Positively Aware?
Conflicting Expert Reviews and Demand for Medical Treatment*

Jorge F. Balat†
Department of Economics, The University of Texas at Austin

Nicholas W. Papageorge‡
Department of Economics, Johns Hopkins University, IZA, and NBER

Shaiza Qayyum§
Economic Advisory, Ernst & Young

June 14, 2019

Abstract: We study the impact of expert reviews on the demand for HIV treatments. A novel feature of our study is that we observe two reviews, provided by both a doctor and an activist in an HIV lifestyle magazine *Positively Aware*, for each HIV drug and focus attention on consumer responses when experts disagree. To establish a causal relationship between reviews and demand, we exploit the arrival of new drugs over time, which provides arguably random variation in reviews of existing drugs. We find that when doctors and activists agree, positive reviews increase demand for HIV drugs. However, doctors and activists frequently disagree, most often over treatments that are effective, but have harsh side effects, in which case they are given low ratings by the activist, but not by the doctor. In such cases, relatively healthy consumers favor drugs with higher activist reviews, which is consistent with a distaste for side effects. This pattern reverses for individuals who are in worse health and thus face stronger incentives to choose more effective medication despite side effects. Findings suggest that consumers demand information from experts whose review is more aligned to their preferences over health versus side effects, which can vary by health status.

Keywords: Health; Information; Product Reviews; Pharmaceutical Demand; HIV/AIDS.
JEL Classification: I12, L15, M3, D12, D83.

*We gratefully acknowledge helpful comments from Tat Chan, Andrew Ching, Michael Darden, Gautam Gowrisankaran, Barton Hamilton, Ginger Jin, Stephanie Heger, Mitchell Hoffman, Jennifer Kohn, Darius Lakdawalla, Michael Makowsky, Patrick McAlvanah, Kevin Thom and Yiyi Zhou along with seminar participants at the FTC, the 2015 SEA meetings, the 2016 NASM of the Econometric Society and the 2016 Yale SOM Marketing Science Conference at Olin Business School. The usual caveats apply.

†jbalat@utexas.edu.
‡[Corresponding author] papageorge@jhu.edu.
§sqayyum2@jhu.edu.

1 Introduction

Consumers facing uncertainty often turn to low-cost sources of information, such as word-of-mouth, advertisements or product reviews generated by other consumers or by experts. In the case of expert reviews, the idea is that individuals turn to a trusted or authoritative information source to help them make decisions. Previous research has demonstrated that expert reviews help drive demand for a number of products, such as movies, wine and books. The impact of expert reviews also extends to higher-stakes contexts, such as financial decisions (Benabou and Laroque, 1992; Reiter and Ziebart, 1991; Cantor and Packer, 1996, 1997).[1]

Despite the importance of expert reviews in several economic contexts, little is known about how consumer demand responds to conflicting expert reviews. Yet, consumers often have access to multiple reviews from different experts who potentially disagree. One possibility, which we explore in this paper, is that individuals facing conflicting reviews rely upon expertise from the source they view as best aligned to their preferences and objectives. Seen this way, individuals facing uncertainty are not passive consumers of available information, but instead appear to actively choose which information source to incorporate into their decisions.

In this paper, we study the impact of expert reviews on the demand for HIV drugs.[2] In our setting, consumers face uncertainty about drug qualities, including treatment efficacy and adverse treatment side effects. Their choices affect their health, well-being and survival. At multiple points in its lifecycle, each HIV drug we study is reviewed by both an HIV physician and an HIV activist, the latter often someone infected with HIV. We demonstrate that favorable expert reviews increase demand for HIV drugs. This finding provides evidence that the influence of expert reviews extends to health investments. Next, we examine patient responses to conflicting reviews, i.e., when the doctor and activist disagree about a given drug. In such cases, consumer responses vary by their current health, with sicker consumers choosing treatments recommended by the doctor and healthier consumers following the activist. To explain this pattern, we argue that consumer responses to expert reviews depend on their incentives to use effective treatments despite adverse side effects — and that these incentives shift with health status.

Examining consumer responses to conflicting reviews requires data that are often lacking in studies linking expert reviews to demand. To study HIV drug reviews and demand,

[1] Coffman et al. (2017) show evidence that social information affects decisions in high-stakes contexts, in their case, career choices.

[2] HIV stands for Human Immunodeficiency Virus, which is a virus that attacks the immune system.

we merge two unique datasets. The first is from a longitudinal study of men infected with HIV (henceforth, HIV-positive or HIV+), which provides detailed information on a variety of health measures and also records each individual's medical treatment consumption decisions. Using this dataset, we can relate patient health outcomes to the treatments they use, which allows us to construct two objective treatment characteristics: a measure of treatment effectiveness against HIV and a measure of treatment side effects. We merge this information with a unique dataset consisting of manually-coded drug reviews. Doctor and activist reviews are disseminated in a comprehensive HIV drug guide published annually in a widely circulated HIV lifestyle magazine called *Positively Aware*. As we explain in detail below, text reviews are scored as positive, negative or mixed.[3] By combining these two datasets, we are able to relate potentially conflicting activist and doctor reviews to drug demand. Moreover, since we observe objective product qualities, we not only control for them, but also use them to better understand how reviews are generated, in particular, why doctors and activists sometimes disagree about a given drug.

We present two main sets of results. First, we estimate a discrete choice model of demand for HIV drugs and provide arguably causal evidence that positive reviews increase demand for HIV drugs. A positive correlation between positive reviews and high demand could be driven by omitted third factors affecting both reviews and demand. We address potential endogeneity problems in two ways. First, we exploit rich data on objective product qualities to control for drug characteristics in our demand equations. Including these in the demand equations means we envision consumers as having some information about drug characteristics, but also facing uncertainty, which means they turn to expert reviews for additional information when making choices.[4] Still, the concern remains that there are unobserved drug characteristics that could simultaneously drive higher reviews and demand. To address this concern, in our main analyses we exploit variation in the choice set over time. We rely on the observation that, as new drugs are introduced, reviews for existing drugs shift in response. Thus, we use the objective qualities of rival drugs on the market, which change over time as the market evolves, to instrument for current-period reviews. This identification strategy follows the spirit of Berry et al. (1995) (henceforth, BLP), as we exploit characteristics of a shifting set of rival products on the market to instrument for a drug's review. Estimates indicate that reviews have a modest positive impact on demand when both experts give the

[3]Drug reviews from *Positively Aware* also contain information on a host of additional drug characteristics, including known interactions, dosage and number of side effects discovered during clinical trials, information which we also use in our empirical analysis.

[4]This logic is formalized in a simple model in Appendix D. Our main qualitative results are robust to the exclusion of observed drug characteristics. However, the magnitude of the impact of reviews declines when we include them, which is in line with the idea that reviews reflect drug characteristics that also drive demand.

same rating to a treatment. However, when the two experts disagree, a higher rating by the activist increases demand by 3.2 percentage points while a doctor's review lowers demand by 2.8 percentage points.[5]

Our second set of results focuses on explaining consumer demand responses to conflicting reviews.[6] We find that, when the reviews of doctors and activists diverge, relatively healthy patients follow the activist's rather than the doctor's advice. Our interpretation is that this pattern reflects patient distaste for drug side effects. To support this view, we provide two pieces of empirical evidence. First, disagreements arise when a treatment is highly effective but has severe side effects, in which case it is given a lower review by the activist, but not by the doctor.[7] If patients favor drugs with fewer side effects and face diverging reviews, they might choose to follow the expert — in this case an HIV activist who is also a fellow patient — who tends to downgrade drugs with harsher side effects. Second, we examine demand responses of HIV+ men who are relatively sick (a condition known as AIDS).[8] Previous research has shown that patients who choose less effective treatments to avoid side effects are more willing to choose effective treatments with adverse side effects when in poor health since the payoff from doing so in terms of improved health is large (Papageorge, 2016). This suggests a way to test the validity of our preferred explanation. If healthier consumers follow the activist in an effort to avoid side effects, we would expect sicker patients to respond more positively to the doctor, the expert who tends to recommend highly effective treatments despite adverse side effects.[9] Indeed, we find that, in contrast to healthier patients, sicker HIV+ men respond positively to higher doctor reviews.

A broad implication of our findings is that it is potentially problematic to view patients as passive consumers of available or posted information. Depending on whether information is relevant to them or is aligned with their preferences, consumers actively select whether they are exposed to the information and the extent of their exposure.[10] This has important implications for how we should interpret estimated effects of information interventions on behavior. Even if individuals are randomly exposed to information, (e.g., in a controlled trial) they may non-randomly select how much and which type of information to incorporate into their decisions, resulting in endogenous selection into treatment. Average treatment effects

[5]The average market share for all combinations on the market, excluding the outside option, is 2.3%.

[6]As reviews are printed on the same page, one after the other, in the same magazine, the presumption is that individuals are exposed to both.

[7]This is in line with research demonstrating that doctors care less about side effects than patients do. In a particularly striking contribution, Ubel et al. (2011) show that doctors, when they fall ill, avoid drugs with side effects despite having recommended them to their patients.

[8]AIDS stands for Acquired Immune Deficiency Syndrome.

[9]To fix ideas, Appendix D presents a theoretical model that formalizes the logic behind our interpretation of the findings.

[10]We thank an anonymous referee for suggesting this point.

from randomized information interventions could therefore obfuscate systematic differences in which types of consumers self select into obtaining and using which type of information. If so, extrapolating estimated effects of information on behavior to new populations could be misleading.

In relating expert reviews to demand, this paper contributes to several strands of literature in economics. The first studies how individuals facing uncertainty rely on a variety of information sources, such as direct-to-consumer advertising (Ackerberg, 2001; Gordon and Hartmann, 2013) or social learning, which includes word-of-mouth and peer effects (Moretti, 2011; Liu, 2006).[11] Other research has studied the effect of online reviews. Anderson and Magruder (2012) demonstrate how *Yelp* reviews affect restaurant choices, and Cabral and Hortacsu (2010) show that *eBay* reputation affects purchasing decisions. In these settings, reviews are usually posted by past purchasers of a product, and reviews may be biased toward the extremes. In our setting, reviewers are chosen by the publication and should be more credible. More closely related to our study, a number of papers show that "report cards" revealing information about product quality can affect choices when quality is uncertain.[12] In our study, we incorporate consumer-level data, which means we can study the impact of reviews on consumer choices and how consumer characteristics determine the way consumers demand information from outside sources.

An advantage of our study is that we incorporate information on objective product qualities along with reviews from multiple, possibly conflicting experts. This allows us to examine how reviews relate to objective product qualities along with heterogeneity in consumer responses to disagreements. We can thus provide novel evidence that the way in which experts weight product qualities in their reviews affects how consumers incorporate these reviews into their decisions. This point relates our study to an emerging literature examining the demand for information. For example, Eliaz and Schotter (2010) show that when making risky decisions, agents pay for information based on the likelihood of information being ex-post optimal. Relatedly, Dranove and Sfekas (2008) show that when hospital report cards provide information that differ from patients' prior beliefs about hospital quality, patients switch to higher-quality hospitals.

[11] The impact of social learning on demand has been shown in a variety of contexts, including the adoption of new crops (Munshi, 2004; Bandiera and Rasul, 2006; Conley and Udry, 2010) and job uptake (Coffman et al., 2017). See Dranove and Jin (2010) for a comprehensive review.

[12] Information in the form of audits or report cards affects election winners (Ferraz and Finan, 2008), stock-buying (Barber and Odean, 2008), Medicare enrollment (Dafny and Dranove, 2008), health plan choice (Chernew et al., 2008), health care provider choice (Wang et al., 2011), hospital patient volumes (Pope, 2009), investments in the housing market (Figlio and Lucas, 2004), restaurant demand (Jin and Leslie, 2003) and education (Andrabi et al., 2017; Hastings and Weinstein, 2008). Fong and Oberholzer-Gee (2011) show that agents are willing to pay for information about charity recipients when agents' charitable giving is responsive to recipient type.

By focusing on disagreements among experts in high-stakes contexts, we also relate to a literature demonstrating that reliance on low-cost information sources, such as expert reviews, can be problematic. For example, Dranove et al. (2003) show that information contained in health care "report cards" decreased patient and social welfare by inducing health care providers to decline treatment to sicker patients. Mayzlin et al. (2014) find that online hotel reviews that affect demand are subject to manipulation. Relatedly, in a study of expert judges of a musical competition, Ginsburgh and Van Ours (2003) show evidence that judges' rankings are often the result of random ordering of the performers and not the underlying performance quality. Yet, judges' rankings affect performers' subsequent careers.[13] The idea is that reviews, either from experts or other users, might not provide useful or accurate information, but could still affect economic decisions and outcomes. The disagreements between reviewers that we examine might suggest that at least one expert is "wrong," which could mean that reliance on reviews could harm patients. Our findings suggest a different interpretation. We argue that disagreements reflect that experts generate reviews that place different weights on multiple drug characteristics. We show that consumers respond differently to divergent reviews, which suggests that they demand different information depending on their current health status and follow conflicting expert reviews accordingly.[14]

Finally, we contribute to a literature examining health investments under uncertainty. For example, Crawford and Shum (2005) show the effects of uncertainty and learning in the demand for anti-ulcer drugs.[15] Coscelli and Shum (2004) model how doctors update their beliefs about drug quality relative to existing drugs after observing the new drug's effects on their patients. Further studies examine how direct-to-consumer advertising (Sinkinson and Starc, 2015), spillover effects from advertising of similar drugs (Shapiro, 2018), detailing (Ching and Ishihara, 2010, 2012) and publicity (Ching et al., 2015) affect demand for pharmaceuticals when drug quality is uncertain.[16] There is also evidence of peer effects in healthcare adoption (Duflo and Saez (2003); Sorensen (2006); Oster and Thornton (2012)).[17]

[13]Relatedly, Bertrand et al. (2010) show that a picture of a smiling woman on a loan brochure affects demand for the loan.

[14]It is possible that reviews are biased in other ways, for example, if reviewers are paid to give higher ratings to low-quality drugs. This would not affect our conclusion that reviewers provide conflicting reviews that affect demand. It would suggest that reviews could lower patient welfare by steering them to low-quality drugs.

[15]Related to learning, Dickstein (2014) designs a framework to analyze how price and promotion influence the learning process of the doctor and the patient and applies his model to depression care.

[16]Related, Liu et al. (2014) study promotion spillovers in demand for HIV drugs.

[17]Theoretical work on social learning from peers can be traced back to Banerjee (1992) and Bikhchandani et al. (1992), who show that informational cascades can explain herd behavior and fads. Schotter (2003) presents a theoretical model of decision making with advice from outside sources (such as word-of-mouth advice and observational learning). Brown et al. (2012, 2013) write a behavioral game-theoretic model to explain limited strategic thinking at the movie box-office.

We show evidence of a novel way that consumers making health investments mitigate uncertainty: by incorporating expertise from potentially conflicting sources in a way that depends on their health objectives.

The rest of this paper is organized as follows. Section 2 discusses our data sources and how we code the expert reviews. In Section 3 we explain how we construct the combination-level data (as HIV drugs are consumed in combination with one another), and presents a preliminary analysis at the drug-combination level. Section 4 describes our econometric model and identification strategy. In Section 5 we present our main results. Finally, Section 6 concludes.

2 Data: Individuals, Drugs and Reviews

In this section, we introduce the individual- and drug-level datasets we use in this study and explain how we code the expert reviews.

2.1 Data Sources

Our data come from two sources. The first is a large panel dataset on HIV+ men's treatment choices and health outcomes. The second is a novel dataset containing expert drug reviews written in the magazine *Positively Aware*.

Data from the Multi-Center AIDS Cohort Study. We use the publicly available dataset from the Multi-Center AIDS Cohort Study (henceforth, MACS), an ongoing study of the natural and treated histories of HIV+ homosexual and bisexual men that was started in 1983.[18] The study is conducted in four cities: Baltimore, Chicago, Pittsburgh, and Los Angeles. There have been three enrollments for MACS. First enrollment was in 1984 (4954 enrolled). Second enrollment in 1987 when 668 additional men enrolled. A third enrollment in 2001 added 1350 men to the sample. The MACS dataset is not a random patient sample.[19] At each semi-annual visit (conducted in March and September of each

[18]The study also follows HIV-negative (henceforth, HIV−) men, but we exclude them from our analysis since over our sample period it is exceedingly rare that uninfected men consume HIV drugs.

[19]Data in this manuscript were collected by the Multi-Center AIDS Cohort Study (MACS) with centers (Principal Investigators) at The Johns Hopkins Bloomberg School of Public Health (Joseph B. Margolick, Lisa P. Jacobson); Howard Brown Health Center, Feinberg School of Medicine, Northwestern University; Cook County Bureau of Health Services (John P. Phair, Steven M. Wolinsky), University of California, Los Angeles (Roger Detels); and University of Pittsburgh (Charles R. Rinaldo). The MACS is funded by the National Institute of Allergy and Infectious Diseases, with additional supplemental funding from the National Cancer Institute. UO1-AI-35042, 5-MO1-RR-00052 (GCRC), UO1-AI-35043, UO1-AI-35039, UO1-AI-35040, UO1-AI-35041. Website located at http://www.statepi.jhsph.edu/macs/macs.html.

year), data are collected on medical treatment choices, health status and a host of socio-demographic measures, including employment, income and completed education. The MACS dataset consists of 6,843 individuals over 50 (semi-annual) visits. We restrict our attention to HIV+ individuals for the time period from 1997 to 2008, which is when drug reviews from the *Positively Aware* Drug Guides — our second data source — are available. Restricting our sample to individuals who were part of the sample between 1997 and 2008 and had no missing information on key variables such as reported side effects, health status, age, education, and work status leaves us with an unbalanced panel of 1,330 individuals consisting of 13,472 observations, where each observation is an individual-visit dyad.

The MACS dataset not only provides us with individual-level drug choices but also includes two measures of health status relevant to individuals with HIV. The first is an objective measure of the individual's immune system health. At each interview, a blood test is conducted to measure the subject's CD4 count, which is defined as the number of white blood cells per mm^3 of blood. Typical CD4 counts range between 500 and 1000 for uninfected (HIV−) individuals and lower counts indicate that the immune system is compromised by HIV. Counts below 300 indicate the individual suffers from AIDS, a condition where the immune system has been compromised to such a degree that it loses functionality and cannot fight off common infections. The second health measure consists of subjects' own reports of their physical ailments, including nausea, headache, fever, diarrhea and drenching sweats. These physical ailments reflect side effects of medical treatments, but can also be symptoms of HIV infection if CD4 counts are low.

We report summary statistics for the variables at the individual level in our sample in Table 1. The average age of individuals in the sample is 47, with 54% of the sample composed of white individuals. Close to 20% of the individuals have only a high school degree, while 50% of the sample has completed college education, and 54% of people work full time. The average CD4 count of individuals in our sample is 536, with 54% of individuals reporting non-decreasing CD4 count from one visit to the next and 63% of the patients reporting no ailments such as fatigue, drenching sweats, and headaches. Relevant to our later analyses, 20% of patients have a CD4 count of less than 300, which indicates that they are living with AIDS.

Data from *Positively Aware*. The second data source we use is a drug guide published annually since 1997 in an HIV lifestyle magazine known as *Positively Aware*, which contains drug reviews for all FDA-approved drugs and drugs nearing approval.[20] While the magazine

[20] *Positively Aware* is a not-for-profit HIV/AIDS treatment journal published by Test Positive Aware Network (TPAN). TPAN is a 501c3, not-for-profit AIDS Service Organization (ASO) based in Chicago, IL.

is issued bimonthly (six regular issues per year), the comprehensive drug guide is published annually joint with the January/February issue. The magazine's contributing writers and columnists are professionals in the field of HIV/AIDS, including HIV specialist physicians from the US, people living with HIV, and advocates. The magazine is widely known in the HIV+ community and distributed for free. To get a sense of their outreach, in their media kit for 2010, the magazine publisher guarantees a minimum circulation of 100,000 copies, with 75,000 copies distributed to more than 1,900 community-based organizations and 700 Walgreens pharmacies across the US, 7,000 copies distributed at more than 200 venues, 5,000 copies distributed at HIV/AIDS conferences and events, 10,000 copies sent to individual subscribers, 1,500 copies delivered to members of the American Academy of HIV Medicine, and 1,500 copies delivered to media, HIV advocates and pharmaceutical representatives.

The aim of the drug guides is to present information about HIV drugs in a form that is easy to decipher and comparable across drugs. It is meant as a guide for patients who are just starting therapy, as well as those who have been on treatment for a long time, helping patients to discuss their treatment options with their doctors and decide whether or not an alternative treatment regimen might be more suitable. From 1997 until 2007, the magazines and the annual drug guides were only available in print. However, starting in 2007, the drug guides have also been available on the magazine's website, `positivelyaware.com`.

The drug guides offer rich information on HIV drug quality. Measures include the number of side effects observed in clinical trials, type(s) of side effects, severity of side effects, food restrictions for each drug, dosage frequency, drug interactions, and the drug's annual cost.[21] Most importantly for this study, the drug guides include reviews for each drug from both an HIV physician and a community activist (see Figure 1 for a sample page from the 2008 drug guide for AZT). While the reviewers might be different from one year to the other, only one physician and one activist review all drugs at any given point in time. To our knowledge, *Positively Aware* provides the only source of expert reviews for all HIV drugs available on the market at a given point in time.[22] Figure 2 shows selected comments from readers about the annual drug guide. These serve as anecdotal evidence that the drug guides are used by patients and doctors to gain knowledge about drugs, and helps patients discuss and choose their treatment options along with their doctor.

[21] A list of all variables constructed using information from the magazines, along with their definitions, is presented in Appendix A.

[22] An online search of HIV drug guides returns a host of resources available for people who want information on HIV drugs. However, none of them publishes expert reviews on all FDA-approved HIV drugs on the market in our time period of analysis. The only source of user reviews for HIV drugs is `drugs.com` but they are only available after our period of analysis.

Table 2 provides summary statistics for drug characteristics from *Positively Aware*. In total, we have data on 27 different drugs produced by 9 unique manufacturing firms that were on the market at some point during the period between 1997 and 2008.[23] In 1997, there were only 9 drugs to choose from, while in the last period of analysis, patients could choose from 25 different drugs. On average, drugs have 13 side effects reported in clinical trials and have molecular interactions with 14 other drugs. The average pill burden for a drug is roughly 2 tablets, taken twice a day.[24]

2.2 Coding Textual Expert Reviews

Typically, in the existing literature on the impact of expert or customer reviews on product demand, the 'expert' review variable is binary (Good or Bad) or categorical (for example, number of stars). As can be seen in Figure 1, our expert reviews are not numerical ratings, but written text. The analysis of text is problematic and open to subjective interpretation. Thus, an important question for us is how to code reviews for subsequent analysis. One of the ways some authors have gotten around this problem is to use the length of the text as a proxy for whether the review is positive or negative, with longer text signifying a "mixed" review (Chevalier and Mayzlin, 2006). However, the reviews for most drugs in the *Positively Aware* drug guides are similar in length and gauging the quality of a review from its length would produce a very noisy measure of the doctors' and activists' valuation of the drug. For some drugs a negative review by the activist is long, as he or she narrates a personal experience, or the experiences of friends, while for other drugs a positive review by a doctor or activist may be long, including for example descriptions of specific experiences when a particular drug helped to save a patient's life. Another option would be to use text analysis software to automate the coding of the reviews. Unfortunately, text analysis software is imperfect and cannot accurately capture the true flavor of the review, especially when the text may be using euphemisms, analogies or subtle sarcasm to convey the message.

To circumvent these problems, we assign a ranking to the reviews manually by developing a numerical scale and by reading each review and assigning a number to it. We use an ordinal rating of 1, 2 or 3 to classify each drug. A rating of 1 signifies a *negative* review and a rating of 3 a *positive* review. A rating of 2 means we cannot assign a 1 or a 3, which thus means that a review is mixed.[25] In what follows, when we mention the doctor's or activist's review,

[23] A detailed description of all the drugs, with information on when each drug entered (and exited) the market, is presented in Table A1 of Appendix A.

[24] The Department of Health and Human Services (DHHS) maintains a list of drugs with 'preferred regimen' status, that is, drugs that have been approved as first-line therapy by the DHHS. On average, 7 out of the 27 drugs on the market were given the preferred status.

[25] To verify that our results are not being driven by the particular way in which the reviews were coded, we

we in fact refer to our numerical interpretation of them. We provide the details of the criteria we followed to construct these numerical variables in Appendix A.

We report descriptive statistics of our expert review variables in Table 2. We observe that the average rating given by doctors is higher than that given by activists (2.02 versus 1.89) and the difference is statistically significant at the 10% level. This suggests that, on average, activists are more critical. This result is reinforced when we compare the fraction of 1's, 2's and 3's given by the two sets of experts, as shown in Figure 3. While an activist gives the lowest rating 36% of the time, the doctor rates a drug 1 only 26.7% of the time. On the other hand, a drug gets the highest rating by a doctor 27.2% of the time, while the activist rates a drug positively 24.7% of the time. In Table 3, we show a cross-tabulation of doctors' and activists' rating to demonstrate how many times the doctor and activist disagree. When the doctor gives the highest rating to a drug, the activist disagrees and gives a lower rating on 30 out of the 63 high doctor rating drugs (48% of time). On the other hand, when the activist gives the highest rating, the doctor disagrees and gives a lower rating 33% of time. While differences in reviews for the same drug could be random or reflect measurement error, as we argue below, they could also reflect how experts weight different drug characteristics when generating reviews. This has implications for how consumers incorporate reviews into their decisions.

3 Combination-Level Data: Preliminary Analysis

HIV drugs are rarely consumed individually and are instead consumed in bundles. Bundles of HIV drugs are sometimes called *cocktails*, *combination therapy*, *combos* or simply *treatments*. At a given point in time, a large majority of HIV patients combine two drugs or more in order to build a regimen that is effective in fighting HIV. Figure 4 shows the distribution of the number of drugs in the combinations. Conditional on taking at least one drug, around 35% of HIV+ individuals take 3 drugs at the same time, while 25% are following monotherapy, i.e., only taking one drug at the time of a visit.

A challenge for our subsequent analysis is that each drug is reviewed individually. In this section, we describe how we construct a dataset for the analysis of demand for combinations, which requires that we aggregate expert reviews for individual drugs into combination-level reviews. Then, we conduct a preliminary data analysis relating expert reviews to the demand for HIV drug combinations. The objective is twofold. First, we present stylized facts. Second,

also employed two undergraduate students at Johns Hopkins University to separately recode the magazine reviews. Results of the paper are robust to differences in coding, and the robustness checks are available upon request from the corresponding author.

we argue that there is exogeneous variation in our data that we can exploit to estimate the causal effect of reviews on demand. We end this section with a brief discussion of alternative interpretations of the observed relationships between drug reviews and drug demand.

3.1 Combination Data Variable Construction

For our combination-level analysis, we construct the choice set, reviews for combos, combo-level objective qualities and combo-level market shares.

Constructing the Choice Set. To study bundling, we return to the individual-level data from MACS and construct a dataset of combination choices. We restrict our attention to individuals who are taking 5 or fewer drugs during one visit.[26] This leaves us with a total of 1,248 unique drug combinations over the entire sample period. At any given visit, the number of unique drug combinations does not exceed 96.[27] A large number of these combinations, however, are taken by a small number of individuals and can be thought of as experimental combos. Therefore, in order to reduce the size of the choice set so that it is manageable from a computational perspective, as well as to be able to construct objective quality measures and market shares for every combination in our choice set, we define a 'fringe' category, in which we bunch together all combinations that are taken by fewer than 25 people in a visit.[28] That leaves us with 79 unique combinations in total across all years in our sample plus the 'fringe' category and the outside option of taking no HIV drug. Note, however, that the choice set is evolving over time. The number of combos over time (excluding the outside option) is illustrated in Figure 5.[29] We see that patients have a minimum of 21 drug combinations to choose from for the first year of our sample (early 1997), and a maximum of 58 drug combinations in late 2004.

Constructing Reviews for Drug Combinations. The doctor and activist reviews are only available for each drug, not all possible drug combinations. Therefore, in order to construct expert reviews for different drug combinations, we average over the reviews of each drug component of the combination.[30] Table 4 presents summary statistics for the

[26]Patients who are taking more than 5 drugs simultaneously are also those who are extremely sick and are probably taking multiple drugs to find one that can decrease their viral load. Since this is not how patients, on average, make medication choices, we exclude these people from our sample. By doing so, we lose less than 2% of observations.

[27]Table E2 shows the number of unique drug combinations per visit for each visit in our sample.

[28]A combination can belong to the fringe category in some visits, but not others.

[29]Over the span of 10 years, different drug combinations fall in and out of favor, especially when new drugs are introduced on the market. The total number of unique alternatives we observe is 81, but not all of these alternatives are encountered in any given time period.

[30]For example, if AZT has a rating of 2 and 3TC has a rating of 1, then the combination AZT-3TC will

combo level variables. Panel (a) shows that the average doctor's rating for a combination is 2.18, while the average activist's rating is 2.07.[31] Consistent with our previous results, doctor's reviews are significantly higher. Using the average of the individual drug reviews as our measure of combo-level reviews may overlook factors that consumers consider, such the minimum or maximum review or the variance. To explore these possibilities, after we present our main results, we assess robustness to alternative ways of aggregating individual drug reviews for combinations. Our main findings are robust to these alternatives.

Objective Qualities: Effectiveness and Side Effects. A key advantage of the MACS dataset is that it allows us to construct objective drug combination quality measures that are crucial for our demand estimation. In particular, we follow Papageorge (2016) and construct two objective quality measures for each treatment at each point in time.[32] The first measure aims to quantify treatment effectiveness at improving underlying health (as measured by CD4 count levels). The second provides a measure of the treatment side effects. We allow these measures to change each period over the lifecycle of a treatment to capture possible differences over time in treatment quality that arise, for example, if HIV mutates.[33]

The way we construct these objective quality measures for the different combinations is as follows. For each combination c, we run a probit regression on demographic characteristics to predict c's probability of non-decreasing CD4 count and probability of no ailment for the entire sample of HIV positive men. To obtain treatment-level objective quality measures, we average over all individuals. As we discuss later, a drug's quality measures might vary over time and so we allow for this by letting the probit coefficients to vary over time. Formally, to construct combo c's measure of effectiveness, we first fit a probit model of the likelihood that a patient will experience an increase in his CD4 count in period $t+1$ when taking combo c at time t, conditional on the patient's characteristics. Letting CD4_{nt} be individual n's CD4 count at time t, we estimate the model

$$\Pr_{ct}(\text{CD4}_{nt+1} \geq \text{CD4}_{nt} | X_{nt}) = \Phi\left(X'_{nt}\beta_{ct}^{\text{CD4}}\right) \tag{1}$$

on the sample of individuals who take combo c at time t, where X_{nt} is a vector of demographic controls including patient n's age, race, education level and work status as well as n's CD4 count at t, and $\Phi(\cdot)$ is the standard normal cdf. We fit the probit for each combo separately

have a rating of 1.5

[31] In Liu et al. (2014), who also study HIV drugs, promotions are studied at the individual drug level even though drugs are prescribed in combinations with others.

[32] From now on we use the terms "treatment" and "combination" interchangeably even though some consumers are observed taking a single drug.

[33] Our results are robust to the use of constant quality measures over time or rolling averages and the results of this robustness check are available upon request from the corresponding author.

(so that all coefficients can vary for each combo), and obtain the predicted probability of non-decreasing CD4 count for all individuals in each visit. In order to get the combo-level predicted probabilities, we average the predicted probabilities over all n at time t. The aim with this procedure is to compute an average treatment effect, which consumers use when choosing a treatment.

Similarly, our measure of combo c's side effects is calculated as the average likelihood that combo c produces no ailments.[34] Let noail_{nt} be a dummy variable that takes the value 1 if patient n experiences no ailments at time t and 0 otherwise. We fit the model

$$\Pr_{ct}(\text{noail}_{nt+1} = 1|X_{nt}) = \Phi(X'_{nt}\beta_{ct}^{\text{noail}}), \tag{2}$$

on the sample of individuals who take combo c at time t and where X_{nt} is the same vector of covariates as above and $\Phi(\cdot)$ is the standard normal cdf. As before, we fit a probit model for each combo and obtain the predicted probability of no ailment for each individual in each visit. In order to get the combo-level predicted probabilities, we then average the predicted probabilities over all n at time t.[35] Table 4, Panel (b) presents the summary statistics for the constructed objective quality measures. The average probability of non-decreasing CD4 count for drug combinations is 57%, while the average probability of no ailment in the period after taking the combination is 60%.[36]

Constructing treatment quality measures using individual-level data stands in contrast to other demand estimation contexts, where product characteristics (e.g., car size or horsepower) are directly observed in the data. A potential concern is that our constructed measures are biased due to patients selecting onto treatments. For example, we might classify a treatment as not very effective if sicker patients choose this treatment, even if the treatment is very effective. In particular, in the context of HIV drugs, a key source of variation in treatment choice is the individual's underlying immune system health. The latter is often not observed in the data and cannot be controlled for, thus a key source of selection bias. In our case, however, we do observe the patients' CD4 counts — a continuous and objective measure of the individual's immune system health — and we control for it when construct-

[34] We define an individual as being free of ailments if he reports no nausea, headache, fever, diarrhea, or drenching sweats in a period.

[35] An alternative approach would use probit coefficients to predict treatment effects for each set of consumer characteristics. We do not follow this approach since the aim is to capture that consumers likely know how drugs work in general, but not necessarily how they work for each set of characteristics, many of which are not observable. However, we note that reduced-form estimates remain unchanged if we allow for consumer-specific treatment effects.

[36] Since the 'fringe' category is composed of different combinations within and across different time periods, each of which have their own objective quality value, we average over different combos within the same time period t to obtain one value per time period for the objective quality measures for 'fringe'.

ing our objective quality measures. Controlling for CD4 count and other consumer-level characteristics helps us eliminate potential selection bias.[37] A second concern might be that we are introducing measurement error into the objective qualities. To the extent that this is a problem, it will cause attenuation bias and thus our estimates of the effects of drug quality on demand are conservative. We return to the discussion of the consequences of using constructed treatment characteristics in Section 4.2, when we discuss our identification strategy.

Combination Market Shares. As mentioned before, the data in the MACS dataset are collected twice a year. Thus, we can construct market shares for each visit, which span six month periods (one for April–September and the other for October–March) in each year.[38][39]

Table 4, Panel (c) provides some summary statistics of combo-level market shares. The average market share of the outside option (taking no drug) is 19%, while the market share of the 'Fringe' group is, on average, 32%. The average market share for combos other than 'fringe' and the outside option is 1%, with a maximum market share of 18%. Figure 6 shows how the market share of the outside option evolves over the time frame of our analysis. The market share for the outside option picks up in October 1999, reaching a peak in April 2003, going down for the next few visits, and then finally reaching a maximum in October 2007. In April 2008, the market share of the outside option falls drastically, from 27% to around 15%. This is because in April 2008, the drug Atripla was introduced on the market, which had a market share of 18% at the time of introduction, suggesting that a large proportion of patients who were off drugs switched to Atripla after its introduction.[40]

3.2 Preliminary Combination-Level Analysis

Having constructed combo-level reviews, objective quality measures and market shares, we now establish key patterns emerging from our data. First, we study objective qualities and

[37] In additional results, we also control for consumer-level fixed effects in constructing treatment quality measures and find that main results do not change.

[38] We construct market shares in the following way. Let C_{nct} be a dummy variable that takes value 1 if patient n responded as having taken combination c at visit t and 0 otherwise. Then, the market share for combination c at time period t is given by:

$$s_{ct} = \frac{\sum_{n=1}^{N_t} C_{nct}}{N_t},$$

where N_t is the total number of HIV+ individuals at visit t.

[39] In the year 2007, for example, we have data on two visits: April 2007 - September 2007, and October 2007 - March 2008.

[40] Other significant changes in the share of the outside option can also be linked to years when new drugs were introduced on the market.

demand to see if individuals prefer higher-quality drugs. Second, we explore the relationship between combo-level reviews and objective qualities. Third, we relate reviews and demand before and after we control for objective qualities to see if reviews have predictive power even after we control for observable drug quality levels. A caveat before we proceed. The results in this Section are useful to establish some of the patterns in the data, but note that what follows is not a proper investigation of demand since it ignores cross-product effects and the potential endogeneity of the reviews. We address formally these two problems in Section 4.

To conduct our preliminary combination-level analysis, we run logit regressions to estimate the effect of objective qualities on the probability of taking that combo. Table 5 shows how objective qualities from the MACS dataset and from the *Positively Aware* drug guide relate to combo demand.[41] Intuitively, Columns (1) through (3) show that if the probability of no ailment from a combo increases (i.e., the side effects from taking that combo go down), the demand for that combo increases. The result is statistically significant even after controlling for the probability of non-decreasing CD4 count. Similarly, if the probability of non-decreasing CD4 count increases (i.e., an increase in the combo effectiveness), then the combo demand is also larger, although the estimated effect is not statistically significant. In columns (4) through (6) we also control for the objective qualities included in the annual guides. As expected, as the number of reported side effects for a combo increases, or if the number of food restrictions for a combo increases, the demand for that combo goes down. Lastly, if the number of reported interactions with other drugs increases, the demand for the drug regimen goes down, although the results are not statistically significant. Therefore, all these results show that people, on average, prefer better quality drugs.

Next, we show how objective qualities from MACS relate to expert reviews. In Table 6 we report the estimates of a regression of reviews on objective qualities by OLS. We find that both the doctor and the activist give a higher rating to combos that have a high probability of non-decreasing CD4 count. The activist also gives higher ratings to drugs with fewer side effects, but there is no significant association between the doctor's review and the probability of no ailment of a combo. Therefore, we find that activist's reviews are higher for combos that are more effective and have lower side effects, but the doctor's reviews are higher only for effective combos, regardless of side effects.

Table 7 presents results from logit regressions showing how the combo demand is related to reviews and combo qualities. In column (2) we see that a doctor's review positively predicts combo demand when we control for objective quality measures, although the estimate is not statistically significant. Columns (3) and (4) show that the activist's review

[41] We construct combo-level qualities using the *Positively Aware* data by averaging across all drugs in a combo.

also positively predicts combo demand, even after controlling for objective quality measures. However, in columns (5) and (6), when we include both experts' reviews together, we see a reversal of sign for the doctor: that is, the doctor's review now negatively predicts combo demand (though the coefficient is not significant when we control for combo qualities). On the other hand, a higher activist's review positively predicts combo-level market share, even after controlling for measures of combo quality and the doctor's review.[42]

We also consider how large changes in reviews for a particular combo over time affect the demand for the combination. Large changes from positive to negative or vice versa may be different from more incremental changes as they might generate more "buzz" via other information sources that disproportionately affects demand. To identify large changes within drug combos we calculate, for each combination, the deviation from the average doctor or activist review for that combo. We consider two measures of large changes: if the deviation from the mean is greater than the median, and if the deviation from the mean is greater than the 75^{th} percentile. Table 8 shows that for both definitions of large changes in reviews within drug combos, combos that experience large changes in expert reviews have larger changes in demand. In particular, column (4) shows that for a combo for which the change in the doctor's review is in the top 25^{th} percentile, an increase in the doctor's review significantly decreases demand. If the change in the activist's review for a combo is in the top 25^{th} percentile, an increase in the activist's review for that combo increases demand more than the increase in demand for combos for which the change in the activist's review is not large.

We end this section by documenting how reviews evolve over time. In our data, drugs are reviewed every year by the two experts and reviews might differ not only across experts but also over time. In Panel (a) of Figure 7, we show how doctors' and activists' reviews evolve over the combo's lifecycle. We find that combo reviews decline as the drug combination ages. This could merely reflect that the objective characteristics of a combo decrease over time as well. Panel (b) shows how the probability of no ailment and probability of non-decreasing CD4 count of combos change over the combination's lifecycle. As the combo ages, the probability of no ailment increases, indicating that side effects decrease as the combination becomes older, while the probability of non-decreasing CD4 count decreases for older combinations, suggesting that old combinations are not as effective as new ones. These patterns are consistent with drugs losing effectiveness as the virus mutates and with patients gaining tolerance to side effects as doctors and patients gain experience with drugs

[42]We also construct objective qualities from MACS at the drug level and relate reviews and drug consumption before and after controlling for these objective qualities. Empirical patterns remain qualitatively similar, which helps alleviate the concern that results might be affected by aggregation. However, since that is not how drugs are actually consumed, we do not report these results as part of our reduced form analysis. These results are available upon request from the authors.

and dosage.

We argue next that the changes in the combo characteristics only partially explain the downward trend in combo reviews. To that end, in Panel (c), we plot residualized ratings after controlling for objective quality measures of the combo, and find that even when we control for the evolution of a combination's quality, reviews are still decreasing over time. One possible reason for this "deflation" could be that reviews are relative to other available drugs in the market. If so, as technology improves and new drugs enter the market, reviewers may lower their reviews for older drugs. What once was regarded as a stellar drug may now be superseded by a newer, better drug. If this is the case, we would expect variation in how much reviews change for a given drug depending on the quality of rival drugs, conditional on a drug's own characteristics (which might also change over time). We test this hypothesis in Section 4.2, and find that higher rival drug qualities lead to lower drug reviews. This finding motivates our identification strategy. The idea is to instrument for reviews using the qualities of the set of rivals at any point in time, where the set of rivals shifts over time due to the emergence of new drugs.

3.3 Expert Reviews and Demand: Alternative Explanations

The previous analysis provides preliminary evidence that expert reviews published in *Positively Aware* predict market shares for HIV drugs. However, there are several alternative explanations which would also explain the correlation between combo reviews and combo demand that we find in the data. One possibility is that reviews do not drive demand but simply reflect a drug's observed qualities which in turn is the demand driver. However, in the previous section, we showed that reviews continue to predict market shares even after we control for various measures of combo quality. Still, it is possible that reviews are not exogenous. One concern is simultaneity. It may be the case the reviews simply reflect demand patterns. This relationship could be contemporaneous – e.g., drugs with high market shares today may receive better reviews today – or not. In our context, there may be inertia or switching costs associated with drug choices. Hence, a high demand for a drug in the past may drive both a high review and high demand for the drug today. To mitigate this problem, in our empirical analysis we control for whether the patient has taken the same drug in the previous period. Another concern is unobserved drug heterogeneity. Magazine reviews may reflect drug qualities that are not observable to the econometrician but are observable to patients and doctors who make treatment decisions and therefore affect demand. We defer the formal treatment of the endogeneity issue to Section 4.

A second possibility is that the impact of reviews on demand for HIV drugs is indeed

causal, but that it is not due to patients reading *Positively Aware*. For example, *Positively Aware* magazines are not the only source of information about drugs available to patients. Other magazines could provide similar information and affect demand. However, to our knowledge, *Positively Aware* drug guides are the only source of information in which patients can read reviews about all FDA approved drugs from a doctor and HIV activist in a systematic way.

The third potential story is that the true demand driver is collective, evolving knowledge about drug quality and the reviews are just reflecting it. We provide some suggestive evidence that this is not the case. We do so by exploiting the timing of the reviews relative to when we observe drug choices. In particular, given that the annual guide is published in January/February and data on drug choices are collected both in April and October, we consider three distinct market share windows for our analysis. Relative to reviews published in Jan/Feb of year t, we can construct market shares realized *before* the magazine is published (i.e., market shares for the window April-September in period $t-1$), market shares for the window that overlaps with the period *during* which magazine is published (i.e., market shares for the window October-March in period $t-1$), and market shares realized *after* the magazine has been published (i.e., market share for the windows April-September and October-March in period t). The timeline of events is illustrated in Figure 8.

If the reviews solely capture evolving social knowledge about drugs, by construction they would only capture knowledge from the 12 months prior to publishing. Thus, we could falsify the social knowledge hypothesis if reviews at period t have no effect on market shares for the *before* and *during* windows at $t-1$. In Table 9, we find that reviews at t for the *before* and *during* periods have no significant effect. Moreover, when we run the regressions of market shares for the *after* window at t and $t+1$ (including both April and October visits), we find that reviews published in period t do have a significant effect for the activist. We also find that the activist's and doctor's review are jointly significant for this period. We interpret this as suggestive evidence that reviews from *Positively Aware* (rather than evolving social knowledge) drive demand for HIV drugs.

4 Econometric Model and Identification

While the results in Section 3.2 relate market shares and drug qualities and reviews, these are not estimates of the demand for combos since the demand for a given drug not only depends on own characteristics and reviews but also on rivals' characteristics and reviews. If we leave the problem unrestricted, this would require us to estimate a very large number of parameters

(i.e., own- and cross-product effects) relative to the sample size. To address this problem, in this section, we specify a structural econometric model of demand for HIV combos that we take to the data. The purpose is twofold. First, the estimates of the coefficients of the structural model will allow us to obtain own- and cross-review elasticities in a parsimonious way. Second, the model makes explicit the identification issue we need to overcome and will help in understanding the logic behind our identification argument.

4.1 Model Specification

We study combination choice using a discrete choice demand model at the combo level. Let $\mathcal{J}_t$ denote the choice set at time period t. To explain choices, we allow the utility of individual i from consuming combination j at time t to depend on the drug characteristics — both observed and unobserved — as well as his demographic characteristics, health status, and unobserved taste shocks.[43] Let x_{jt} be a K-dimensional vector of observed product characteristics — including the doctor's and activist's reviews, along with objective qualities of a combo — at time t and let ξ_{jt} denote the unobserved product characteristic (or combo j's demand shock at time t).[44] [45] Also, let z_{it} be an R-dimensional vector of individual i's characteristics at time t, including age, education (dummies for whether the individual is a high school or college graduate), work status (dummy for full-time work), race (dummy for black), AIDS status, and whether or not the individual was taking the same combination in the last period. We can then write the utility i gets from consuming alternative j at time t as

$$u_{ijt} = \sum_k x_{jtk}\tilde{\beta}_{ik} + \xi_{jt} + \epsilon_{ijt}, \tag{3}$$

with

$$\tilde{\beta}_{ik} = \bar{\beta}_k + \sum_r z_{irt}\beta_{kr}, \tag{4}$$

[43]The model we specify here is used to estimate the impact of reviews on demand. Following literature on advertising, which uses a similar framework, our model treats reviews as an additional product characteristic that drives demand by affecting the utility of a given product. Alternatively, a fully specified structural demand model could treat individuals as not having preferences over reviews, but as relying on reviews for additional information about drug characteristics over which they do have preferences, but do not fully observe. If this is the case, in our current setup, we are recovering a reduced-form relationship between reviews and demand.

[44]Note that we treat each combination j at time t as a separate product, so that AZT–3TC in 1997 is a different product than AZT–3TC in 1998.

[45]While the list prices for these drugs differ considerably, health insurance covers the full cost of the treatment and several other government assistance programs are available to uninsured patients. As part of the MACS data we observe out-of-pocket expenditures on HIV drugs and they tend to be negligible. We have included this variable as a control and it is always not statistically and economically significant. Hence, we exclude this variable from the analysis.

where $\tilde{\beta}_{ik}$ is individual $i's$ taste for product characteristic k, which depends on his observed individual-level characteristics z_i, and ϵ_{ijt} represents a shock to preferences which we assume is distributed Type-I extreme value and independent across choices, individuals, and time. Letting

$$\delta_{jt} = \sum_k x_{jtk}\bar{\beta}_k + \xi_{jt} \tag{5}$$

denote the mean utility level, we can rewrite the utility as

$$u_{ijt} = \delta_{jt} + \sum_{k,r} x_{jtk} z_{ir} \beta_{kr} + \epsilon_{ijt}. \tag{6}$$

Market-level aggregate consumer behavior is obtained by aggregating the choices implied by the individual utility maximization over the population distribution of individual characteristics. Let $\mathcal{P}(\mathbf{w})$ denote the distribution of $\mathbf{w}$ in the population, where $\mathbf{w} = (\mathbf{z}, \boldsymbol{\epsilon})$ is the vector of observed and unobserved individual characteristics. Then, conditional on product characteristics, the fraction of individuals who choose combination j at time t is given by integrating over the set of individual characteristics that imply a preference for combo j at time t:

$$s_{jt}(\boldsymbol{\delta}, \boldsymbol{\beta}; \mathbf{x}, \mathcal{P}(\mathbf{w})) = \int_{A_{jt}(\boldsymbol{\delta}, \boldsymbol{\beta}; \mathbf{x})} \mathcal{P}(\mathbf{w}) d\mathbf{w}. \tag{7}$$

where

$$A_{jt}(\boldsymbol{\delta}, \boldsymbol{\beta}; \mathbf{x}) = \{\mathbf{w} : \max_{p \in \emptyset \cup \mathcal{J}_t} [u_{ipt}(\mathbf{w}; \boldsymbol{\delta}, \boldsymbol{\beta}, \mathbf{x})] = u_{ijt}\}. \tag{8}$$

We estimate the demand model by GMM. Details about the estimation are presented in Appendix B.

4.2 Identification

We know from Section 3.2 that doctors' and activists' reviews reflect observed combo characteristics. An endogeneity problem might arise if reviews for treatment j also reflect j's unobserved combo quality, ξ_j (that is, a combo characteristic that is observed by the patients and reviewers but not by the econometrician). This problem is analogous to the endogeneity of price that arises in traditional demand estimation (see, e.g., Berry (1994)).

In order to establish a causal relationship between reviews and market share, we leverage the fact that the choice set is evolving over time, with new drugs entering the market every period. If combo entry is exogenous and reviews are relative, then the entry of new combos provides exogenous variation in reviews over a combo's lifecycle.[46] Given that we have two

[46]Note that in our setting, and in contrast to the setting in BLP, the characteristics of the incumbent

potentially endogenous regressors – namely, the doctor's and activist's reviews – we need two instruments. Specifically, we use the averages of the two observable qualities of rival combos on the market as instruments for the reviews. The intuition is that the quality of rival drugs should change the reviewer's relative valuation of an incumbent drug's quality, and will hence affect the review for that drug. In an analog way to a first-stage regression, Table 10 shows how the doctor's and activist's reviews of a combo relate to the average quality of rival combos on the market. As expected, results show that (i) an increase in the objective qualities of a combo is positively correlated with its reviews; and, more importantly, (ii) an improvement in the average probability of no ailment or the average probability of non-decreasing CD4 count of rival combinations leads to a decrease in the reviews for the combination. A joint test of the rivals' objective quality measures show that both the average probability of non-decreasing CD4 count of rival combinations and the average probability of no ailment of rival combinations significantly affect doctors' and activists' ratings for a combination.

Validity of these instruments relies on the identifying assumption that rival treatments affect reviews by experts but are uncorrelated with incumbent treatment unobserved characteristics. That is, we require the observable characteristics of the entrants to be orthogonal to the unobserved characteristics of the incumbent drugs. Note that the logic behind our instruments is similar in spirit to the one in BLP. In BLP, prices are endogenous and need to be instrumented. Prices are set in equilibrium by oligopolistic firms, and therefore prices not only depend on a given product's characteristics but also on the characteristics of its rivals. Therefore, rivals' characteristics are valid instruments under the assumption that product characteristics — other than price — are exogenous.

We acknowledge two important concerns. First, observed characteristics of new entrants could be correlated with incumbents' unobservable quality. For example, pharmaceutical companies presumably innovate and bring new drugs to the market in response to existing drugs' qualities. Arguably, incumbents' effectiveness, side effects, drug interactions, food restrictions, and so on, affect how pharmaceutical firms develop new drugs. Note, however, that this is only a concern for identification if new drugs' characteristics are correlated with unobserved incumbent characteristics that are not captured by the host of observed characteristics we include in our analyses. In other words, identification relies on the assumption that new entrants' observed characteristics are uncorrelated with incumbents' unobserved characteristics, conditional on incumbents observed characteristics. In further support of

treatments also change over time. For this variation to be exogenous, the same conditions we impose on new entrants is required. We need to assume that the observed characteristics of treatment j's rivals are independent of j's unobserved characteristic (conditional on j's observed characteristics).

this assumption, we note that the FDA approval process is often lengthy and fraught with uncertainty which introduces some randomness into the timing of drug entry, which would further mitigate concerns that new drug characteristics correlate to existing drugs' unobservable qualities. Finally, innovation itself entails uncertainty and pharmaceutical firms are not always able to fully determine the characteristics of their drugs in response to the market conditions. For example, Hamilton et al. (2018) show that many new entrants in the HIV drug market are dominated both in terms of efficiency and side effects by existing drugs, a fact that is at odds with the idea that firms have full control over the innovation process. This introduces further randomness into the entry process, which supports the idea that new entrants provide exogenous variation in reviews.[47]

A second concern with identification, one that is not shared by instruments in BLP, is that we construct the treatment characteristics, and hence the instruments, using our patient-level data as described in Section 3.1. To the extent that there is selection into treatments based on unobserved patient characteristics, this could undermine the validity of our instruments. The main worry here is that sicker patients might select into more effective drugs, introducing bias into our estimation of average treatment effects. To mitigate the effects of selection, we control for patient demographics and health status when we construct the objective drug characteristics and in our specification of the utility function. In particular, the inclusion of a continuous measure of underlying health (CD4 count) is crucial because it allows us to control for a key source of variation in health decisions, which is often omitted and thus a key source of bias.

5 Findings

This section presents our main findings. We begin with estimates from a baseline demand model, which relates demand to drug characteristics and reviews using the IV logit model presented in the previous section. Details about the estimation procedure are presented in Appendix B. Results are qualitatively similar to the reduced-form estimates we obtained previously. Higher activist reviews increase demand, whereas higher doctor reviews lower demand, even after controlling for objective treatment characteristics. Next, we estimate a set of models that explicitly distinguishes cases where the two reviews are in agreement versus disagreement. We show that higher reviews increase demand when doctors and activists

[47] A related concern, which was discussed earlier, is that reviews reflect demand shocks driven by consumer preferences for unobserved characteristics. This would induce correlation between reviews and unobserved treatment characteristics and thus undermine our identification strategy. As discussed in Section 3.3, this would suggest timing that is not consistent with the data.

agree. However, when they disagree, healthier consumers tend to follow the activist's review, while less healthy patients follow the doctor. The remainder of this section provides evidence that these patterns reflect how consumers trade off their demand for long run health and their distaste for treatment side effects.

5.1 Estimates of the Baseline Model and Robustness

We begin by estimating the parameters of the demand model without heterogenous coefficients, treating reviews from both experts as additional treatment characteristics, instrumenting for both using the averages of the rivals' objective qualities. Table 11 reports the IV logit coefficients. Column (1) shows that a higher doctor's review on its own decreases demand. A one-unit increase in the doctor's review decreases the likelihood the treatment is chosen by 1.6 percentage points.[48] This result also holds when we control for objective treatment qualities (see column (2)). Similarly, columns (3) and (4) show that a higher activist's review for a combination increases consumption. A one-unit increase in the activist's review for a combination increases demand by 2.4 percentage points.

When we include both reviews together, we find that a higher review from the doctor lowers demand, while a higher review from the activist raises demand (see columns (5) and (6)). This finding is in line with our previous reduced-form estimates. Keeping objective qualities and the activist's review fixed, an increase in the doctor's rating of one unit leads to a 3.6 percentage point decrease in demand, while an increase in the activist's rating, keeping the doctor's review fixed, raises demand by 3.8 percentage points.[49] Column (7) shows that even after controlling for FDA combo characteristics, and keeping objective qualities and the activist's review fixed, an increase in the doctor's rating of one unit leads to a 2.7 percentage point decrease in demand, while an increase in activist's rating, keeping the doctor's review fixed, raises demand by 1.6 percentage points.

How should we interpret our estimates given varying sets of controls? Using a reduced-form interpretation, the coefficients we estimate quantify the arguably causal impact of expert reviews on demand. Given concerns about omitted variable bias, we not only instrument for reviews, but also include in our demand equations a host of observables that

[48]The percentage point change in the probability of choosing a combo alternative is calculated for each combo-time dyad and then averaged across the entire sample.

[49]Note that in our IV logit specifications, once we control and instrument for activist's reviews, the coefficient on probability of non-decreasing CD4 count is negative. This negative coefficient captures how patients with different attributes (for example, those who are working full-time) may prefer combinations with fewer side effects but lower efficacy. In fact, in our demand model with individual attributes, we show that once we explicitly account for differences in patients' attributes such as race, work status etc., both objective qualities affect utility positively.

we suspect might simultaneously drive demand and reviews. We find that reviews affect demand in a statistically and economically significant way. We also find that the magnitude of these effects declines by about one third when we include observable characteristics, but remain economically meaningful and statistically significant.

Imposing a more structural interpretation of our estimates leads to a different interpretation of which controls we should include in our demand equations. If we include a control variable in our demand estimation, we are imposing the assumption that the information captured by the variable is known by the patient. For example, if consumers have information on average effectiveness and side effects of different treatments, we need to include these measures as covariates in our specification of utility. Thus, the model in column (5) is consistent with consumers not having access to information on treatment characteristics and when they make their choices. In contrast, the models in columns (6) and (7) are consistent with consumers having information about the drug qualities we control for, yet still facing uncertainty about the treatments' qualities and hence relying on the reviews. This logic is formalized in the simple model we present in Appendix D. While we cannot be sure which of the two alternatives is the right one, our preferred specification is the one in column (7). To the extent that patients do not have information about our controls, controlling for these observables would underestimate the effect of the reviews. Thus, it is reassuring that our results are robust to the inclusion of different sets of controls, which amounts to imposing different assumptions on consumer information sets.

In a similar vein, we also do not know how patients aggregate the reviews of the individual drugs that form the different treatments. We thus assess whether our results are robust to different ways of constructing and controlling for combo-level reviews. First, we generate reviews for combos by calculating the percentage of drugs that have a rating of 3 in the combination. This relaxes the implicit cardinality assumption arising from our use of averages. Summary statistics for this measure of expert ratings is presented in Table 4. Demand estimates using this definition of reviews are given in Panel (a) of Table 12. Notice that results are qualitatively similar to our original specification. As before, we find that doctor's and activist's reviews positively predict demand when including them one at a time; however, when we control for both at the same time, we find that a higher doctor's review lowers demand. In our second specification, we construct indicator variables for whether the doctor and activist's review are higher than 2, to capture the effect of 'high' expert reviews on demand. We report the results in Panel (b) of Table 12. Again, we find that when we control for both experts' reviews at the same time, a high doctor's review decreases demand while a high activist's review increases demand for the combo.

Our third alternative specification includes the average review across all drugs in a com-

bination as well as the standard deviation of reviews within each combination. The aim is to capture how patients value both the mean and the variance of individual product attributes (drug-level reviews) in the bundles they consume (see, e.g., Farquhar and Rao (1976), Bradlow and Rao (2000)).[50] Results using this specification are shown in Panel (c) of Table 12. For the doctor's review, after controlling for the average review, an increase in the standard deviation is negatively related to demand. For the activist's review, the standard deviation of the reviews has a positive but insignificant relationship to demand once we control for objective qualities of the combination. In columns (5) and (6), we again see the reversal in sign for the average doctor's review when we include both the activist and doctor's review, although our estimates are imprecisely estimated. Since we find no evidence that the standard deviation of reviews affects demand significantly once we control for the average reviews, we omit the standard deviations in our main specification.[51] More generally, and similar to the question of deciding which controls to include in demand equations, as we cannot be sure how consumers aggregate information about treatments multiple treatments, it is reassuring that our main results are robust to a variety of different specifications.

5.2 Disagreements and Demand

At face value, it seems puzzling that demand responds negatively to higher doctor's reviews. To further explore this result, we consider how consumers respond to reviews when the doctor and activist agree versus when they disagree. In fact, disagreements occur quite frequently: for 70% of combination-time dyads.[52]

To understand disagreements better, we first assess how they evolve over the age of the combination. For this exercise, we generate an indicator variable that takes the value 1 if the activist's and doctor's review differ for any one of the drugs in the combination. Panel (a) of Figure 9 depicts disagreements over drug age, and 0 otherwise. Not surprisingly, most of the disagreements between the two experts occur when the combination is 'new', i.e., the combination has only been on the market and consumed by patients for three years or less. The experts disagree 75% of the time when the combination is new, but over time, specifically, when the combination has been part of the choice set for more than 6 years,

[50] Farquhar and Rao (1976) and Bradlow and Rao (2000) describe individual choices among an assortment of multi-attributed items in which the assortment could be made from a subset of all items available to individuals. In their model, they allow a mean level of attributes for the assortment as well as the dispersion of attributes to affect utility.

[51] Additional robustness checks are presented in Appendix C.

[52] In Appendix D, we present a simple theoretical model outlining how demand should shift when experts disagree and patients choose drugs that align with their preferences.

the frequency of disagreements between the two experts declines.[53] We also consider the magnitude and direction of the disagreements. Panel (b) of Figure 9 shows the distribution of the difference between the activist and doctor's review. When the two experts disagree, we are more likely to see a higher average review for the combination from the doctor than the activist.

Next, we explore the effect of disagreements on demand by interacting the doctor's and activist's review with a dummy for disagreements, and interacting the doctor's review with a dummy for agreements.[54] The coefficient on the interaction between agreements and the doctor's review captures the relationship between reviews and demand when the experts agree, while the coefficients on the interactions between disagreement and the two expert reviews capture which expert patients follow when experts disagree. The estimates are shown in Table 13 (for comparison, column (1) reproduces column (6) of Table 11). In column (2), we see that, on average, if both experts agree and the combination gets a higher review, then demand rises. This finding means that patient demand rises when both the activist and the doctor ratings for a treatment are high. However, this increase in demand is not economically and statistically significant. On average, a one unit increase in experts' rating when both experts agree leads to an increase in the probability of taking a combination by approximately 0.66 percentage point.[55] When the experts disagree, however, a higher activist's review for a combination increases demand by 3.2 percentage points, while a higher doctor's review lowers demand by 2.8 percentage points. This set of results is intuitive. The informational content of the expert reviews, above and beyond the objective treatment qualities, is small when the experts agree. Put differently, it is when the experts disagree that the reviews are more valuable to inform demand.

These new results suggest that the negative coefficient on the doctor's review from our baseline model is driven by cases when the doctor's review is at odds with the activist's. To explore this point further, we also assess potential asymmetries in how patients respond to conflicting expert reviews. We calculate the difference between the activist's and doctor's review for each drug, generate a dummy for whether this difference is positive (the activist gives a higher review compared to the doctor) or negative (the activist's review is lower than the doctor's) and interact these dummies with the two experts' reviews. Results are shown in column (3) of Table 13. Estimates show that when reviewers disagree, there is a significant

[53] An exception is a high proportion of disagreements occurring when combo age is 11. This is driven by a set of combinations of old drugs (d4T, 3TC and Nevirapine). Removing these combinations does not affect our results.

[54] Note that when the experts agree, the activist's and doctor's reviews take the same value, so interacting the dummy for agreement with the activist review is redundant.

[55] We calculate the average marginal effect by first calculating the marginal effect for each combo-time dyad, and then averaging across all combo-time dyads.

effect on demand when the activist's review is lower than the doctor's. In particular, an increase in the doctor's (activist's) review has a negative (positive) effect on demand when the reviews differ and the activist's review is lower than the doctor's review. When the activist's review is higher than the doctor's, the effect of reviews on demand is insignificant. This result provides further nuance to our baseline estimates. The overall negative reaction to the doctor's review arises when doctors and activists disagree and, moreover, when the activist downgrades a drug that the doctor does not. It is still a puzzle why patients seem to defy the doctor's recommendation in certain situations, a point we explore in the next section.

5.3 Conflicting Reviews, Side Effects and Demand for Expertise

Having established the importance of disagreements in explaining how patients respond to expert reviews, we now turn to understanding patient responses to disagreements.

5.3.1 Disagreements and Objective Qualities

We begin by exploring the relationship between experts' ratings and objective qualities of treatments (probability of no ailment and probability of non-decreasing CD4 count) in the choice set to see if there are differences, on average, in how experts respond to these qualities when they disagree. In Table 14, we regress doctors' and activists' ratings on the objective qualities of own and rival combos for the sample of combos for which the two experts disagree. We find that when the two experts disagree, the doctor's review increases if the probability of non-decreasing CD4 count of a combo increases, while the probability of no ailment has a negative but statistically insignificant effect on the doctor's rating. On the other hand, the activist responds positively to both objective quality measures. This then means that a drug that is highly effective but has harsh side effects will be ranked high by the doctor but not necessarily by the activist.

5.3.2 Individual Characteristics and Demand for Expertise

Results until now suggest that patient responses to conflicting reviews could reflect their attempt to choose treatments with fewer side effects. Effective treatments with side effects are downgraded by the activist, but not by the doctor. Consumers with a distaste for side effects may understand this and utilize expertise accordingly. In particular, consumers may turn to the activist — a fellow patient whose review responds to side effects — when choosing treatments under uncertainty. A test for this explanation would consider the behavior of

patients who are not necessarily seeking drugs with fewer side effects, but instead aim to use the most effective treatments possible.[56] Presumably, such patients would be more likely to follow the doctor's review. In fact, using the same dataset, Papageorge (2016) shows that sicker patients are more willing to suffer side effects. The reason is that they face stronger incentives to make costly health investments and use treatments despite their drawbacks. If patient responses to reviews reflect a distaste for side effects, we might expect sicker patients to respond more positively to doctors' reviews in comparison to relatively healthy patients.

To explore this possibility, we allow parameters on reviews to depend on patient characteristics as formulated in equation (4). Results are presented in Table 15. Many results are similar to our baseline estimates. On average, individuals prefer combinations that have a higher probability of increasing CD4 count in the next period as well as those that increase the probability of suffering no ailments. The parameters on doctors' and activists' reviews are both statistically significant, showing that reviews matter for treatment choice. For a black individual with no AIDS, working full-time, high school education, average age of 47, and who has taken an experimental drug in the previous visit, a higher activist's review increases combo demand while a higher doctor's review lowers demand. [57]

Turning to individual characteristics, we find that different types of patients react differently to reviews. The most striking finding is that sicker patients — defined as those living with AIDS — respond positively to both the doctor's and activist's reviews. In other words, for patients with AIDS, we find a reversal in sign in how patients respond to the doctor's review. While healthier patients respond positively to the activist and negatively to the doctor, sicker patients respond positively to both. This finding provides strong evidence for our preferred explanation of patient responses to conflicting reviews. When doctors and activists agree, their reviews lead to increases in demand for HIV treatments. When they disagree, healthier patients use information from the reviewer who downgrades effective treatment with harsh side effects. However, sicker patients who face strong incentives to invest in their health despite harsh side effects do the opposite. They utilize expertise from the doctor, the expert reviewer who recommends treatments based on their effectiveness and largely ignores side effects.

Interacting demand responses to expertise with individual characteristics provides several more nuanced lessons about how individuals incorporate possibly conflicting expert reviews

[56] Appendix D presents a very simple theoretical model that formalizes the logic behind this falsification test.

[57] Note, though, that we are already controlling for combos' objective characteristics. Therefore, when we say that the patient's preferences align with the activist's or do not align with the doctor's, this statement is conditional on objective characteristics. In other words, patient's preferences (do not) align with the activist's (doctor's) above and beyond objective effectiveness and side effects measures.

into their decisions. We show that the coefficient on full-time work is negative and significant, meaning that full-time workers are more likely to avoid medication altogether. This is consistent with the idea that individuals may choose not to take life-saving treatments if the side-effects interfere with daily functions (see Papageorge (2016)). Moreover, full-time work predicts a relatively large increase in demand due to a high activist's versus a high doctor's review. This suggests that full-time workers are somewhat more likely to use information from the activist, which makes sense if they aim to use treatments with fewer side effects. We also find that older consumers respond positively to doctor's reviews.[58]

We also find evidence of differences by race in how consumers respond to expert reviews. In particular, our estimates suggest that black men are just as likely to follow the activist's review as are white men, but are less likely to follow the doctor. This is consistent with distrust of the medical establishment among African Americans, which has been documented in many studies (see, e.g., Alsan and Wanamaker (2017)). A similar pattern emerges for individuals without a college degree: they place more weight on the activist's review. In other words, apart from health differences in how individuals respond to different sources of information, there may also be socioeconomic gradients. One concern with the pattern we find is that it suggests that lower-educated and non-white individuals may put their long-run health at more risk compared to white men with higher education. Patients may follow the activist's review in an effort to use medical treatments that make side effects less probable. However, when they become ill, they turn to the doctor's review in an effort to recover their health. Indeed, following the activists review when in relatively good health makes most sense if patients switch gears when in poor health. If less educated or non-white individuals are less likely to switch to following the doctor's review when in poor health, they may be less likely to recover. If so, the expertise provided by the activist may be more harmful to blacks as compared to whites. If so, patient advocates (in our case, encapsulated in the activist's review) may provide information that is more helpful to more highly educated individuals at the expense of others. Future research could further explore how various information sources affect demand and health outcomes for different socioeconomic groups.

[58]The estimated positive coefficient on the dummy variable 'same combo last period - other', even though not significant, can be interpreted as capturing switching costs or, alternatively, as learning-by-doing (i.e., experience). That is, if a patient was taking a combo (other than the fringe) in the previous period, it is more likely that the patient will continue taking that same combo in the current period. On the other hand, if the patient was taking a combo from the fringe class in the previous period, it is more likely that the patient will switch out of the fringe in the current period. This could be interpreted as a cost associated with continuing experimenting with a rarely used treatment. The interactions also indicate that college-educated individuals respond positively and significantly to doctors' reviews, but not to activists' reviews. While estimating a dynamic model of drug consumption is beyond the scope of this paper, including the dummies for the consumption of the combo in the previous period serves to mitigate the potential endogeneity problem arising from persistence in the demand and the reviews reflecting that.

6 Conclusion

We have demonstrated that expert reviews affect demand in a high-stakes context: the market for HIV treatments. Much research on low-cost information and decision-making overlooks the idea that consumers often have access to multiple information sources. Exploiting rich data that includes objective drug qualities, individual-level health outcomes and multiple reviews, we show that consumer responses depend on their health along with other observable factors. We argue that these responses provide evidence that consumers demand information that is aligned to their preferences over health and side effects, which can vary depending on their current health state. According to our results, patients are not passive consumers of low-cost information sources, but actively incorporate information from different experts to make more informed decisions. This point has important implications in terms of how researchers should interpret the effect of information interventions on demand. Our results suggest that consumers endogenously choose whether they are exposed to the information and the extent of the exposure based on their priors over whether or not the information will be relevant to them or is aligned with their preferences.

In our study we use observational data and suitable econometric tools to shown a new set of patterns that, to the best of our knowledge, existing studies have not been able to show. A natural next step would be to design a randomized trial to not only corroborate our findings about diverging expert reviews, but to also extend our results. Future work could continue to investigate how individuals respond to conflicting information, for example, when reviews are side-by-side, as in our case, versus when they are not. Another example could be to investigate how consumers incorporate information into their choices when information acquisition from additional, possibly conflicting, information sources is costly. Moreover, future research could further explore heterogeneity in how individuals respond to various information sources when making decisions under uncertainty. An experimental setting could be used to vary not only the source of the information, but also its content. Moreover, though we have emphasized health differences in responses to doctors' versus activists' reviews, future work could focus on socioeconomic differences in how individuals respond to conflicting information sources. Such work could allow for an assessment of how such differences in the incorporation of information contribute to well-established health disparities.

References

Ackerberg, Daniel A, "Empirically Distinguishing Informative and Prestige Effects of Advertising," *The Rand Journal of Economics*, 2001, pp. 316–333.

Alsan, Marcella and Marianne Wanamaker, "Tuskegee and the Health of Black Men," *Quarterly Journal of Economics*, 2017, *133* (1), 407–455.

Anderson, Michael and Jeremy Magruder, "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database," *The Economic Journal*, 2012, *122* (563), 957–989.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," forthcoming at *American Economic Review* 2017.

Bandiera, Oriana and Imran Rasul, "Social Networks and Technology Adoption in Northern Mozambique," *The Economic Journal*, 2006, *116* (514), 869–902.

Banerjee, Abhijit V, "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 1992, pp. 797–817.

Barber, Brad M and Terrance Odean, "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *Review of Financial Studies*, 2008, *21* (2), 785–818.

Benabou, Roland and Guy Laroque, "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *The Quarterly Journal of Economics*, 1992, *107* (3), 921–958.

Berry, Steven, "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262.

_ , James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, July 1995, *63* (4), 841–890.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman, "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment," *The Quarterly Journal of Economics*, 2010, *125* (1), 263–306.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 1992, pp. 992–1026.

Bradlow, Eric T and Vithala R Rao, "A Hierarchical Bayes Model for Assortment Choice," *Journal of Marketing Research*, 2000, *37* (2), 259–268.

Brown, Alexander L, Colin F Camerer, and Dan Lovallo, "To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office," *American Economic Journal: Microeconomics*, 2012, *4* (2), 1–26.

_ , _ , and _ , "Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The Case of Withheld Movie Critic Reviews," *Management Science*, 2013, *59* (3), 733–747.

Cabral, Luis and Ali Hortacsu, "The Dynamics of Seller Reputation: Evidence from ebay," *The Journal of Industrial Economics*, 2010, *58* (1), 54–78.

Cantor, Richard and Frank Packer, "Multiple Ratings and Credit Standards: Differences of Opinion in the Credit Rating Industry," *Federal Reserve Bank of New York Staff Reports*, 1996.

__ **and** __ , "Differences of Opinion and Selection Bias in the Credit Rating Industry," *Journal of Banking & Finance*, 1997, *21* (10), 1395–1417.

Chernew, Michael, Gautam Gowrisankaran, and Dennis P Scanlon, "Learning and the Value of Information: Evidence from Health Plan Report Cards," *Journal of Econometrics*, 2008, *144* (1), 156–174.

Chevalier, Judith A and Dina Mayzlin, "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 2006, *43* (3), 345–354.

Ching, Andrew and Masakazu Ishihara, "The Effects of Detailing on Prescribing Decisions under Quality Uncertainty," *Quantitative Marketing and Economics*, 2010, *8* (2), 123–165.

__ **and** __ , "Measuring the Informative and Persuasive Roles of Detailing on Prescribing Decisions," *Management Science*, 2012, *58* (7), 1374–1387.

__ **, Robert Clark, Ignatius Horstmann, and Hyunwoo Lim**, "The Effects of Publicity on Demand: The Case of Anti-Cholesterol Drugs," *Marketing Science*, 2015, *35* (1), 158–181.

Coffman, Lucas C, Clayton R Featherstone, and Judd B Kessler, "Can Social Information Affect What Job You Choose and Keep?," *American Economic Journal: Applied Economics*, 2017, *9* (1), 96–117.

Conley, Timothy G and Christopher R Udry, "Learning About a New Technology: Pineapple in Ghana," *American Economic Review*, 2010, pp. 35–69.

Coscelli, Andrea and Matthew Shum, "An Empirical Model of Learning and Patient Spillovers in New Drug Entry," *Journal of Econometrics*, 2004, *122* (2), 213–246.

Crawford, Gregory S and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

Dafny, Leemore and David Dranove, "Do Report Cards Tell Consumers Anything they Don't Already Know? The Case of Medicare HMOs," *The Rand Journal of Economics*, 2008, *39* (3), 790–821.

Dickstein, Michael J, "Efficient Provision of Experience Goods: Evidence from Antidepressant Choice," Working Paper 2014.

Dranove, David and Andrew Sfekas, "Start Spreading the News: A Structural Estimate of the Effects of New York Hospital Report Cards," *Journal of Health Economics*, 2008, *27* (5), 1201–1207.

__ **and Ginger Zhe Jin**, "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 2010, *48* (4), 935–963.

__ **, Daniel Kessler, Mark McClellan, and Mark Satterthwaite**, "Is More Information Better? The Effects of "Report Cards" on Health Care Providers," *Journal of Political Economy*, 2003, *111* (3), 555–588.

Duflo, Esther and Emmanuel Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *The Quarterly Journal of Economics*, 2003, *118* (3), 815–842.

Eliaz, Kfir and Andrew Schotter, "Paying for Confidence: An Experimental Study of the Demand for Non-instrumental Information," *Games and Economic Behavior*, 2010, *70* (2), 304–324.

Farquhar, Peter H and Vithala R Rao, "A Balance Model for Evaluating Subsets of Multiattributed Items," *Management Science*, 1976, *22* (5), 528–539.

Ferraz, Claudio and Frederico Finan, "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," *The Quarterly Journal of Economics*, 2008, *123* (2), 703–745.

Figlio, David N and Maurice E Lucas, "What's in a Grade? School Report Cards and the Housing Market," *American Economic Review*, 2004, *94* (3), 591–604.

Fong, Christina M and Felix Oberholzer-Gee, "Truth in Giving: Experimental Evidence on the Welfare Effects of Informed Giving to the Poor," *Journal of Public Economics*, 2011, *95* (5), 436–444.

Ginsburgh, Victor A and Jan C Van Ours, "Expert Opinion and Compensation: Evidence from a Musical Competition," *American Economic Review*, 2003, *93* (1), 289–296.

Gordon, Brett R and Wesley R Hartmann, "Advertising Effects in Presidential Elections," *Marketing Science*, 2013, *32* (1), 19–35.

Hamilton, Barton, Andres Hincapie, Robert Miller, and Nicholas W. Papageorge, "Innovation and Diffusion of Medical Treatment," *NBER working paper w24577*, 2018.

Hastings, Justine S and Jeffrey M Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *The Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Jin, Ginger Zhe and Phillip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *The Quarterly Journal of Economics*, 2003, *118* (2), 409–451.

Liu, Hongju, Qiang Liu, and Pradeep K Chintagunta, "Promotion Spillovers: Drug Detailing in Combination Therapy," forthcoming at *Management Science* 2014.

Liu, Yong, "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 2006, *70* (3), 74–89.

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier, "Promotional Reviews: An Empirical Investigation of Online Review Manipulation," *American Economic Review*, 2014, *104* (8), 2421–55.

Moretti, Enrico, "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales," *The Review of Economic Studies*, 2011, *78* (1), 356–393.

Munshi, Kaivan, "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 2004, *73* (1), 185–213.

Oster, Emily and Rebecca Thornton, "Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up," *Journal of the European Economic Association*, 2012, *10* (6), 1263–1293.

Papageorge, Nicholas W., "Why Medical Innovation is Valuable: Health, Human Capital, and the Labor Market," *Quantitative Economics*, 2016, *7* (3), 671–725.

Pope, Devin G, "Reacting to Rankings: Evidence from America's Best Hospitals," *Journal of Health Economics*, 2009, *28* (6), 1154–1165.

Reiter, Sara A and David A Ziebart, "Bond Yields, Ratings, and Financial Information: Evidence from Public Utility Issues," *Financial Review*, 1991, *26* (1), 45–73.

Schotter, Andrew, "Decision Making with Naive Advice," *American Economic Review*, 2003, *93* (2), 196–201.

Shapiro, Bradley T, "Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants," *Journal of Political Economy*, 2018, *126* (1), 381–437.

Sinkinson, Michael and Amanda Starc, "Ask Your Doctor? Direct-to-Consumer Advertising of Pharmaceuticals," *NBER working paper*, 2015.

Sorensen, Alan T, "Social Learning and Health Plan Choice," *The Rand Journal of Economics*, 2006, *37* (4), 929–945.

Ubel, P.A., A.M. Angott, and B.J. Zikmund-Fisher, "Physicians Recommend Different Treatments for Patients Than They Would Choose for Themselves," *Archives of Internal Medicine*, 2011, *171* (7), 630–634.

Wang, Justin, Jason Hockenberry, Shin-Yi Chou, and Muzhe Yang, "Do Bad Report Cards have Consequences? Impacts of Publicly Reported Provider Quality Information on the CABG Market in Pennsylvania," *Journal of Health Economics*, 2011, *30* (2), 392–407.

Tables and Figures

Table 1: Summary Statistics: MACS Dataset (Individual Level Variables)

	Mean	Std. Dev.	Min	Max
CD4 Count	536.4	283.9	5	3819
Non-decreasing CD4	0.54	0.50	0	1
No Ailment	0.63	0.48	0	1
AIDS	0.20	0.40	0	1
Age	47.15	8.21	19.5	80
Work Full-time	0.54	0.50	0	1
White	0.54	0.50	0	1
High School	0.19	0.39	0	1
College	0.50	0.50	0	1
Nobs.	13,472			

Notes: Summary statistics for the Multi-center AIDS Cohort Study (MACS) variables, which consists of 13,472 patient-visit observations. We restrict our sample to the years 1997-2008.

Table 2: Summary Statistics: Positively Aware Drug Guides Data

	Mean	Std. Dev.	Min	Max
Annual Cost	6690	4180	875	28007
No. of Side Effects	13.16	6.10	1	33
No. of Drug Interactions	14.26	10.34	0	43
Food Restrictions	0.34	0.48	0	1
Pill Burden (per take)	2.15	1.86	1	8
Dosage Frequency (per day)	1.94	0.65	1	3
DHHS Preferred	0.27	0.25	0	1
Publicly Traded Manuf.	0.90	0.28	0	1
Doctor's Rating	2.02	0.74	1	3
Activist's Rating	1.89	0.77	1	3
Disagreement	0.39	0.49	0	1
Nobs.	197			

Notes: Summary statistics for drug-level variables constructed using the Positively Aware annual drug guide, which consists of 197 drug-year observations. We restrict our sample to the years 1997-2008, and to drugs that have been FDA approved and can be matched to treatments observed in the MACS dataset. Doctor and activists' rating can take values 1, 2 or 3.

Table 3: Tabulation of Doctor's and Activist's Reviews

	Activist's Review = 1	Activist's Review = 2	Activist's Review = 3
Doctor's Review = 1	39	10	0
Doctor's Review = 2	20	49	16
Doctor's Review = 3	9	21	33

Notes: This table reports the number of observations for each combination of doctor's and activist's reviews.

Table 4: Summary Statistics: Combo Level

	Mean	Std. Dev.	Min	Max
(a) Reviews				
Doctor Average	2.18	0.57	0	3
Activist Average	2.07	0.56	0	3
Doctor Std. Dev.	0.51	0.38	0	1.41
Activist Std. Dev.	0.60	0.36	0	1.41
% of 3's - Doctor	0.37	0.34	0	1
% of 3's - Activist	0.32	0.30	0	1
% of 2's - Doctor	0.47	0.34	0	1
% of 2's - Activist	0.45	0.31	0	1
% of 1's - Doctor	0.14	0.23	0	1
% of 1's - Activist	0.21	0.26	0	1
Disagreement	0.70	0.46	0	1
(b) Objective Qualities				
Probability of Non-decreasing CD4	0.57	0.09	0.23	0.93
Probability of No Ailment	0.60	0.13	0.14	0.88
(c) Market Shares				
Combos	0.01	0.02	0	0.18
Fringe	0.32	0.05	0.23	0.42
Outside Option (No Drug)	0.19	0.04	0.12	0.27
Nobs.	1,086			

Notes: Panel (a) reports summary statistics for combo-level variables constructed using the Positively Aware annual drug guide. Panels (b) and (c) report combo-level variables constructed using the MACS dataset. The probability of non-decreasing CD4 count and probability of no ailment are constructed by averaging data across all individuals for each combo in every visit. Combos in the 'Fringe' category at a particular visit are taken by fewer than 25 individuals in that visit.

Table 5: Qualities and Combo Demand

	(1)	(2)	(3)	(4)	(5)	(6)
Prob of No Ailment	2.11***		2.10***	1.73***		1.71***
	(0.29)		(0.29)	(0.27)		(0.27)
Prob of Non-decreasing CD4		0.27	0.15		0.69*	0.59
		(0.42)	(0.41)		(0.39)	(0.38)
FDA Combo Characteristics:						
No. of Side Effects				−0.08***	−0.08***	−0.08***
				(0.01)	(0.01)	(0.01)
No. of Drug Interactions				−0.01	−0.01	−0.01
				(0.01)	(0.01)	(0.01)
Food Restrictions				−0.45***	−0.52***	−0.45***
				(0.16)	(0.16)	(0.16)
Nobs.	1,086	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. The table reports logit coefficients. Probability of no ailment and probability of non-decreasing CD4 count are combo characteristics constructed using the MACS dataset, while FDA combo characteristics are constructed using the Positively Aware annual drug guide by averaging across all drugs in a combo. The FDA combo characteristics include number of side effects, number of drug interactions and food restrictions for the combo.

Table 6: Qualities and Reviews

	Doctor			Activist		
	(1)	(2)	(3)	(4)	(5)	(6)
Prob of No Ailment	0.15		0.11	0.34***		0.30**
	(0.13)		(0.13)	(0.12)		(0.12)
Prob of Non-decreasing CD4		0.80***	0.79***		0.81***	0.78***
		(0.19)	(0.19)		(0.17)	(0.17)
FDA Combo Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Nobs.	1,086	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. The left-hand-side variable is either Doctor's or Activist's review (taking values between 0 and 3, where expert review = 0 for the outside option). The FDA combo characteristics include number of side effects, number of drug interactions and food restrictions for the combo.

Table 7: Reviews and Combo Demand with FDA Combo Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Doctor's Review	−0.01	0.08			−0.23**	−0.10
	(0.07)	(0.06)			(0.10)	(0.09)
Activist's Review			0.15**	0.18***	0.32***	0.25***
			(0.07)	(0.06)	(0.10)	(0.09)
Prob of No Ailment		1.71***		1.69***		1.68***
		(0.27)		(0.27)		(0.27)
Prob of Non-decreasing CD4		0.50		0.33		0.34
		(0.39)		(0.39)		(0.39)
FDA Combo Characteristics:						
No. of Side Effects		−0.08***		−0.08***		−0.08***
		(0.01)		(0.01)		(0.01)
No. of Drug Interactions		−0.01		−0.01		−0.01
		(0.01)		(0.01)		(0.01)
Food Restrictions		−0.45***		−0.48***		−0.49***
		(0.16)		(0.16)		(0.16)
Nobs.	1,086	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. The table reports logit coefficients. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. Both experts' reviews are constructed by averaging over drug reviews in each combo.

Table 8: Controlling for Big Changes within Combos

	(1)	(2)	(3)	(4)
Doctor's Review	−0.12	−0.04	−0.09	−0.01
	(0.09)	(0.09)	(0.09)	(0.09)
Activist's Review	0.31***	0.30***	0.25***	0.20**
	(0.09)	(0.10)	(0.09)	(0.09)
Doctor's Review Change > 50 pctile	−0.33***			
	(0.07)			
Activist's Review Change > 50 pctile	−0.16**			
	(0.07)			
Doctor's Review Change > 50 × Doctor's Review		−0.14		
		(0.09)		
Activist's Review Change > 50 × Activist's Review		0.27***		
		(0.09)		
Doctor's Review Change > 75 pctile			−0.41***	
			(0.09)	
Activist's Review Change > 75 pctile			0.10	
			(0.09)	
Doctor's Review Change > 75 × Doctor's Review				−0.15*
				(0.09)
Activist's Review Change > 75 × Activist's Review				0.29***
				(0.09)
Prob of No Ailment	1.63***	1.67***	1.70***	1.71***
	(0.27)	(0.27)	(0.27)	(0.27)
Prob of Non-decreasing CD4	0.23	0.23	0.26	0.27
	(0.38)	(0.39)	(0.39)	(0.39)
FDA Characteristics	Yes	Yes	Yes	Yes
Nobs.	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. Both experts' reviews are constructed by averaging over drug reviews in each combo.

Table 9: Timing Results

	Before	During	After
Doctor's Review	0.01	−0.06	−0.10
	(0.14)	(0.14)	(0.09)
Activist's Review	0.12	0.08	0.25***
	(0.14)	(0.14)	(0.09)
Prob of No Ailment	1.29***	1.67***	1.68***
	(0.42)	(0.41)	(0.27)
Prob of Non-decreasing CD4	−0.22	0.62	0.34
	(0.59)	(0.59)	(0.39)
FDA Combo Characteristics:			
No. of Side Effects	−0.08***	−0.09***	−0.08***
	(0.01)	(0.01)	(0.01)
No. of Drug Interactions	−0.00	−0.00	−0.01
	(0.01)	(0.01)	(0.01)
Food Restrictions	−0.80***	−0.55**	−0.49***
	(0.24)	(0.24)	(0.16)
Nobs.	464	460	1,086
F-value	0.81	0.17	4.69

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. F-test values are reported for joint significance of doctor's and activist's reviews. Column (1) reports results for market shares before the reviews are published, column (2) reports results for market shares during the period when reviews are published, and column (3) reports results for market shares realized after the reviews are published. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. Both experts' reviews are constructed by averaging over drug reviews in each combo.

Table 10: Reviews and Own and Rival Objective Qualities

	Doctor				Activist			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prob of No Ailment	0.15		0.11	0.03	0.34***		0.30**	0.26**
	(0.13)		(0.13)	(0.13)	(0.12)		(0.12)	(0.12)
Prob of Non-decreasing CD4		0.80***	0.79***	0.83***		0.81***	0.78***	0.80***
		(0.19)	(0.19)	(0.19)		(0.17)	(0.17)	(0.17)
Avg Rivals' Prob of No Ailment				−4.43**				−2.03
				(1.91)				(1.77)
Avg Rivals' Prob of Non-dec CD4				2.71*				−0.63
				(1.63)				(1.51)
FDA Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs.	1,086	1,086	1,086	1,086	1,086	1,086	1,086	1,086
F-stat	14.2	17.3	15.0	12.4	39.9	42.8	37.8	29.6

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. The left-hand-side variable is either Doctor's or Activist's review for a combo.

Table 11: Main Results — Baseline IV Logit Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Doctor's Review	−1.60***	−1.19**			−4.37***	−3.66***	−2.70***
	(0.58)	(0.58)			(1.09)	(0.98)	(0.89)
Activist's Review			2.44***	1.99***	5.39***	3.84***	1.57**
			(0.76)	(0.59)	(1.25)	(0.87)	(0.74)
Prob of No Ailment		2.00***		1.93***		1.46***	1.55***
		(0.33)		(0.37)		(0.47)	(0.38)
Prob of Non-decreasing CD4		1.53*		−2.64***		−1.04	1.39*
		(0.81)		(0.98)		(1.27)	(0.84)
FDA Combo Characteristics:							
No. of Side Effects							−0.07***
							(0.01)
No. of Drug Interactions							0.03*
							(0.02)
Food Restrictions							−0.69***
							(0.25)
No. of Individuals	13,472	13,472	13,472	13,472	13,472	13,472	13,472
Combo-time dyads	1,086	1,086	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. The table reports the logit coefficients. Doctor's and Activist's reviews have been instrumented using the average probability of no ailment and non-decreasing CD4 count of rival combos as well as the standard deviation of the probability of no ailment and non-decreasing CD4 count of rival combos. Combo-visit dyad is the unit of analysis. The total number of combo-visit observations used for the estimation is 1,086, which are constructed using data on 13,472 individuals.

Table 12: Main Results — Baseline Model Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
(a) Percentage of High Reviews						
% of 3's - Doctor	1.72***	1.57***			−3.20***	−2.41***
	(0.44)	(0.43)			(0.85)	(0.75)
% of 3's - Activist			2.66***	2.59***	5.58***	4.69***
			(0.46)	(0.46)	(0.94)	(0.81)
Prob of No Ailment		1.78***		1.62***		1.51***
		(0.29)		(0.30)		(0.37)
Prob of Non-decreasing CD4		−0.36		−0.62		−0.57
		(0.44)		(0.44)		(0.54)
(b) Indicator Variables for High Reviews						
Doctor's Review - High	0.68	−1.07			−1.39	−2.46**
	(0.64)	(0.76)			(1.09)	(1.15)
Activist's Review - High			4.58***	4.16***	5.04***	4.68***
			(0.88)	(0.86)	(0.97)	(0.97)
Prob of No Ailment		2.33***		1.05**		1.43***
		(0.35)		(0.48)		(0.55)
Prob of Non-decreasing CD4		0.10		−0.69		−0.91
		(0.45)		(0.63)		(0.69)
(c) Review Average and Standard Deviation						
Doctor's Review	−0.29	−0.58			−2.32	−1.93
	(0.43)	(0.53)			(1.93)	(1.26)
Doctor's Review SD	−5.86***	−5.68***			−4.43	−4.57
	(1.55)	(1.75)			(3.65)	(3.57)
Activist's Review			1.18***	1.15***	2.88	2.44
			(0.35)	(0.38)	(2.13)	(1.77)
Activist's Review SD			−3.36*	0.63	−5.43	−4.64
			(2.00)	(2.25)	(4.80)	(5.98)
Prob of No Ailment		0.86		1.90***		0.32
		(0.63)		(0.41)		(1.28)
Prob of Non-decreasing CD4		2.73**		−1.50**		0.39
		(1.35)		(0.72)		(2.66)
No. of Individuals	13,472	13,472	13,472	13,472	13,472	13,472
Combo-time dyads	1,086	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. The table reports the logit coefficients. In Panel (a), we use the percentage of drugs that receive a rating of 3 in a combo as a measure of 'high' reviews. In Panel (b) we add indicator variables for whether the average review of experts across drugs in a combo are high. Doctor's Review - High is an indicator variable that is 1 if the doctor's review for the combo is greater than 2. Similarly, Activist's Review - High is an indicator variable that is 1 if the activist's review for the combo is greater than 2. Doctor's and Activist's reviews have been instrumented using the average probability of no ailment and non-decreasing CD4 count of rival combos as well as the standard deviation of probability of no ailment and non-decreasing CD4 count of rival combos. In Panel (c), our measure of reviews for the two experts includes the average across all drugs in a combo, as well as the standard deviation of reviews across drugs in a combo. In all cases, we use the average objective qualities (probability of no ailment and probability of non-decreasing CD4 count) of rival combos as instruments for reviews. Combo-visit dyad is the unit of analysis. The total number of combo-visit observations used for the estimation is 1,086, which are constructed using data on 13,472 individuals.

Table 13: Main Results — Disagreements and Demand

	(1)	(2)	(3)
Doctor's Review	-3.66^{***}		
	(0.98)		
Activist's Review	3.84^{***}		
	(0.87)		
Agree × Doctor's Review		0.65	1.08
		(1.69)	(1.19)
Disagree × Activist's Review		3.19^{***}	
		(0.85)	
Disagree × Doctor's Review		-2.78^{***}	
		(0.79)	
Agree		−0.85	1.97
		(3.39)	(4.34)
Positive Disagreement × Doctor			3.74
			(3.60)
Negative Disagreement × Doctor			-3.91^{***}
			(1.16)
Positive Disagreement × Activist			−1.25
			(3.96)
Negative Disagreement × Activist			7.15^{***}
			(2.26)
Prob of No Ailment	1.46^{***}	1.68^{***}	1.71^{***}
	(0.47)	(0.48)	(0.65)
Prob of Non-decreasing CD4	−1.04	−1.49	−1.99
	(1.27)	(1.77)	(1.34)
No. of Individuals	13,472	13,472	13,472
Combo-time dyads	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. The table reports the logit coefficients. Doctor's and Activist's review have been instrumented using the average probability of no ailment and average probability of non-decreasing CD4 count of rival combos. The variable 'Agree' is a dummy which is 1 if both experts give the same rating to a combo. The variable 'Disagree' is a dummy which is 1 if each expert gives a different rating to a combo. Finally, the variable 'Positive Difference' is a dummy which is 1 if the doctor's review is lower than the activist's review, while the variable 'Negative Difference' is a dummy which is 1 if the doctor's review is higher than the activist's review. Combo-visit dyad is the unit of analysis. The total number of combo-visit observations used for the estimation is 1,086, which are constructed using data on 13,472 individuals.

Table 14: Reviews and Objective Qualities when Experts Disagree

	Doctor	Activist
Prob of No Ailment	-0.10	0.23^{**}
	(0.11)	(0.11)
Prob of Non-decreasing CD4	0.44^{***}	0.40^{**}
	(0.17)	(0.17)
Avg Rivals' Prob of No Ailment	-8.04^{***}	-4.97^{***}
	(1.19)	(1.20)
Avg Rivals' Prob of Non-dec CD4	-3.74^{***}	-5.45^{***}
	(1.34)	(1.35)
PA Characteristics	Yes	Yes
Nobs.	671	671

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Combo-visit dyad is the unit of analysis. The sample is restricted to cases in which the two experts' ratings are different from each other. The left-hand-side variable is either Doctor or Activist's review.

Table 15: Demand Model with Individual Attributes

	Estimate	Std. Error
Means (β)		
Doctor's Review	−5.86	0.00
Activist's Review	3.15	0.04
Constant	5.94	0.13
Prob of No Ailment	0.70	0.00
Prob of Non-decreasing CD4	0.40	0.01
Individual Attributes		
AIDS	0.06	0.05
Age	0.20	0.13
Full-time Work	−1.59	0.67
Black	−0.17	0.18
College	0.50	0.03
Same Combo Last Period - Fringe	−0.32	0.19
Same Combo Last Period - Other	2.00	1.90
Interactions with Individual Attributes		
Doctor's Review × AIDS	9.50	2.38
Doctor's Review × Age	0.26	0.01
Doctor's Review × Full-time work	−1.32	0.99
Doctor's Review × Black	−2.05	2.02
Doctor's Review × College	3.41	1.25
Doctor's Review × SC - Fringe	−3.72	0.49
Doctor's Review × SC - Other	3.32	1.43
Activist's Review × AIDS	1.48	1.25
Activist's Review × Age	0.38	0.31
Activist's Review × Full-time Work	0.55	0.22
Activist's Review × Black	−0.06	0.05
Activist's Review × College	−0.30	0.28
Activist's Review × SC - Fringe	−5.47	0.84
Activist's Review × SC - Other	4.33	1.76

Notes: The table reports coefficients for the IV-logit demand model with individual characteristics. Combo-visit dyad is the unit of analysis. Doctor's review, activist's review, probability of no ailment and probability of non-decreasing CD4 count vary only over combo and visit. The variable 'Same Combo Last Period - Fringe' is a dummy for whether the individual taking a fringe combo was also taking a combo from the fringe group (combinations taken by less than 25 individuals in a visit) in the last visit, and 'Same Combo Last Period - Other' is a dummy which is 1 if the individual was taking the same combo (including the outside option) last visit that he is taking in the current period. The model is estimated using Generalized Method of Moments (GMM).

Nucleoside Reverse Transcriptase Inhibitor

Brand Name: Retrovir

Common Name: zidovudine (ZDV) or AZT

Class: nucleoside analog (also called nucleoside reverse transcriptase inhibitor, NRTI or nuke)

Standard dose: One 300 mg tablet twice-a-day (12 hours apart); two 100 mg capsules three times a day also available, no food restrictions (may be taken with or without food). Clear, strawberry-flavored liquid available for pediatric use. Take missed dose as soon as possible, but do not double up on your next dose. Generic Retrovir (zidovudine) is available.

AWP: $432.88 (generic $315) / month

Manufacturer contact: GlaxoSmithKline, www.treathiv.com, 1 (888) 825–5249

AIDSInfo: 1 (800) HIV–0440 (448–0440), www.aidsinfo.nih.gov

Potential side effects and toxicity: Most common side effects include headaches, fever, chills, muscle soreness, fatigue, nausea, and fingernail discloration. Zidovudine (AZT) has been associated with alteration of various cells in the blood through bone marrow suppression resulting in anemia (low red blood cells) and/or neutropenia (low white blood counts), particularly in people with advanced HIV during the first three months. Potential for severe anemia requiring blood transfusion, erythropoietin injections, or hospitalization when used on its own or in combination with hydroxyurea. Prolonged use of high doses of zidovudine has been associated with symptomatic myopathy (muscle damage). Rare but potentially fatal toxicity with all NRTIs is pancreatitis (inflammation of the pancreas), hepatomegaly (enlarged liver) with steatosis (fat) and lactic acidosis (accumulation of lactate in the blood and abnormal acid-base balance). Lactic acidosis has been seen in patients taking NRTIs but is more common and more severe in women, people who are obese, and people who have been taking nukes for a long time; and more common in people with liver disease, but can occur in people without a history of liver damage. People with lactic acidosis may experience persistent fatigue, abdominal pain or distension, nausea/vomiting, and difficulty breathing or shortness of breath; and enlarged, fatty liver. Pancreatitis can be life-threatening and may cause pain in the stomach and back, along with nausea, vomiting and blood in the urine. Risks for pancreatitis include: higher than recommended doses of NRTIs, advanced HIV, and alcohol use. The risk for pancreatitis with zidovudine is low compared to ddI.

Potential drug interactions: Biaxin, Mycobutin, and rifampin (under various brand names) may decrease zidovudine blood levels. Benemid (probenecid), Dilantin (phenytoin), and Depakote (valproic acid) may increase zidovudine blood levels and decrease zidovudine clearance, but no dosing adjustments are recommended. Zidovudine and Zerit should not be used together due to evidence that one limits the other's effectiveness. Also, bone marrow supression should be monitored with use of Cytovene (ganciclovir), Valcyte, amphotericin B, pentamidine, dapsone, flucytosine, sulfadiazine, interferon-alpha, ribavirin (Rebetol), and with cancer treatments such as hydroxyurea and doxorubicin. Ribavirin and zidovudine may cancel each other out, so this combination should be monitored closely. New Procrit or Epogen warning: if hemoglobin target is above manufacturer's recommendation (12 g/dL), the risk for serious and life-threatening cardiovascular complications significantly increases. For zidovudine patients, measure hemoglobin once a week after starting the anemia drugs until hemoglobin has stabilized. Notify healthcare provider if experiencing pain and/or swelling in the legs, worsening in shortness of breath, increases in blood pressure, dizziness or loss of consciousness, extreme tiredness, or blood clots in hemodialysis vascular access ports. Do not take with Combivir or Trizivir, since zidovudine is already in these medications.

Tips: In combination with Epivir, zidovudine is recommended as a preferred NRTI agent in U.S. HIV treatment guidelines in people on HIV therapy for the first time. The not-so-good news for people adding zidovudine: the fatigue and the potential anemia. You can start taking erythropoietin (Procrit or Epogen) for some anemias, but that's adding an expensive weekly injectable. Some doctors would prefer switching out the zidovudine for another drug. Also, some clinicians avoid the "T" drugs, or thymidine analogs (zidovudine and Zerit) because of implication in lipoatrophy. Zidovudine has for years been associated with "AZT butt," a disheartening flatness that happens gradually. Taking with food may minimize upset stomach. Please see package insert for more complete potential side effects and interactions.

Doctor

Retrovir, more commonly called AZT, was the first drug approved for the treatment of HIV infection, and it prolonged many lives back in the late '80's and early '90's. It got a new life in the form of Combivir after 3TC became available, experienced another resurrection as part of Trizivir, a once popular "triple-nuke" combination, and has been a cornerstone of therapy in the HAART era. However, AZT's time has finally passed. Compared to the nukes we're using now (namely tenofovir and abacavir), it's weaker, is dosed twice a day, is harder on the stomach, is more prone to resistance, and causes anemia and mitochondrial toxicity, including lipoatrophy. I still have a *few* patients still taking AZT because of resistance to other drugs (it becomes stronger if you have mutations that cause resistance to 3TC, FTC, abacavir, or tenofovir), but that may change as newer, safer agents become available. So long, AZT, and congratulations on a good, long run!—Joel Gallant, M.D.

Activist

Retrovir/AZT was the first drug developed for the treatment of HIV. In subsequent years, activists fought many battles to speed up the drug development process, but the history of AZT demonstrates that the mechanisms and ability to quickly test and approve drugs were present all along. What was lacking, except in the case of AZT, was the will to do it. AZT certainly has served a useful place in the history of treatment for HIV, but it has always come at a price. There is almost a cultural memory of the early and often severe side effects, but people don't always remember that this was primarily the result of overdosing. When dosed properly, AZT can still have side effects but they are seldom severe. Still, many people today believe it is time to reconsider the whole class of drugs that AZT comes from. Most of them have potentially significant side effects that derive from the very nature of what they are doing. It is difficult to conceive of a drug of this type that would be completely free of side effects. With so many new and relatively non-toxic drugs becoming available in recent years, it may be time to ask whether we can build fully effective regimens that don't rely on the old paradigm of "two nukes and a protease inhibitor" or "two nukes and a non-nuke." When this paradigm first became standard in 1996, it wasn't chosen because this was inherently the right or best way to treat HIV. Rather, it was simply the only kind of combination available at the time.—Martin Delaney

Figure 1: Sample Page from the 2008 Positively Aware Drug Guide

"Am I ready to take the medications, as prescribed, every day for the rest of my life?" When people ask me how I've lived so long, what I have to say is, "Good doctors who listen, taking my meds when and like I am supposed to, and bucking up when I don't like the side effects, *and*, yes, LUCK and a good attitude doesn't hurt either.

—*Greg*
VIA THE INTERNET

THANKS FOR THE DRUG GUIDE

I just wanted to write to thank you for putting together your annual Drug Guide.

As someone who was diagnosed in 2009, it has been an anchor in the storm for me.

The first issue of POSITIVELY AWARE that I read was the January/February issue which, I have to say, blew my mind. I live in a small town and your article about access to care in rural areas really hit home. I travel almost 200 miles to see a doctor, but thanks to that article, I researched my area until I found the one HIV specialist within driving distance.

And then I got the Drug Guide. My doctor had put me on Atripla right away and though I've had no problem with it and my viral load has come down, reading about the other drugs, the side effects, and the drugs being developed really opened my eyes. Before, I guess I was just in shock and willing to do anything to stay healthy. After reading the Drug Guide, I'll do my "homework" before just agreeing to any particular treatment. It has also helped me in talking to my doctor, who told me he always has a copy on hand and the chart picturing the drugs hangs in his office.

You guys do a great job and I will look forward to reading every issue of POSITIVELY AWARE.

—*Marlen P.*
LACEY, WA

A CAPTIVE AUDIENCE

Greetings from a California state prison. I am truly blessed to receive a subscription to your invaluable magazine.

Not only have I been HIV-positive since 2001, a prisoner, but I am also a male-to-female transgender person. Your magazine is surely a God-send.

I was wondering if you could do an article on hormone treatment and HIV and also HIV meds. I know there are many who could use this information.

Thank you for all your insights, hints, and articles—I truly appreciate this much-needed resource.

We don't have a death sentence; we just share our bodies with a virus. I always say, "I don't live with HIV—it has to live with me!"

—*Nathan*

DRUG GUIDE KUDOS

Bravo for your 12th Annual HIV Drug Guide (January/February)! I have been an HIV social worker for four years and found the issue extremely easy to read for myself and my clients. I even gave a copy of it to the physicians in the clinic! Each medication was explained by your staff, Dr. Gallant and Martin Delaney in a way that most people can understand. The general information on each of the drugs is so helpful—I keep it in my top drawer to pull out when I forget the generic name or the dosage of a medication. I loved the medication class color coding, the Drug Interactions Chart, the Side Effects Chart, and especially the "centerfold" Drug Chart. It is the only one I have gotten with all the new medications. Thank you, and keep up the amazing work!

Figure 2: READER COMMENTS ABOUT DRUG GUIDE: Figure 2 shows reader's comments about the annual drug guides and presents anecdotal evidence that patients read the drug guides and learn about new drugs on the market.

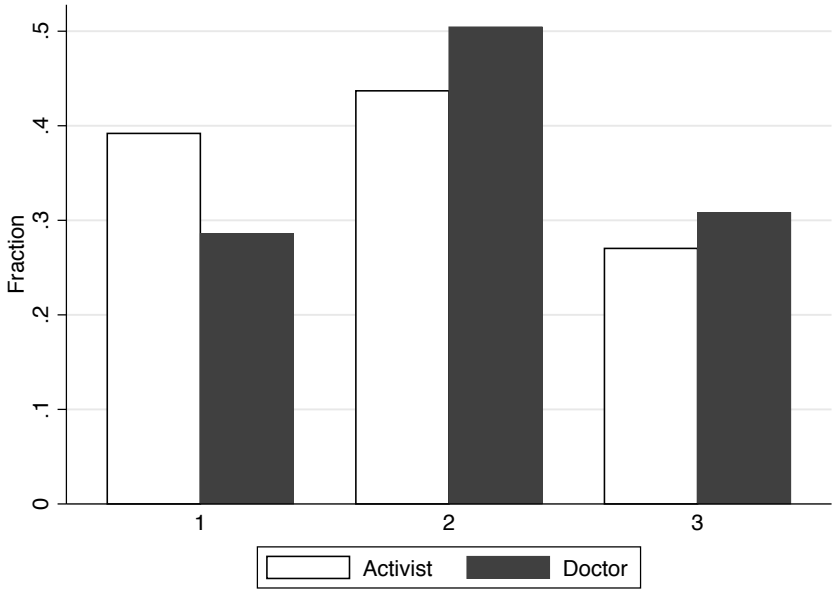


Figure 3: COMPARISON OF DOCTOR AND ACTIVIST RATINGS: The Figure plots the fraction of 1's, 2's and 3's given to individual drugs, by expert.

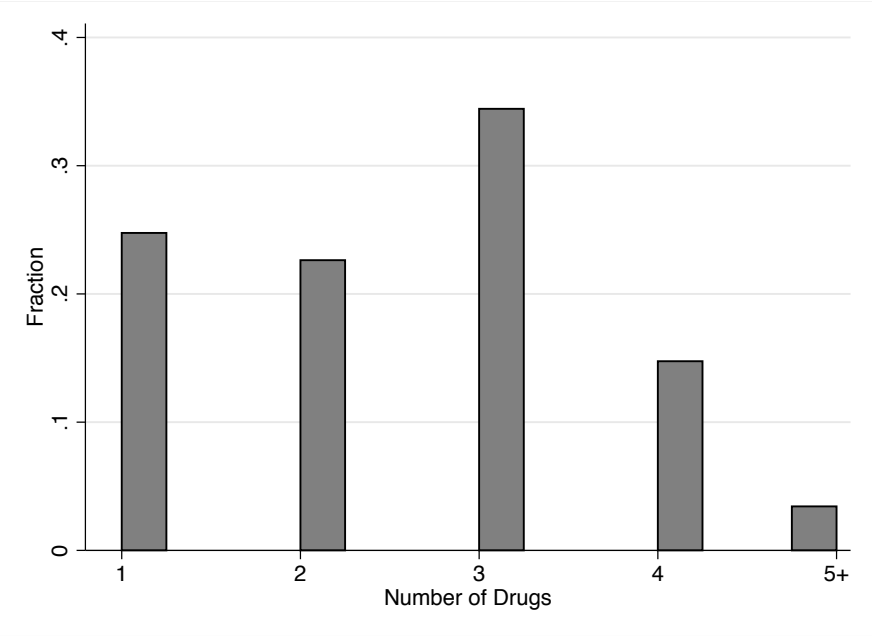


Figure 4: Distribution of Number of Drugs Taken Together: The Figure plots the distribution of drugs taken together in a combo.

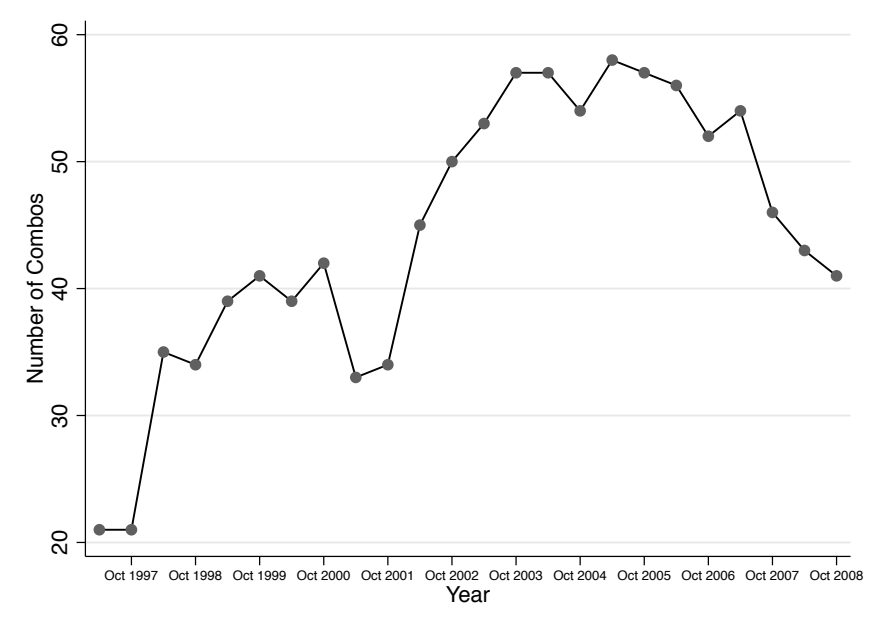


Figure 5: Total Number of Combos over Time: The Figure shows how the total number of combos (including 'Fringe') observed in the data evolves over the period of analysis.

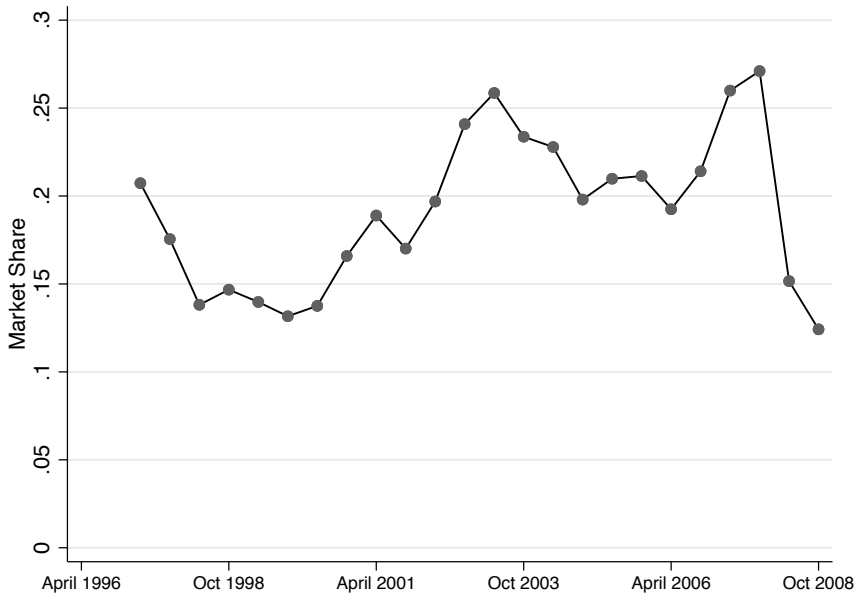
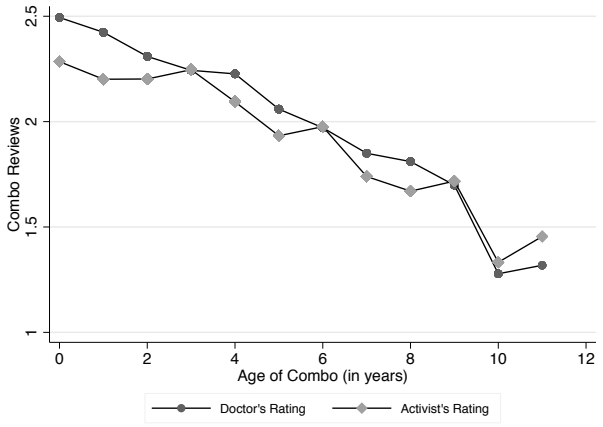
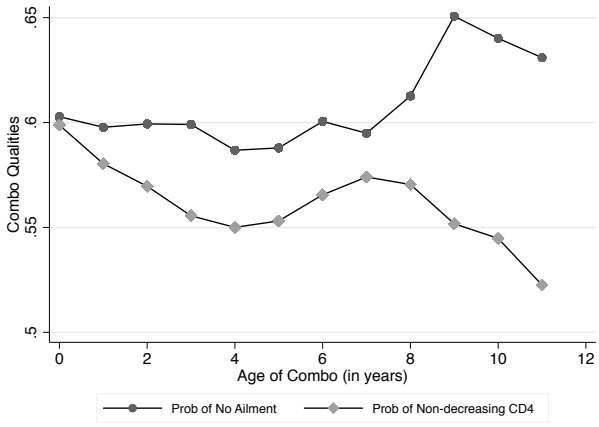


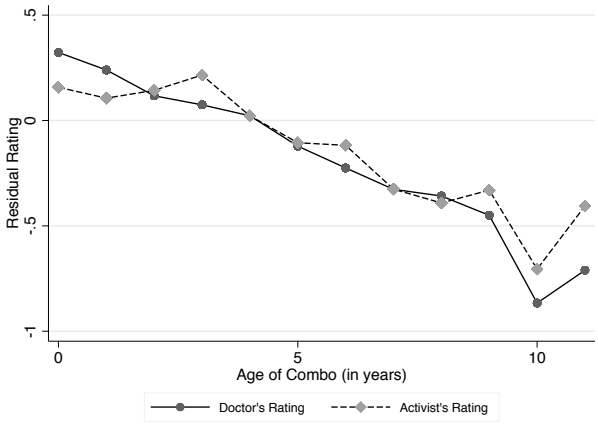
Figure 6: Outside Option Market Share: The Figure plots how the market share of the outside option, defined as taking no HIV treatment, evolves over the period of analysis.



(a) Ratings over Combo Age



(b) Combo Qualities over Combo Age



(c) Residual Ratings over Combo Age

Figure 7: Combo Reviews and Qualities over the Life Cycle: Figure 7 (a) shows how the average combo ratings of the two experts evolves over the age of the combo. Figure 7 (b) plots the evolution of objective qualities of combos, probability of no ailment and probability of non-decreasing CD4 count, over combo age. Lastly, Figure 7 (c) plots residual ratings for combo over combo age, where the residual ratings are the residual of an OLS regression of combo ratings on two objective qualities, probability of no ailment and probability of non-decreasing CD4 count.

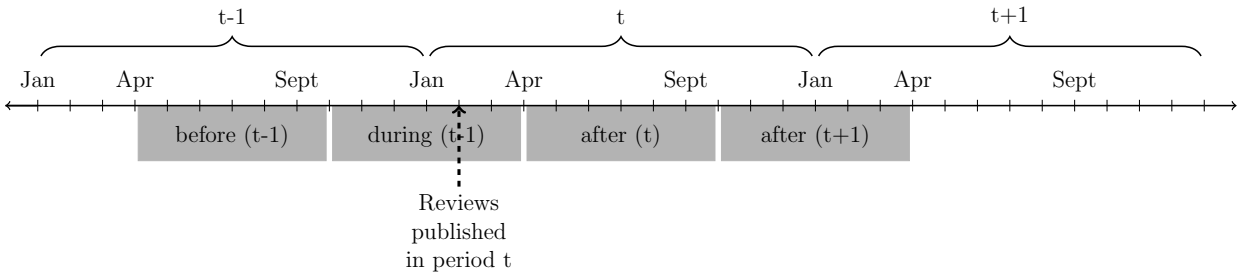
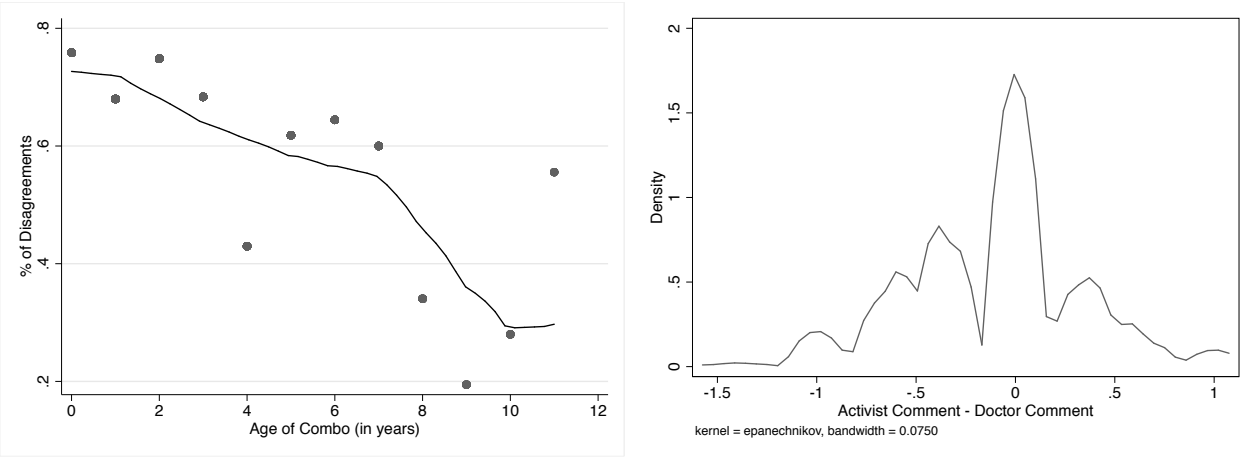


Figure 8: Timeline of Events: The Figure shows the timeline of events studied in the paper. Market share data is available for two six month windows, spanning from April to September and October to March. *PA* annual drug guides are published in January/February of every year, which coincides with the October-March window from the MACS data.



(a) Disagreement over Combo Age **(b)** Distribution of Differences

Figure 9: Disagreements: Figure 9 (a) plots the percentage of disagreements between the doctor and activist about the rating of the combo over the age of the combo, where the variable disagreement is a dummy which is 1 if the activist and the doctor have a different rating for the combo. Figure 9 (b) plots the distribution of the difference in combo ratings between the activist and the doctor.

Appendix A Data Collection

A.1 *Positively Aware* Data Dictionary

In this section, we present a data dictionary for the constructed dataset from the *Positively Aware* magazines. Below is a list of variables that we derived from the magazines, along with a description of what that variable measures.

- Common Name - This codes the generic name of the drug.
- Brand Name - This variable codes the brand name under which the drug is sold.
- Class - Class of drugs that the drug belongs to.
- Manufacturer - Name of the manufacturer.
- Public - A binary variable, indicating whether the drug company is publicly traded.
- Year - Year the magazine was published.
- No. of Side Effects - Number of side effects for the drug listed in the drug guide.
- No. of Drug Interactions - Number of drug interactions with other drugs listed in the drug guide.
- Pill Burden - Number of tablets that need to be taken together.
- Dosage Frequency - Number of times a day the drug dose needs to be taken.
- Food Restrictions - A binary variable indicating whether drug intake has any food restrictions.
- Annual Cost - Average Wholesale Price of drugs, as specified by the manufacturer
- DHHS Preferred - A binary variable, indicating whether the drug has been approved as first-line therapy by the Department of Health and Human Services.
- Doctor's Rating - A categorical variable that encapsulates a doctor's rating of the drug on a scale of 1 to 3.
    1. Doctor mainly uses negative words or phrases to describe the drug.
    2. Doctor says positive things, with some qualifications.
    3. Doctor says mostly positive things.

- Activist's Rating - A categorical variable that encapsulates the activist's rating of the drug on a scale of 1 to 3.

    1. Activist mainly uses negative words or phrases to describe the drug.
    2. Activist says positive things, with some qualifications.
    3. Activist says mostly positive things.

- Doctor - The variable codes the name of the doctor who has reviewed for the current issue of the drug guide.

- Activist - The variable codes the name of the activist who has reviewed for the current issue of the drug guide.

Table A1 presents a summary of all the drugs in the dataset, along with their manufacturer details and year of entry and exit.

Doctor and Activist Reviews

In order to create a ranking system for the reviews, we use the following set of criteria:

- Assign a rating of 1 if mostly negative words or phrases have been used to describe the drug. For example, comments such as "*There is* ***not much to say*** *about ddC anymore.*" ... "***hard to get excited about it****, and these days it's often not prescribed.*" ... "*The role for delavirdine* ***remains unclear****.*", or an activist's comments such as "*ddC has* ***never lived up to its initial promise****" ... "****overall, not a very useful drug***" ... "*Invirase was* ***extraordinarily weak ... not much reason to take it.*** " would be assigned a rank of 1.

- Assign a rating of 2 if the doctor or advocate points out the positive as well as the negative aspects of the drug, but does not give an absolute recommendation of whether the drug is good or bad. For example, comments of the form "*The new soft-gel formulation achieves* ***much better drug levels*** *... but if you are going to use Fortovase as a sole PI,* ***you will have to take a lot of pills****.*", and "*It may not be the best bet to include in first-line treatment ...* ***but it remains a solid antiviral****.*"

- Assign a rank of 3 to drugs with reviews that mostly use positive words to describe the drug. For example, "*3TC is a* ***potent, convenient and well-tolerated drug****"* or, "*3TC, with its* ***minimal side effects, easy dosing schedule and high potency****, may be the* ***most useful of the nucleosides*** " would receive a rank of 3.

Table A1: Drug Information

	Manufacturer	Year of Introduction	Year of Discontinuation
(a) NRTI			
Retrovir	GlaxoSmithKline	1987	-
Videx	Bristol-Myers Squibb	1997	-
Hivid	Hoffman-LaRoche	1997	2006
Zerit	Bristol-Myers Squibb	1997	-
Epivir	GlaxoSmithKline	1997	-
Combivir	GlaxoSmithKline	1998	-
Ziagen	GlaxoSmithKline	1999	-
Viread	Gilead Sciences	2000	-
Trizivir	GlaxoSmithKline	2001	-
Emtriva	Gilead Sciences	2004	-
Epzicom	GlaxoSmithKline	2004	-
Truvada	Gilead Sciences	2004	-
(b) NNRTI			
Viramune	Boehringer Ingelheim	1997	-
Rescriptor	Agouron Pharmaceuticals	1997	-
Sustiva	Bristol-Myers Squibb	1998	
(c) PI			
Norvir	Abbott Laboratories	1997	-
Crixivan	Merck & Company	1997	-
Viracept	Agouron Pharmaceuticals	1997	-
Saquinavir	Hoffman-LaRoche	1997	-
Agenerase	GlaxoSmithKline	1999	-
Kaletra	Abbott Laboratories	2000	-
Aptivus	Boehringer Ingelheim	2001	-
Reyataz	Bristol-Myers Squibb	2002	-
Lexiva	GlaxoSmithKline	2004	-
Prezista	Tibotec Therapeutics	2004	-

Notes: The table lists details about all drugs in the sample, grouped by drug type. HIV drugs belong to three drug types: Nucleoside Reverse Transcriptase Inhibitor (NRTI), Non-nucleoside Reverse Transcriptase Inhibitor (NNRTI) and Protease Inhibitor (PI). During our period of analysis, only one drug was discontinued.

Appendix B Demand Estimation

We estimate the demand model by GMM, matching the moments predicted by the model to the sample moments. We match two sets of moments to their sample analogue: (1) the market shares for all combinations, and (2) the covariance of the observed product characteristics, $\mathbf{x}$, with the observed individual-level characteristics, $\mathbf{z}$.

For computational ease, we assume that the ϵ_{ijt}'s have an independently and identically distributed extreme value distribution, which leads to the familiar closed-form for the model's choice probabilities conditional on $\mathbf{z}$:

$$\Pr_t(y = j|\mathbf{x}, \mathbf{z}, \boldsymbol{\theta}) = \frac{\exp\left(\delta_{jt} + \sum_{kr} x_{jtk} z_{ir} \beta_{kr}\right)}{1 + \sum_q \exp\left(\delta_{qt} + \sum_{kr} x_{qtk} z_{ir} \beta_{kr}\right)} \tag{9}$$

In order to compute our moments, we first find the value of $\boldsymbol{\delta}$ that makes the market shares from the data, s_{jt}^N, equal to the market shares predicted by the model,[59] $s_{jt}(\boldsymbol{\delta}, \boldsymbol{\beta}; .)$, for each guess at $(\boldsymbol{\beta})$. We then substitute that $\boldsymbol{\delta}(\boldsymbol{\beta}, s_{jt}; .)$ for δ into the model's prediction for the micro moments, making them a function of $(\boldsymbol{\beta}, \boldsymbol{\delta}(\boldsymbol{\beta}, s_{jt}; ,))$. Lastly, we search over $(\boldsymbol{\beta})$ to minimize the distance between model's predictions for the micro moments and the data.

Recall that we also need to address the endogeneity problem of the reviews, since we expect reviews and ξ_{jt} to be correlated. The instruments we use are the average combo characteristics of rival drugs on the market. Let $\boldsymbol{Z} = [Z_1, Z_2]$ be the set of instruments, where Z_1 is the average probability of no ailments for the rival drugs on the market, and Z_2 is the average probability of non-decreasing CD4 count for the rival drugs on the market.

We now describe our estimation algorithm in detail:

1. Let $\mathbf{z}_d$, for $d = 1, \ldots, ns$, be the individual-level characteristics for the ns individuals in visit t from the individual level data from MACS. We then define $\boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta})$ as the value of $\boldsymbol{\delta}$ for a given value of $\boldsymbol{\beta}$ that sets

$$g_1^{ns,N}(\boldsymbol{\theta}) = s_{jt}^N - \frac{1}{ns} \sum_{d=1}^{ns} \Pr_t\left(y = j|\mathbf{x}, \mathbf{z}_d, \boldsymbol{\beta}, \boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta})\right) \tag{10}$$

   equal to $\mathbf{0}$.

2. Calculate the model's prediction for the covariances between the characteristics of the chosen combination and individual-level attributes. In particular, to form the sample

[59] For the logit specification, that is simply equal to the log market share of combo c minus the log of the share of the outside option (taking no drugs).

moment, we interact the average attributes of the individuals that chose combination j at time t with the characteristics of the combination at time t, and then average over all available combinations in that time period. Formally, the second moment is defined as:

$$g_2^{n,ns}(\boldsymbol{\theta}) \approx \frac{1}{n}\sum_j n_j x_{kj}\left\{\frac{\sum_{i_j=1}^{n_j} z_{i_j}}{n_j} - E\left[\mathbf{z}|y=j,\boldsymbol{\beta},\boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta})\right]\right\} \tag{11}$$

where

$$E[\mathbf{z}|y=j,\boldsymbol{\beta},\boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta})] = \frac{(ns)^{-1}\sum_d \mathbf{z}_d \Pr_t(y=j|\mathbf{x},\mathbf{z}_d,\mathbf{v}_d,\boldsymbol{\beta},\boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta}))}{s_{jt}^n}, \tag{12}$$

n_j is the number of individuals taking combination j, $n = \sum_j n_j$ and $\Pr_t(y=j|\mathbf{x},\mathbf{z}_d,\mathbf{v}_d,\boldsymbol{\beta},\boldsymbol{\delta}^{ns,N}(\boldsymbol{\beta}))$ is given by equation (9).

3. Calculate $\bar{\beta}_k$ using the IV GMM formula, and then, using $\boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta})$ from step 1, calculate the error term as

$$\omega_{jt}(\boldsymbol{\theta}) = \boldsymbol{\delta}^{ns,n}(\boldsymbol{\beta}) - \sum_k x_{jtk}\bar{\boldsymbol{\beta}}, \tag{13}$$

to calculate the third moment, which is given by:

$$g_3 = E[\boldsymbol{Z}\omega(\boldsymbol{\theta})] = 0 \tag{14}$$

4. Find the generalized method of moments estimator of $(\boldsymbol{\theta}_{GMM}) = (\boldsymbol{\beta_{GMM}}, \bar{\boldsymbol{\beta}}_{\boldsymbol{GMM}})$ from stacking g_2 and g_3 into a single vector f. In particular, we use a two-step estimation procedure with

$$(\boldsymbol{\beta_{GMM}}, \bar{\boldsymbol{\beta}}_{\boldsymbol{GMM}}) = argmin\left(\frac{1}{n}\sum_{i=1}^n f(\theta)\right)^T \hat{W}\left(\frac{1}{n}\sum_{i=1}^n f(\theta)\right) \tag{15}$$

where $W = E[f(\theta)f(\theta)']$. With the optimal weight matrix, the variance-covariance of the parameters $\boldsymbol{\theta_{GMM}}$ is given by:

$$\hat{V}(\boldsymbol{\theta_{GMM}}) = (\hat{G}^T\hat{W}\hat{G})^{-1} \tag{16}$$

Appendix C Additional Robustness Checks

For additional robustness checks, we begin controlling for expert and manufacturer fixed effects. Table C1, column (3) shows that even after controlling for doctor and activist fixed effects, we find that doctor's review negatively affects demand while activist's review positively affects demand. The results remain unchanged even when we control for manufacturer fixed effects. To account for the fact that a combo may be composed of drugs manufactured by different companies, we create dummies for every manufacturer combination. We also test to see if the effect of reviews is different for the time periods when drug guides were published online, along with print publication. In years 2007 and 2008, Positively Aware drug guides were also available online on the Positively Aware website. We would expect that due to higher dissemination of these reviews through the online publication, the effect of reviews during those years would be higher. Column (5) shows that for the years the drug guides were also available online, the effect of both experts reviews was highly significant, and higher than the effect of reviews during years when only print publications were available.

Next, we see the effect of pooling the doctor and activist reviews. Table C2 presents the results of the logit with instruments for two ways of pooling the reviews: adding the two reviews for each combination, and taking the maximum of the two reviews for each drug. For both measures, we find that even after controlling for objective qualities, an increase in reviews leads to an increase in the likelihood of choosing the drug combination.

In Table C3, we report results for the specification in which we control for individual and time fixed effects when predicting the probabilities of non-decreasing CD4 count and no ailment for each individual. As before, doctors' and activists' reviews positively predict demand independently; however, in the specification in which we control for both the activists' and doctors' reviews together and control for the combination's objective qualities, we find that a higher review from the doctor decreases the probability of choosing that combination while a higher review by the activist for a combination leads to an increase in the probability of that combination being demanded. The disagreement results are the same, yet in this specification the interaction between the doctors' review and disagreement is not significant.

Lastly, we also check if our mechanism for explaining the negative coefficient on doctor's review is robust to how we define the reviews. Therefore, we use the definition for reviews in which we calculate the percentage of drugs in a combination that have a rating of 3 as our measure of combo-level reviews and run the specification with agreements and disagreements between the two experts. Table C4, column (1) replicates the results for this definition of reviews with which we find that after we control for the activist's review and the objective

qualities, the doctor's review negatively affects demand. In column (2), we find that if the experts agree about a combination, then a higher review increases the likelihood of taking that combination. However, in the case of a disagreement, a higher activist's review leads to an increase in the likelihood of taking the combination while a higher doctor's review decreases the likelihood of taking that combination (though the effect is not significant). In column (3), we explore the non-linearities in disagreements and find that if the activist gives a lower review to the combination than the doctor (i.e., a smaller percentage of drugs in the combo receive a rating of 3 from the activist), and the activist's review increases, then the probability of consuming that combination increases.

Table C1: IV Logit Estimates - Expert and Manufacturer Fixed Effects

	(1)	(2)	(3)	(4)	(5)
Doctor's Review	0.46		-2.03^{***}	-4.45^{*}	-3.31^{***}
	(0.31)		(0.64)	(2.62)	(0.87)
Activist's Review		2.14	2.68^{*}	6.51^{**}	3.17^{***}
		(1.39)	(1.61)	(2.72)	(0.77)
Prob of No Ailment	1.53^{***}	0.90	1.15^{*}	0.14	1.59^{***}
	(0.26)	(0.56)	(0.65)	(0.88)	(0.43)
Prob of Non-decreasing CD4	1.07^{***}	-0.20	0.08	0.09	-0.63
	(0.39)	(0.81)	(0.94)	(1.37)	(1.14)
Online × Doctor's Review					-4.79^{***}
					(1.59)
Online × Activist's Review					4.42^{***}
					(1.31)
Doctor Fixed Effects	Yes	No	Yes	No	No
Activist Fixed Effects	No	Yes	Yes	No	No
Manufacturer Fixed Effects	No	No	No	Yes	No
No. of Individuals	13,472	13,472	13,472	13,472	13,472
Combo-time dyads	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Doctor and activist fixed effects control for the respective expert's ID. Manufacturer fixed effects control for manufacturer ID for the combo. Online is a dummy which is 1 for years 2007 and 2008, when the drug guides were also published online. The total number of observations used for the estimation is 1,086, which are constructed using data on 13,472 individuals. Objective qualities include the probability of no ailment and probability of non-decreasing CD4 count of the combo.

Table C2: IV Logit Estimates - Pooling Reviews

	(1)	(2)	(3)	(4)
Total	0.42***	0.55***		
	(0.14)	(0.15)		
Max			0.62***	0.81***
			(0.21)	(0.22)
Objective Qualities	No	Yes	No	Yes
No. of Individuals	13,472	13,472	13,472	13,472
Combo-time dyads	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Doctor's review and Activist's review have been pooled together and instrumented using the average probability of no ailment and average probability of non-decreasing CD4 count of rival combos. Columns (1) and (2) show results for the specification in which the two experts' reviews have been pooled by adding up the reviews, while columns (3) and (4) show results for the specification in which the maximum of the two experts' reviews is used as the measure of combo review. The total number of observations used for the estimation is 1,086, which are constructed using data on 13,472 individuals. Objective qualities include the probability of no ailment and probability of non-decreasing CD4 count of the combo.

Table C3: Objective Qualities with Individual and Time Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Doctor's Review	1.64***	1.49***			−0.79	−2.60***	
	(0.38)	(0.34)			(0.77)	(1.00)	
Activist's Review			2.01***	1.08***	2.63***	4.26***	
			(0.36)	(0.21)	(0.71)	(0.93)	
Agree × Review							1.60***
							(0.46)
Disagree × Activist's Review							3.00***
							(0.56)
Disagree × Doctor's Review							−1.10
							(1.01)
Agree							0.49
							(2.47)
Objective Qualities	No	Yes	No	Yes	No	Yes	Yes
Nobs.	1,086	1,086	1,086	1,086	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Objective qualities include the probability of no ailment and the probability of non-decreasing CD4 count of the combo, which are constructed by controlling for individual and time fixed effects when predicting the probabilities using individual-level data from MACS. Doctor's and Activist's review have been instrumented using the average probability of no ailment and average probability of non-decreasing CD4 count of rival combos. The variable 'Agree' is a dummy which is 1 if both experts give the same rating to a combo. The variable 'Disagree' is a dummy which is 1 if each expert gives a different rating to a combo.

Table C4: Disagreements

	(1)	(2)	(3)
Doctor's Review	-2.41^{***}		
	(0.75)		
Activist's Review	4.69^{***}		
	(0.81)		
Agree × Review		2.07^{***}	2.05^{***}
		(0.56)	(0.61)
Disagree × Activist's Review		2.08^{***}	
		(0.63)	
Disagree × Doctor's Review		-0.40	
		(1.08)	
Agree (% High)		-0.18	-1.09
		(0.47)	(0.70)
Positive Difference × Doctor			2.61
			(1.94)
Negative Difference × Doctor			-3.41^{*}
			(1.92)
Positive Difference × Activist			-0.99
			(1.50)
Negative Difference × Activist			8.28^{***}
			(3.14)
Objective Qualities	Yes	Yes	Yes
No. of Individuals	13,472	13,472	13,472
Combo-time dyads	1,086	1,086	1,086

Notes: *, **, *** denote p-value < 0.10, 0.05, and 0.01, respectively. Standard errors are given in parentheses. Doctor's and Activist's review have been instrumented using the average probability of no ailment and average probability of non-decreasing CD4 count of rival combos. The total number of observations used for the estimation is 1,086, which are constructed using data on 13,472 individuals. The variable 'Agree' is a dummy which is 1 if both experts give the same rating to a combo. The variable 'Disagree' is a dummy which is 1 if each expert gives a different rating to a combo. Finally, the variable 'Positive Difference' is a dummy which is 1 if the doctor's review is lower than the activist's review, while the variable 'Negative Difference' is a dummy which is 1 if the doctor's review is higher than the doctor's review. Objective qualities include the probability of no ailment and probability of non-decreasing CD4 count of the combo.

Appendix D Theoretical Model

Let drug $d's$ unobserved quality $\theta \in \mathbb{R}^2$ have two dimensions: drug effectiveness $h \in \mathbb{R}$ and how well it represses side effects $s \in \mathbb{R}$. The utility an individual gets from consuming drug d, conditional on all observed objective qualities $\mathbf{X}$ is given by:[60][61]

$$u_d(h, s|\mathbf{X}) = \alpha h + \beta s + \gamma(\text{AIDS} \cdot h), \tag{17}$$

where AIDS is a dummy for whether the individual is suffering from AIDS and $\alpha > 0, \beta > 0, \gamma > 0$.[62] We assume that the individual does not observe θ and uses reviews from doctors and activists as signals of the true unobserved quality. Let us assume that h and s can take one of two values, $h \in \{h^H, h^L\}$ and $s \in \{s^H, s^L\}$, where H denotes high quality and L denotes low quality, and doctor and activist comments can either be high or low, i.e., $D, A \in \{0, 1\}$ where 0 denotes low comment and 1 denotes high comment. Then, we can define probabilities for observing quality $t \in \{H, L\}$, conditional on doctor and activist comments as:

$$P_d(h = h^H|R = r) = p_R^r, \tag{18}$$

$$P_d(s = s^H|R = r) = q_R^r, \tag{19}$$

$R \in \{D, A\}, r \in \{0, 1\}$. Moreover, we assume that conditional on both observed and unobserved drug characteristics doctor's and activist's comments are independent. Given this setup, we can now derive theoretical predictions that can be tested empirically.

Proposition 1. *When the doctor and activist agree, individuals choose the drug that gets a high comment, provided that comments are informative.*

Proof. Individuals will choose the drug that gives them the highest expected utility. Suppose drug k gets high comments from both experts, while drug j gets low comments from both experts. An individual, regardless of his AIDS status, will choose drug k over j when

$$E[u_k(h, s|\mathbf{X}, D, A)] > E[u_j(h, s|\mathbf{X}, D, A)] \tag{20}$$

[60] We write our theoretical model after conditioning on all observed characteristics of the drug to understand how drug demand relates to unobserved qualities of the drug and expert comments. We categorize the drug's unobserved qualities into two dimensions, effectiveness and side effects, which may be correlated with observed measures of drug effectiveness (probability of non-decreasing CD4 count) and side effects (probability of no ailment).

[61] We have suppressed the individual subscript i to simplify notation.

[62] This restriction on preference parameters assumes that individuals prefer drugs that are more effective and have less side effects, and that these are state-dependent preferences for effectiveness, in that individuals with AIDS prefer more effective drugs more (Papageorge, 2016).

$$\Leftrightarrow (\alpha+\gamma\text{AIDS})h^H(p_D^1-p_D^0+p_A^1-p_A^0)+\beta s^H(q_D^1-q_D^0+q_A^1-q_A^0)> \quad (21)$$
$$(\alpha+\gamma\text{AIDS})h^L(p_D^1-p_D^0+p_A^1-p_A^0)+\beta s^L(q_D^1-q_D^0+q_A^1-q_A^0).$$

The last inequality is always true when $p_D^1 > p_D^0$, $p_A^1 > p_A^0$, $q_D^1 > q_D^0$ and $q_A^1 > q_A^0$. In words, both experts are more likely to give a higher rating to drugs that are better on both dimensions. □

Proposition 2. *When the doctor and activist disagree, we will observe differences in responses to conflicts depending on health status if and only if*

1. *individuals without AIDS value low side effects more than high effectiveness ($\beta > \alpha$),*
2. *individuals with AIDS value high effectiveness more than low side effects ($\beta < (\alpha+\gamma)$),*
3. *the activist puts more weight on side effects than the doctor ($q_D^0 > q_D^1$ and $q_A^1 > q_A^0$),*
4. *the relative probability that the activist gives a high rating to a drug that has high h is lower than the relative probability of the doctor doing the same ($(p_A^1-p_A^0) < (p_D^1-p_D^0)$).*

Proof. Suppose the doctor gives a low comment to drug k and a high comment to drug j, while the activist gives a high comment to drug k and a low comment to drug j. Then, an individual without AIDS will choose drug k when

$$\implies \alpha h^H(p_D^0-p_D^1+p_A^1-p_A^0)+\beta s^H(q_D^0-q_D^1+q_A^1-q_A^0)> \quad (22)$$
$$\alpha h^L(p_D^0-p_D^1+p_A^1-p_A^0)+\beta s^L(q_D^0-q_D^1+q_A^1-q_A^0)$$

Given that $h^H > h^L$ and $s^H > s^L$, under these assumptions, equation (22) will be satisfied if $(p_A^1-p_A^0) > (p_D^1-p_D^0)$. If $(p_A^1-p_A^0) < (p_D^1-p_D^0)$, then for equation (22) to be satisfied, $\beta > \alpha$, so that the expected marginal utility from higher s is greater than the expected marginal utility from higher h.

An individual with AIDS = 1 will choose drug j over drug k if

$$(\alpha+\gamma)h^H(p_D^0-p_D^1+p_A^1-p_A^0)+\beta s^H(q_D^0-q_D^1+q_A^1-q_A^0)< \quad (23)$$
$$(\alpha+\gamma)h^L(p_D^0-p_D^1+p_A^1-p_A^0)+\beta s^L(q_D^0-q_D^1+q_A^1-q_A^0)$$

It is easy to see that equation (23) will be satisfied when $(p_A^1-p_A^0) < (p_D^1-p_D^0)$, $\alpha,\beta,\gamma > 0$, and $\beta < (\alpha+\gamma)$, so that the expected marginal utility from higher s is lower than the expected marginal utility from higher h.

Now lets suppose $(p_A^1 - p_A^0) < (p_D^1 - p_D^0)$, $q_D^0 > q_D^1$, $q_A^1 > q_A^0$ and that for people without AIDS $\beta > \alpha$ while for people with AIDS $\beta < (\alpha + \gamma)$.

An individual without AIDS will choose drug k (for which the activist's comment is higher than the doctor's) when equation (22) is satisfied. Given our assumption that $h^H > h^L$ and $s^H > s^L$ and the above conditions, we can see that since $\beta > \alpha$, the LHS of the equation (22) is greater than the RHS. Individuals with AIDS, however, will choose drug j (for which the doctor's comment is higher than the activist's) when equation (23) is satisfied. Given that we assume that $\alpha, \beta, \gamma > 0$, and following the above conditions, we can see that equation (23) is satisfied.

□

Appendix E State of the Market

Table E1: New Drugs

Date of Entry	Name	Market Share at time of entry
April, 1997	Videx	4.40%
April, 1999	Efavirenz	5.84%
April, 1999	Ziagen	0.76%
October, 2000	Kaletra	0.28%
October, 2001	Viread	0.62%
April, 2002	Trizivir	1.67%
October, 2003	Reyataz	0.71%
October, 2003	Emtriva	0.71%
April, 2005	Lexiva	0.56%
April, 2005	Truvada	6.60%
April, 2005	Epzicom	1.88%
October, 2006	Prezista	0.37%
April, 2008	Atripla	19.0%

Notes: The table lists all new drugs that enter the HIV drug market during our period of analysis (1997-2008), along with the market share of those drugs at the time of entry. Market share is calculated at the combo level; i.e. for each of the drugs listed, the market share for drug i is the combined market share of all combinations that include drug i.

Table E2: Unique Combinations per Visit

Visit	Number of Unique Combos
April, 1997	81
October, 1997	32
April, 1998	68
October, 1998	30
April, 1999	73
October, 1999	43
April, 2000	52
October, 2000	44
April, 2001	38
October, 2001	41
April, 2002	57
October, 2002	51
April, 2003	54
October, 2003	86
April, 2004	74
October, 2004	50
April, 2005	91
October, 2005	65
April, 2006	58
October, 2006	37
April, 2007	29
October, 2007	24
April, 2008	39
October, 2008	31

Notes: The table lists the number of unique combinations observed in the dataset in each visit.

Table E3: Objective Qualities and Reviews of New Entrants and Rivals at Time of Entry

Reviews				
	Doctor		Activist	
	Own	Rival	Own	Rival
Videx	2.42	2.37	2.50	2.42
Efavirenz	2.78	2.28	2.16	2.15
Ziagen	2.92	2.32	2.00	2.16
Kaletra	2.33	1.97	2.67	2.41
Viread	2.83	2.52	2.33	2.38
Trizivir	3.00	2.26	2.17	2.09
Reyataz	2.20	2.36	2.00	2.11
Emtriva	2.17	2.36	2.17	2.11
Lexiva	2.33	2.20	2.33	1.91
Truvada	2.70	2.16	2.63	1.85
Epzicom	2.71	2.17	2.12	1.90
Prezista	2.33	1.91	2.33	1.80
Atripla	2.00	2.04	3.00	2.07
Objective Qualities				
	Non-Dec. CD4		No Ailment	
	Own	Rival	Own	Rival
Videx	0.54	0.57	0.56	0.63
Efavirenz	0.55	0.56	0.65	0.55
Ziagen	0.61	0.56	0.61	0.56
Kaletra	0.55	0.49	0.73	0.55
Viread	0.54	0.54	0.53	0.59
Trizivir	0.54	0.53	0.56	0.61
Reyataz	0.69	0.55	0.71	0.61
Emtriva	0.52	0.56	0.86	0.60
Lexiva	0.76	0.55	0.74	0.63
Truvada	0.62	0.55	0.70	0.62
Epzicom	0.64	0.55	0.61	0.63
Prezista	0.93	0.56	0.90	0.63
Atripla	0.61	0.60	0.81	0.60

Notes: The table reports the average reviews for each expert and objective qualities (probability of non-decreasing CD4 count and probability of no ailment) for the new entrants and their rivals at the time of entry. For any new entrant drug i, the columns labeled 'Own' report the average reviews (or objective quality measure) for all combinations that contain drug i. The columns labeled 'Rival' report the average review (or objective quality measure) for all combos other than the combos that contain drug i.