off-label uses is two-fold: it reduces monetary waste on inappropriate prescriptions, and it allows patients to switch to more appropriate treatments. The FCA was not created to act as a signal for inappropriate treatment; its purpose was to prevent fraud to the government. A positive externality of FCA suit, however, is that it is a visible industry notice in an area that is not directly regulated by the government. When used to target inappropriate off-label promotion, a happy consequence of the FCA suit is that it provides information to other third-party payers that an off-label use might not be appropriate.

This Chapter finds evidence that off-label prescriptions decline after an FCA suit: Part IV studies FCA settlement, and Part V studies other landmarks in the FCA case. While this is a

relatively positive result, indicating that the FCA suit can serve as an industry-wide signal, there are important caveats. Further legal changes might result in two effects. First, the increasingly broad powers of the FCA and a RICO win for one third-party payer may accentuate the observed difference in payer response. Second, the FCA's increasingly expansive powers may dilute the signal of the FCA, such that inappropriate and appropriate off-label uses are subjects for settlements and settlement is less informative.

Additionally, some functions of information spreading that the FCA engages in, specifically publicizing obscure scientific findings, might be better done by a government agency than by lawsuit. The government does not formally regulate off-label promotion, but if government suits function as information-spreading devices, it may be better to regulate promotion directly by providing a digest summarizing the scientific information publicly available on a particular drug use.

Other functions, such as punishing fraudulent statements, may require the use of the False Claims Act. Attempts to directly regulate off-label *promotion* through misbranding often encounter First Amendment challenges; the FCA has currently avoided such challenges. Until the government develops a clear way to require appropriate off-label promotion through misbranding, punishing fraudulent promotion through the FCA remains an important option.

The False Claims Act is a powerful tool for controlling off-label promotion; although its purpose is as a reimbursement tool for the government, this Chapter finds that it may also provide a signal to other payers and physicians that the particular off-label use is inappropriate. Given the difficulties the government faces in regulating off-label uses, this function is important; whether the expansion of the FCA will undermine this function, however, remains to be seen.

VII. References

Abbott, Ryan and Ian Ayres. 2014. "Evidence and Extrapolation: Mechanisms for Regulating Off-Label Uses of Drugs and Devices" (February 21, 2014). *Duke Law Journal* 64: 377, available at: http://ssrn.com/abstract=2399511. [hereafter Abbot and Ayres(2014b)].

Abbott, Ryan and Ian Ayres. 2014. "Can Bayesian Extrapolation Improve FDA Regulation of Off-Label Uses of Drugs and Devices?" *Food and Drug Policy Forum,* 4(5):1–12. [hereafter Abbot and Ayres(2014a)].

Armstrong, David. 2003. "How Drug Directory Helps Raise Tabs for Medicaid and Insurers." *The Wall Street Journal*, October 23, 2003, available at http://online.wsj.com/article/0,,SB106685564225943200,00.html.

Barker, Thomas and Ross Margulies. 2014. "Centers for Medicare & Medicaid Services Proposes Changes to Six Protected Class Rule under Medicare Part D." *Foley Hoag LLP,* available at http://www.foleyhoag.com/publications/alerts-and-updates/2014/january/cms-proposes-changes-to-six-protected-class-rule-under-medicare-part-d.

Barger Jr, James F., Pamela H. Bucy, Melinda M. Eubanks, and Marc S. Raspanti. 2005. "States, Statutes, and Fraud: An Empirical Study of Emerging State False Claims Acts." *Tulane Law Review* 80: 465.

Blair, Katherine A. 2010. "In Search of the Right RX: Use of the Federal False Claims Act in Off-Label Drug Promotion Litigation." *Health Lawyer* 23: 44.

Center for Disease Control (CDC). 2012. "Smoking and Tobacco Use." http://www.cdc.gov/tobacco/data_statistics/fact_sheets/fast_facts/ (accessed March 25, 2013).

Center for Medicare Advocacy (CMA). 2010. *CMA Report: Medicare Coverage for Off-Label Drug Use*, available at: http://www.medicareadvocacy.org/cma-report-medicare-coverage-for-off-label-drug-use/#_edn11.

Chan, Tat, Chakravarthi Narasimhan, and Ying Xie. 2013. "Treatment Effectiveness and Side Effects: A Model of Physician Learning." *Management Science* 59(6): 1309–1325.

Chandra, Amitabh, David Cutler, and Zirul Song. 2011. "Who Ordered That? The Economics of Treatment Choices in Medical Care." In *Handbook of Health Economics*, Mark V Pauly, Thomas G. McGuire, and Pedro Pita Barros (eds.), 398–432. Amsterdam, NLD: North Holland.

Chintagunta, Pradeep, Ronald Goettler, and Minki Kim. 2012. "New Drug Diffusion when Forward-Looking Physicians Learn from Patient Feedback and Detailing." *Journal of Marketing Research* 49(6): 807–821.

Cooney, J. Gordon, Jr., John P. Lavelle, Jr., and Bahar Shariati. 2010. Back to the Future: Civil Rico in Off-Label Promotion Litigation, *Defense Counsel Journal* 77: 168.

Covell, David G., Gwen C. Uman, and Phil R. Manning. 1985. "Information Needs in Office Practice: Are They Being Met?" *Annals of Internal Medicine* 103(4), 596–599.

Cox, Jack. 2007. "Court Rules for Pfizer in Patent Infringement Case on Neurontin." *Pfizer*, *available at* http://www.pfizer.com/news/press-release/press-release-archive-detail/court_rules_for_pfizer_in_patent_infringement_case_on_neurontin (accessed May 28, 2014).

Department of Justice (DOJ). 2004. "Warner-Lambert to Pay $430 Million to Resolve Criminal & Civil Health Care Liability Relating to Off-Label Promotion." May 13, 2004. http://www.justice.gov/opa/pr/2004/May/04_civ_322.htm (accessed March 11, 2013).

Eichel, Kristin McCreary. 2011. "Focusing on Fraud: The Federal Government Expands Its Use of the False Claims Act to Police Off-Label Pharmaceutical Promotion." *Indiana Health Law Review* 8: 399.

Ely, John W., Jerome A. Osheroff, M. Lee Chambliss, Mark H. Ebell, and Marcy E. Rosenbaum. 2005. "Answering Physicians' clinical questions: Obstacles and Potential Solutions." *Journal of the American Medical Informatics Association* 12(2): 217–224.

Fisk, Margaret Cronin, Elizabeth Lopatto and Jef Feeley. 2012. "Lilly Sold Drug for Dementia Knowing It Didn't Help, Files Show." *Bloomberg*, June 12, 2012. http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aTLcF3zT1Pdo (accessed March 13, 2013).

Food and Drug Administration (FDA). 2004. "Equetro - Drugs@FDA." available at: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/PediatricAdvisoryCommittee/UCM234474.pdf (accessed March 15, 2013).

Food and Drug Administration (FDA). 2005. "Neurontin – Safety Information." available at: http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm151865.htm (accessed April 14, 2013).

Food and Drug Administration (FDA). 2008. "Stavzor - Drugs@FDA." available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2008/022152lbl.pdf (accessed March 15, 2013).

Food and Drug Administration (FDA). 2009. "Good Reprint Practices for the Distribution of Medical Journal Articles and Medical or Scientific Reference Publications on Unapproved New Uses of Approved Drugs and Approved or Cleared Medical Devices."

available at: http://www.fda.gov/RegulatoryInformation/Guidances/ucm125126.htm (accessed March 5, 2013).

Food and Drug Administration (FDA). 2010. "2005 Safety Alerts for Human Medical Products." available at: http://www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm151239.htm (accessed April 14, 2013) [hereafter (FDA 2010a)].

Food and Drug Administration (FDA). 2010. "Lamictal Labeling Changes Overview." available at: http://www.accessdata.fda.gov/drugsatfda_docs/label/2006/020241SLRs10s21s25s26s27,020764SLRs3s14s18s19s20lbl.pdf (accessed January 28, 2016) [hereafter (FDA 2010b)].

Food and Drug Administration (FDA) 2012. "The FDA's Drug Review Process: Ensuring Drugs Are Safe and Effective." http://www.fda.gov/drugs/resourcesforyou/consumers/ucm143534.htm (accessed April 15, 2013).

Food and Drug Administration (FDA). 2014. "Guidance for Industry Distributing Scientific and Medical Publications on Unapproved New Uses — Recommended Practices." available at: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm387652.pdf (accessed January 20, 2016).

Fox, Peter D. 2003. "Prescription Drug Benefits: Cost Management Issues for Medicare." *Health Care Financing Review, 25*(2), 7.

Frye, Mark A., Terence A. Ketter, Timothy A. Kimbrell, Robert T. Dunn, Andrew M. Speer, Elizabeth A. Osuch, David A. Luckenbaugh, Gabriela Corá-Locatelli, Gabriele S.

Leverich, and Robert M. Post. 2000. "A Placebo-Controlled Study of Lamotrigine and Gabapentin Monotherapy in Refractory Mood Disorders." *Journal of Clinical Psychopharmacology* 20(6): 607–614.

Fullerton, Catherine A., Alisa B. Busch, and Richard G. Frank. 2010. "The Rise and Fall of Gabapentin for Bipolar Disorder: A Case Study on Off-Label Pharmaceutical Diffusion." *Medical Care* 48(4): 372–379.

Gibbons, David C. 2015. "A Victory for Amarin Further Erodes FDA Regulation of Off-Label Promotion." *FDA Law Blog*, August 10, available at http://www.fdalawblog.net/fda_law_blog_hyman_phelps/2015/08/a-victory-for-amarin-further-erodes-fda-regulation-of-off-label-promotion.html.

Greene, Stephanie M. 2014. "After Caronia: First Amendment Concerns in Off-Label Promotion." *San Diego L. Rev.* 51: 645.

Greene, Stephanie M., and Lars Noah. 2014. "Debate: Off-Label Drug Promotion and the First Amendment." *University of Pennsylvania Law Review* 62: 239–267. available at SSRN: http://ssrn.com/abstract=2458391

Greene, Stephanie M. 2005. "False Claims Act Liability for Off-Label Promotion of Pharmaceutical Products." Penn St. L. Rev. 100: 41.

Hirschfeld, Robert, Charles L. Bowden, Michael J. Gitlin, Paul E. Keck, Trisha Suppes, Michael E. Thase, and Roy H. Perlis. *Practice Guideline for the Treatment of Patients with Bipolar Disorder*. American Psychiatric Association, 2002.

Hoffman, Megan L. 2009. "The Substantial Weight Test: A Proposal to Resolve the Circuits' Disparate Interpretations of Materiality Under the False Claims Act." *University of Kansas Law Review* 58: 181.

Johnson, Sandra H. 2007. “Polluting Medical Judgment? False Assumptions in the Pursuit of False Claims Regarding Off-Label Prescribing.” *Minnesota Journal of Law, Science, and Technology*, 9(1), 61–124.

Kesselheim, Aaron S., Devan Darby, David M. Studdert, Robert Glynn, Raisa Levin, and Jerry Avorn. 2011. "False Claims Act prosecution did not deter off-label drug use in the case of Neurontin." *Health Affairs* 30(12): 2318–2327.

Kesselheim, Aaron S., Christopher T. Robertson, Jessica A. Myers, Susannah L. Rose, Victoria Gillet, Kathryn M. Ross, Robert J. Glynn, Steven Joffe, and Jerry Avorn. 2012. “A Randomized Study of How Physicians Interpret Research Funding Disclosures.” *New England Journal of Medicine* 367(12), 1119–1127.

Lenzer, Jeanne. 2004. “Pfizer Pleads Guilty but Drug Sales Continue to Soar.” *BMJ*, May 20, 2004, available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC416587/ (accessed March 25, 2013).

Longstreth, Andrew. 2014. “Pfizer Agrees to $190 million settlement over generic Neurontin.” *Reuters.com*, April 21. http://www.reuters.com/article/2014/04/21/us-usa-antitrust-pfizer-idUSBREA3K17420140421 (accessed May 30, 2014).

Majumdar, Sumit R., Thomas S. Inui, Jerry H. Gurwitz, Matthew W. Gillman, Thomas J. McLaughlin, and Stephen B. Soumerai. 2001. “Influence of Physician Specialty on Adoption and Relinquishment of Calcium Channel Blockers and Other Treatments for Myocardial Infarction.” *Journal of General Internal Medicine*, 16(6): 351–359.

Mapes, Roy E. A. 1977. “Physicians’ Drug Innovation and Relinquishment.” *Social Science and Medicine* 11(11): 619–624.

Martin, Benjamin S. 2004. “Medicaid Coverage for Drugs for Off-Label Uses.” Arnold & Porter LLP. Available at National Institute of Mental Health http://www.ehcca.com/presentations/pharmacongress5/martin_2.pdf

National Institute of Mental Health (NIMH). 2013. “How is Bipolar Disorder Treated?” available at: http://www.nimh.nih.gov/health/publications/bipolar-disorder/how-is-bipolar-disorder-treated.shtml (access April 7, 2013).

Neafsey, Patricia J. 2005. “Medication News: Lyrica (Pregabalin): Neurontin Replacement.” *Home Healthcare Nurse* 23(9): 563-564.

Obrocea, Gabriela V., Robert M. Dunn, Mark A. Frye, Terence A. Ketter, David A. Luckenbaugh, Gabriele S. Leverich, Andrew M. Speer, Elizabeth A. Osuch, Kamal Jajodia, and Robert M. Post, 2002. "Clinical Predictors of Response to Lamotrigine and Gabapentin Monotherapy in Refractory Affective Disorders." *Biological Psychiatry* 51(3): 253–260.

Pande, Atul C., Jerri G. Crockatt, Carol A. Janney, John L. Werth, and Georgia Tsaroucha. 2000. "Gabapentin in Bipolar Disorder: A Placebo-Controlled Trial of Adjunctive Therapy." *Bipolar Disorders* 2(3): 249–255.

Philip, Elissa. "United States v. Carolina: How True Does Truthful Have to Be." *Vanderbilt Law Review En Banc* 67 (2014): 157.

Prakash, Snigdha. “Court Files Yield New Information in Suit Against Drugmaker.” *NPR* (November 2, 2002).

Prakash, Snigdha. “Neurontin Lawsuit.” *NPR* (June 18, 2002).

Purse, Marcia. 2012. “Suit: Neurontin Marketed Illegally for Bipolar Disorder” available at: http://bipolar.about.com/cs/neurontin/a/neurontin_suit.htm (accessed May 1, 2014).

Radley, David C., Stan N. Finkelstein, and Randall S. Stafford. 2006. "Off-Label Prescribing Among Office-Based Physicians." *Archives of Internal Medicine* 166(9): 1021–1026, available at: http://archinte.jamanetwork.com.proxy.library.vanderbilt.edu/article.aspx?articleid=410250.

Riley, James B., Jr., and P. Aaron Basilius. 2007. "Physicians' Liability for Off-Label Prescriptions." *Hematology & Oncology News & Issues* available at: http://www.mcguirewoods.com/news-resources/publications/health_care/Off_Label.pdf (access March 5, 2013).

Robertson, Christopher. 2014. "When Truth Cannot Be Presumed: The Regulation of Drug Promotion Under an Expanding First Amendment." *Boston University Law Review*, 94: 545.

Rogoff, Michael, Manvin Mayell, and Paula Ramer, 2014. *The Aftermath of* Caronia *in Pursuing Off-Label Cases,* INSIDECOUNSEL, archived at http://perma.cc/92QE-AHPR.

Stafford, Randall S. 2008. "Regulating Off-Label Drug Use – Rethinking the Role of the FDA." *New England Journal of Medicine*, 358(14): 1427–1429.

Vieta, Eduard, José Manuel Goikolea, Anabel Martinez-Aran, Merce Comes, Katia Verger, Xavier Masramon, Jose Sanchez-Moreno, and Francesc Colom, 2006. "A Double-Blind, Randomized, Placebo-Controlled, Prophylaxis Study of Adjunctive Gabapentin for Bipolar Disorder." *Journal of Clinical Psychiatry* 67(3): 473–477.

Viscusi, W. Kip. 1997. "Alarmist Decisions with Divergent Risk Information." *The Economic Journal* 107(445): 1657–1670.

Viscusi, W. Kip and Richard J. Zeckhauser. 2015. “Regulating Ambiguous Risks: The Less than Rational Regulation of Pharmaceuticals.” *Journal of Legal Studies* 44(2): S387–S422.

Williamson, John W., Pearl S. German, Robin Weiss, Elizabeth A. Skinner, and Frederick Bowes, 1989. "Health Science Information Management and Continuing Education of Physicians: A Survey of US Primary Care Practitioners and Their Opinion Leaders." *Annals of Internal Medicine* 110(2): 151–160.

Williams, John W., Leah Ranney, Laura C. Morgan, and Lynn Whitener. 2009. "How Reviews Covered the Unfolding Scientific Story of Gabapentin for Bipolar Disorder." *General Hospital Psychiatry* 31(3): 279–287.

Wittich, Christopher M., Christopher M. Burkle, and William L. Lanier 2012. “Ten Common Questions (and Their Answers) About Off-Label Drug Use.” *Mayo Clinic Proceedings* 87(10): 982–990.