

Prescription Drug Advertising and Drug Utilization: The Role of Medicare Part D

Abby Alpert^{† a,b}, Darius Lakdawalla^{c,b}, Neeraj Sood^{c,b}

^a University of Pennsylvania; 3641 Locust Walk, Colonial Penn Center, Philadelphia, PA 19104, United States

^b NBER, National Bureau of Economic Research; 1050 Massachusetts Ave., Cambridge, MA 02138, United States

^c University of Southern California; 635 Downey Way, Verna & Peter Dauterive Hall (VPD), Los Angeles, CA 90089, United States

[†] Corresponding Author: alpertab@wharton.upenn.edu

Email Addresses: alpertab@wharton.upenn.edu; dlakdawa@usc.edu; nsood@usc.edu

ABSTRACT

This paper examines how direct-to-consumer advertising (DTCA) for prescription drugs influences utilization by exploiting a large and plausibly exogenous shock to DTCA driven by the introduction of Medicare Part D. Part D led to larger increases in advertising in geographic areas with higher concentrations of Medicare beneficiaries. We examine the impact of this differential increase in advertising on non-elderly individuals to isolate advertising effects from the direct effects of Part D. We find that exposure to advertising led to large increases in treatment initiation and improved medication adherence. Advertising also had sizeable positive spillover effects on non-advertised generic drugs. Our results imply significant spillovers from Medicare Part D on the under-65 population and an important role for non-price factors in influencing prescription drug utilization. *Keywords: direct-to-consumer advertising, prescription drugs, medication adherence, Medicare Part D; JEL Codes: H51, I10, I18*

1. Introduction

Direct-to-consumer advertising (DTCA) of prescription drugs is a salient and controversial issue in the U.S. Spending on this form of advertising has increased dramatically in the last few decades from \$150 million in 1993 to \$6 billion in 2016 (Dave, 2013; Schwartz and Woloshin, 2019). This rise was precipitated by a 1997 FDA policy change that relaxed restrictions on DTCA.¹ Most DTCA occurs on television, where pharmaceuticals represented the third highest category of advertising expenditures in 2014 (behind automotive and fast food restaurant advertising).² Nielsen estimates that an average of 80 pharmaceutical ads air every hour on American television.³ Since Americans aged 50+ watch an average of more than 40 hours of live television per week (Nielsen, 2014), pharmaceutical advertising may have especially large effects on the behavior of older individuals with high rates of prescription drug use. Indeed, Figure 1 shows that the dramatic rise in advertising that occurred over the last two decades has coincided with a similarly striking increase in spending on prescription drugs.

The rise of DTCA has generated considerable policy debate about its effects on patient welfare. In 2015, the American Medical Association (AMA), the physicians' professional association, called for a ban on DTCA in the U.S. Most other countries (with the exception of New Zealand) already ban this type of advertising. The AMA cited "concerns that a growing proliferation of ads is driving demand for expensive treatments despite the clinical effectiveness of less costly alternatives."⁴ On the other side, proponents of DTCA argue that advertising is informative as it educates patients about available treatments and encourages them to seek care

¹ Prior to 1997, ads were required to include essentially all of the information on the product label (which is unlikely to fit in a 30-second television or radio spot), but after 1997 only the *major* risks and benefits needed to be included.

² See Nielsen, 2014: <http://www.nielsen.com/us/en/insights/news/2015/tops-of-2014-advertising.html>

³ Nielsen estimate reported in FiercePharma "Top 10 DTC Pharma Advertisers – H1 2013" available at: <http://www.fiercepharma.com/special-reports/top-10-dtc-pharma-advertisers-h1-2013>

⁴ For the AMA's position, see: <https://www.ama-assn.org/press-center/press-releases/ama-calls-ban-dtc-ads-prescription-drugs-and-medical-devices>

for underdiagnosed conditions.⁵ Advertisements may also serve to remind patients to take their existing medications, promoting better drug adherence (Donohue et al., 2004; Wosinska, 2005).

There is a lack of consensus on whether DTCA serves primarily to inform or persuade (Berndt, 2005), which matters for assessing its value to patients. This distinction hinges partly on the extent to which DTCA impacts drug utilization and on the mechanisms underlying advertising's impacts, such as whether the effects of DTCA stem from changes in the initiation of therapy versus changes in adherence and whether the marginal consumer seeking treatment due to advertising is appropriate for therapy. There is also limited evidence on whether DTCA causes substitution towards expensive treatments and away from cheaper alternatives— a question at the heart of the policy debate. Answering these questions and identifying DTCA's causal effects has been challenging empirically, since demand factors often influence both the amount of advertising and the timing of advertisements. Some studies have tried to address these endogeneity concerns with instrumental variable strategies, though it is difficult to find appropriate instruments given the close relationship between demand and advertising decisions.

We address these challenges by introducing a new quasi-experimental approach to estimating the effects of DTCA. We exploit a large shock to DTCA driven by the introduction of Medicare Part D in 2006. Our instrumental variable strategy exploits variation across geographic areas in the share of the population that is covered by Medicare (ages 65+) to predict changes in advertising exposure across areas. We show that there was a large relative increase in advertising exposure immediately following the introduction of Part D in geographic areas with a high share of elderly compared to areas with a low elderly share. Prior to Part D, both the levels and trends in advertising exposure across high and low elderly share areas were nearly identical.

⁵ See for example, PhRMA's position: <https://www.reuters.com/article/us-pharmaceuticals-advertising/u-s-doctor-group-calls-for-ban-on-drug-advertising-to-consumers-idUSKCN0T62WT20151117>

Since advertising cannot be perfectly targeted to the elderly, we use the sudden differential increase in advertising exposure for non-elderly that live in elderly-dominated areas to study the effects of advertising on drug use. This strategy hinges on the observation that non-elderly individuals are exposed to the increase in DTCA but do not receive Part D insurance coverage, which may independently impact drug utilization.

This paper makes several contributions. First, we exploit a major policy change to identify the effects of advertising on drug utilization. The use of policy shocks as natural experiments has been scarce in the existing advertising literature and the shock to advertising due to Part D is unusually large in terms of its size and breadth relative to most advertising changes that have been studied. Second, this policy shock provides an ideal setting for estimating a broad array of behavioral responses to advertising on both the extensive and intensive margins, including drug initiation and adherence. Prior studies on the revenue consequences of advertising have largely focused on overall utilization and spending and there is limited prior evidence on drug adherence. Third, we use novel pharmaceutical advertising data from Nielsen “ratings,” which we observe by age group. While almost all prior DTCA studies use advertising spending or the number of ads to quantify advertising, Nielsen ratings are a more direct measure of actual advertising exposure. This measure is more often used outside of the DTCA literature to measure exposure to other types of television programming (e.g., Kanazawa and Funk, 2001; Kearney and Levine, 2015). Finally, we quantify spillover effects of Part D on the non-elderly population. Numerous studies have examined the effects of Part D on the elderly but few have considered the effects on the non-elderly.⁶ One mechanism through which Part D may have an effect on the non-elderly is through advertising and we find strong evidence of these spillovers.

⁶ Prior studies on the spillover effects of Part D have examined pharmaceutical R&D investments (Blume-Kohout and Sood, 2013) and negotiated drug prices (Duggan and Scott Morton, 2010; Lakdawalla and Yin, 2015).

We find that drug utilization is highly responsive to advertising exposure. Following Part D, there was a 6 percent increase in the number of prescriptions purchased by the non-elderly in areas with high elderly share, relative to areas with low elderly share. Event study results show that this differential effect coincided precisely with the implementation of Part D in 2006. The event study also confirms that there were no differential pre-trends in utilization across higher and lower elderly share areas. Our results show that a 10 percent increase in advertising views leads to a 5.4 percent increase in total prescriptions filled for advertised drugs, which amounts to an advertising exposure elasticity of 0.54, and an implied advertising expenditure elasticity of 0.23. This is larger than prior estimates in the literature (e.g., Rosenthal et al., 2003; Shapiro 2018, 2022; Sinkinson and Starc, 2019), which may reflect the unusually large shock to advertising driven by Part D and our use of data on advertising views which is a more direct measure of exposure than prior studies' measures of advertising.

About 70% of the total effect of advertising is due to new initiation of prescription drugs, while increased adherence to drug therapy accounts for the remaining 30%. We find that those who initiate treatment due to advertising are on average less compliant, which mitigates some of the health gains from advertising. On the other hand, we find that adherence increases for existing patients in response to advertising. This suggests that advertised drug treatments might be less appropriate for marginal patients than existing ones. We find that DTCA on net does not switch consumers away from cheaper generics to expensive advertised branded drugs; instead, it leads to an overall net *increase* in the use of generics. The likely mechanism is that advertising triggers a visit to a physician for the condition and the physician prescribes a therapeutically similar generic drug. This suggests that the presence of a learned intermediary—the physician—might introduce informational effects, even if advertising is designed to be persuasive.

We use these estimates to conduct a back-of-the-envelope calculation estimating the welfare consequences of increased exposure to advertising. We show that the benefits of increased exposure to advertising can outweigh the costs for cost effective drugs such as statins. Our findings support a strong informative role of advertising for these types of drugs. Finally, this paper also shows that by increasing insurance coverage for one population, Part D had the unintended effect of generating additional demand for individuals outside of the Medicare program. These demand increases were themselves large and economically important.

While Part D triggered a number of changes in the prescription drug marketplace, the bulk of our estimated effects seem linked to advertising in particular. We find limited evidence in favor of alternate causal channels. First, Part D did not differentially reduce drug prices in high elderly share areas, ruling out concurrent price effects that could independently impact drug use. Second, changes in promotion directed to physicians (“detailing”) after Part D appear to be unrelated to elderly share. Finally, in a placebo test estimating the effects of Part D for classes of drugs that *do not* advertise, we find utilization effects that are null or small relative to the effects for classes that do advertise. This provides evidence that the observed changes in utilization are predominantly due to advertising. Accounting for the potentially small effects of other mechanisms, we compute a lower bound on the advertising expenditure elasticity ranging from 0.14 to 0.19; this represents between sixty to eighty percent of our estimated elasticity.

2. Background and Related Literature

2.1. Why Would Medicare Part D Increase Advertising Exposure?

Medicare is a federal program that provides health insurance to the elderly, ages 65 and over, and to qualifying non-elderly disabled individuals. On January 1, 2006, Medicare expanded to include coverage of outpatient prescription drugs through the introduction of Part D,

representing the largest expansion of the Medicare program since its inception. Part D substantially lowered out-of-pocket costs and increased drug utilization for the elderly (e.g., Lichtenberg and Sun, 2007; Ketcham and Simon, 2008; Yin et al., 2008).

The widespread changes brought about by Part D significantly altered pharmaceutical firms' incentives to advertise. As shown in earlier theoretical work (Lakdawalla, Sood, and Gu, 2013), insurance expansions such as Part D can increase the return to advertising through two mechanisms. First, more profitable markets generate greater returns to capturing new consumers, and in turn stimulate more intense advertising effort. Thus, the returns to advertising are higher when there are more insured consumers in the market, because insured consumers face lower out-of-pocket costs that induce greater spending. Second, insurance coverage might alter the responsiveness of consumers to advertising. Intuitively, an undecided consumer might be more likely to try a new drug after seeing an ad if the cost of trying the drug is lower.

Given this result, we would expect drug advertising to increase more after Part D in geographic areas with a higher share of elderly individuals, where there was a greater expansion in insurance coverage. We will show that this prediction is borne out in the data. Previous research (Lakdawalla, Sood, and Gu, 2013) found that Part D led to a relative increase in national advertising for the types of drugs differentially used by Medicare beneficiaries. We build on this previous work by introducing a new strategy exploiting geographic variation in advertising exposure across markets and characterizing the causal utilization effects of advertising using administrative claims data.

2.2. Previous Literature on Advertising Effects

Our paper contributes to a large literature on the impacts of DTCA on drug use (see Dave, 2013 for a thorough survey). The majority of studies in this literature find positive demand

effects of advertising, although the estimated elasticities vary widely. While studies consistently find evidence of significant market expansion effects from advertising (e.g., Berndt et al., 1995; Rosenthal et al., 2003; Iizuka and Jin, 2005; Bradford et al., 2006), evidence of market stealing—gaining market share from competitors—is mixed. Some studies find no effect, and others find small but statistically significant effects (e.g. Wosinska, 2002; Dave, 2013).

A persistent challenge for this literature has been in identifying a source of variation in advertising that is orthogonal to demand factors.⁷ Estimates could be biased upwards if firms target advertising to markets (or time periods) where demand for the drug is already high or biased downwards if firms aim to stimulate demand where it would otherwise be low.⁸ Our study overcomes this problem by using a natural experiment—the introduction of Part D—to study the effects of DTCA on drug utilization among those unaffected by the insurance expansion. To our knowledge, there are only two other studies that provide natural experiment evidence on the effects of DTCA on utilization. Most similar to our study, Sinkinson and Starc (2019) exploit changes in DTCA due to political election advertising (which temporarily displace DTCA) to examine the effects of advertising on firm revenue for statins. The estimated elasticities in our study are larger. This may be partially explained by differences in identification strategies, with Sinkinson and Starc (2019) exploiting temporary reductions in advertising and our study exploiting a large and permanent increase in advertising. Given the long-lasting effects of advertising, temporary reductions in advertising could have more muted effects on drug use than

⁷ Most previous studies of DTCA have had to rely on cross-sectional or time-series variation in advertising to identify the effect on drug utilization. Studies that consider the endogeneity concern have instrumented for DTCA using variables such as the age of the drug, time until patent expiration, advertising expenditures by the same company in an unrelated drug class, and national advertising costs.

⁸ For example, there could be an upward bias if DTCA targets high prevalence diseases that would naturally have higher sales without advertising. There could be a downward bias if advertising targets under-diagnosed or under-treated diseases. Also, firms may increase advertising when a competing product enters the market. This could also create a downward bias, since, in the absence of advertising, the demand for the older product may have declined as patients substituted towards the new product. It is unclear which of these effects would dominate.

advertising increases.⁹ In another study, Shapiro (2018) compares differences in advertising expenditures at television market boundaries to estimate effects of DTCA on antidepressant use. This paper exploits idiosyncratic misalignment in the targeting of advertising at market borders, another distinct source of geographic variation with potentially different implications. Empirical approaches using different shocks to DTCA will find different local average treatment effects (Imbens and Angrist, 1994), but our policy experiment bears directly on market-wide and permanent changes to DTCA which speaks to the debate on regulating DTCA.

Our study offers several contributions to the literature. First, the size, breadth, and permanence of the DTCA shock driven by Part D's introduction is unusually large relative to other advertising changes captured in previous studies. Second, we use novel data that measure actual exposure to advertising using Nielsen ratings rather than indirect measures that are used in most DTCA studies such as advertising spending or number of ads.¹⁰ Third, we decompose the total utilization effects of advertising into distinct channels of behavioral response including drug initiation and adherence, which is important for assessing welfare implications of DTCA. Since much of the pharmaceutical advertising literature has focused on the revenue consequences of advertising, less is known about the behavioral mechanisms underlying advertising effects and their welfare implications. Specifically, there is little evidence on how advertising impacts drug adherence; the few existing studies find very small or null effects (Wosinska, 2005; Donohue et al., 2006). Fourth, we estimate the effects of DTCA for a large number of drugs across several conditions. Prior studies typically focus on a single drug class. Given that FDA advertising

⁹ Sales decay may occur more slowly than sales growth, since individuals will already have experience with the drug and may continue to get refills. Memories of prior advertising impressions may also decay slowly.

¹⁰ One exception is Avery et al. (2012), which uses survey data from Simmons National Consumer Survey and Kantar/TNS Media Intelligence to impute individual-level exposure to ads for anti-depressants (a similar method is used in Dubois et al. (2018) for non-drug advertising). In contrast to that paper, which studies self-reported anti-depressant use in the past 12 months, we have administrative pharmacy claims that enable us to construct comparatively rich measures of utilization. We also observe actual exposure to advertisements at the market level.

regulations tend to consider all types of drugs uniformly, our estimates are more generalizable for such policy considerations. Finally, our results are relevant for understanding the broader consequences of insurance expansions on demand for unaffected populations.

3. Data Sources

3.1. Advertising Data

We use the Nielsen Ad*Views™ database from 2001-2010 to measure pharmaceutical advertising in local markets. We focus on television advertising, which accounts for more than two-thirds of total DTCA expenditures (Avery et al., 2012). Nielsen collects data on the universe of television commercials shown in 210 “Designated Market Areas” (DMAs) that span the entire U.S. Each DMA is comprised of one or more counties that share the same home-market television stations; thus, households in each DMA view the same television programming and advertising. Nielsen viewing stations located in each DMA record all commercials shown and can identify “national” ads shown in all 210 DMAs and “local” ads shown in a subset of markets. We use data on local ads since there is scope for targeting different amounts of advertising to different markets. Local ads can be shown during network, syndicated, or local television programming. We obtained local advertising data for the top 100 DMAs (86.5% of TV viewers) and the top 200 advertised brand-name prescription drugs, which account for 96% of advertising spending.

Our measure of DTCA exposure is Nielsen gross rating points (GRPs). Rating points are derived from data collected on actual viewership of television commercials for a sample of households in each DMA. Using meters attached to participants’ televisions or paper diaries, Nielsen records who in the household is watching and what they are watching 24 hours a day. “Rating points” are essentially the fraction of the sample that watched a particular commercial.

The data we obtained provide rating points for commercials of each brand-name prescription drug aggregated by DMA, quarter, and for two age groups (ages 2-64 and ages 65+), which is defined as follows:

$$(1) \textit{Rating Points}_{jmat} = \frac{\# \textit{of views}_{jmat}}{\# \textit{of persons}_{mat}} \times 100$$

Where $\textit{Rating Points}_{jmat}$ are computed as the total number of views of commercials for brand-name drug j in market (DMA) m , age-group a , and quarter t divided by the number of individuals in the sample in that group, multiplied by 100. We divide rating points by 100 to interpret this measure as average views per person. Rating points can increase if the number of commercials increases, commercials become better targeted (e.g., primetime vs. late night), or individuals watch more television. Nielsen rating points are the industry standard for measuring television viewership and have the advantage of being a more direct measure of advertising exposure than total advertising expenditures or the number of ads, which have been the predominant measures of advertising in the DTCA literature to date. While in recent years, a variety of alternative methods for watching television have been introduced, such as time shifted (DVR) and Internet viewing, traditional live television remains the dominant medium.¹¹

3.2. Identifying Variation in Advertising Exposure

We construct our instrument for advertising exposure based on the elderly share in each DMA. We compute the share of the population that is 65 and over (i.e., eligible for Medicare) using the 2000 Census. Television advertising is purchased at the DMA-level, since all households within the DMA view the same ads. Therefore, what is relevant to the advertising

¹¹ In 2014, adults ages 50-64 watched on average 43.2 hours of TV per week, of which only 3.8 hours were time-shifted and an additional 1.2 hours were spent watching video on the Internet (Nielsen, 2014). Since most of our study period from 2004-2010 precedes the widespread adoption of time-shifted viewing and the introduction of Internet streaming services (e.g. Netflix, YouTube), we expect that the majority are watching TV live. Nielsen accounts for time-shifted viewing by including recorded programs watched within 7 days of its initial release.

decision is the DMA-level elderly share. We hold the elderly share constant at the DMA's 2000 value so that no identification originates from changes in elderly share. There is substantial heterogeneity in elderly share across markets, ranging from 8% in the Houston DMA to 26% in Fort Myers-Naples DMA (see Table 1).

3.3. Drug Utilization Data

We construct measures of drug utilization using a database of insurance claims from more than 40 large national employers, including many Fortune 500 companies, for 2004-2010.¹² These data were compiled by a prominent health benefits consulting company and cover approximately 18 million person-years during the study period. The claims dataset is described in more detail in previous studies (e.g., Goldman et al., 2004; Goldman et al., 2007). The pharmacy claims include information on all outpatient prescription drug purchases. Limited demographic information is provided, including gender, age, marital status, and the three-digit ZIP code of residence. We restrict our analysis to individuals with full-year insurance coverage and aged 40-60.¹³ This group is closer in age to Medicare eligibility and thus more likely to be using similar types of prescription drugs as Medicare beneficiaries. We only include individuals who live in the top 100 Nielsen DMAs, which represents about 95 percent of pharmacy claims.

Each person in the claims data is assigned to a Nielsen DMA based on their three-digit ZIP code of residence to determine their potential advertising exposure. One limitation of our data is that some three-digit ZIP codes overlap multiple Nielsen DMAs, so it is not possible to assign these individuals to a single DMA with certainty. Instead we assign these individuals the population-weighted average of DMA-level advertising exposure across all of the possible

¹² Data prior to 2004 is not defined in a consistent way with data after 2004, so we cannot use it in our analysis.

¹³ We exclude ages 61-64 out of concern that individuals close in age to Medicare eligibility may change their drug utilization behavior in anticipation of future Part D coverage (Alpert, 2016).

DMAs where they could reside.¹⁴ Consequently, we use the three-digit ZIP code rather than the DMA as the level of analysis, since all individuals residing in a three-digit ZIP code will have the same advertising exposure. As we will show in Section 5.2.4, the results are similar if we restrict the data to the subsample of individuals with a single DMA match.

The main outcomes are total number of prescriptions purchased and total days supplied. For our main analyses, we focus on drugs that treat five chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. There are 50 drugs for these conditions that advertised during our study period (see list of drugs in Appendix Table B.1). We collapse the data to the three-digit ZIP code level by condition and quarter, computing the mean prescriptions purchased and days supplied, to conduct our analyses at the level of variation in advertising exposure. Zeros are included for individuals who were enrolled in health insurance but did not purchase any drugs for the condition.¹⁵ This results in 107,345 ZIPcode-by-condition-by-quarter observations. Since Part D affected advertising incentives for all drugs and due to the possibility of spillovers across drugs treating the same condition, we do not conduct a drug-level analysis and instead perform our analysis at the condition-level.

We also construct a measure of drug adherence. We measure adherence between the individual's first drug claim through their last drug claim for a given condition. Adherence is measured quarterly as the medication possession ratio (MPR), which is a widely used method for measuring medication compliance with claims data (Andrade et al., 2006). The MPR is calculated as the number of days with drug on-hand (days supplied) divided by the number of days in the quarter (see Appendix Section A.2 for further details).

3.4. Descriptive Statistics

¹⁴ About 30 percent of individuals receive this probabilistic measure of advertising exposure.

¹⁵ Each person gets 5 observations for each condition in each quarter. If they do not use a drug for that condition, the observation is zero. We do not condition on having a diagnosis as advertising might affect the rate of diagnosis.

In Appendix Table B.2, we present sample means for the advertising variables by elderly share before and after Part D. We split the 100 DMAs into above vs. below-median elderly share markets. The elderly view more pharmaceutical ads relative to non-elderly viewers. For example, in low elderly share markets in 2005, elderly viewers saw on average 1,184 pharmaceutical ads per year compared to 387 ads for the non-elderly. After Part D, views increased more in high elderly share markets than in low elderly share markets.

4. Empirical Strategy

To understand the impact of DTCA on drug utilization, we exploit quasi-experimental variation in advertising exposure after the introduction of Part D for the non-elderly. We capture the differential change in DTCA exposure across high and low elderly share areas by estimating the following difference-in-differences equation, our first-stage relationship:

$$(2) DTCA_{mct, <65} = \beta(Share65_m \times Post_t) + \gamma_t + \mu_m + \alpha_c + \epsilon_{mct}$$

Where $DTCA_{mct, <65}$ is per capita views (rating points) for non-elderly individuals ages 2-64 in market m in quarter t for ads related to condition c . Our main analysis sample focuses on five chronic conditions. The market m is the three-digit ZIP code.¹⁶ $Share65_m$ is the share of population 65+ in market m in 2000, and $Post_t$ is an indicator that equals 1 in the post-Part D period (2006-2010).¹⁷ Thus $Share65_m \times Post_t$ is our instrument for DTCA exposure for the non-elderly. In some specifications we use an alternative form of this instrument, $HighElderlyShare_m \times Post_t$, where $HighElderlyShare_m$ is an indicator that equals 1 if the

¹⁶ As discussed in Section 3.3, we collapse the data to the 3-digit ZIP code level rather than the DMA-level, since a subset of the sample resides in 3-digit ZIP codes that cross multiple DMAs and cannot be matched to a single DMA. For this subset, we assign them the population-weighted average DMA-level advertising exposure across all DMAs where they might reside. Thus, our constructed advertising exposure measure is constant across individuals within the same 3-digit ZIP code, but not always within a DMA, which necessitates analysis at the 3-digit ZIP code level.

¹⁷ The 65+ share is computed at the DMA level since this is the relevant market from the advertiser's perspective. For individuals whose 3-digit ZIP code cannot be matched to a single DMA, they receive the population weighted average of the 65+ share across all the DMAs where they might reside.

market m has an above-median elderly share.¹⁸ This latter instrument corresponds to the graphical representation of the results presented in the paper. We use the share of the population 65+ as our instrument to reflect advertising incentives.¹⁹ We expect that the greater increase in demand in high elderly share areas after Part D led to a greater increase in advertising.²⁰

Since areas with a high and low share of elderly are demographically different, all of our analyses condition on market fixed effects (μ_m). We include time fixed effects (γ_t) to account for secular time trends and national-level shocks (e.g., patent expirations, treatment guideline changes, etc.) that would affect high and low elderly share areas similarly. We also include condition fixed effects (α_c) to account for differences in utilization and returns to advertising across conditions. Standard errors are clustered at the three-digit ZIP code level (market m) to account for serial correlation within areas and correlation across conditions.

Next, we estimate a reduced form equation comparing changes in drug utilization for the non-elderly across areas with a high elderly share relative to a low elderly share, as follows:

$$(3) Y_{mct, <65} = \theta(Share65_m \times Post_t) + \gamma_t + \mu_m + \alpha_c + \epsilon_{mct}$$

Where $Y_{mct, <65}$ is mean total prescriptions (or other drug utilization measures) for non-elderly individuals in market m in quarter t for advertised drugs that treat condition c . The reduced form model estimates how outcomes for the non-elderly evolve after Part D in high versus low elderly share areas.

¹⁸ In the Appendix, we also define the instrument as quartiles of elderly share. This allows for non-linear effects.

¹⁹ We focus on per capita views as our outcome variable (which is what Nielsen ratings measure) so per capita demand is the relevant measure of the returns to advertising rather than total demand under this specification.

²⁰ Other factors may also affect advertising decisions, which will add noise to the relationship between elderly share and changes in advertising, but this should not affect the validity of elderly share as an instrument.

Finally, we estimate the effect of advertising on prescriptions purchased using 2SLS. The second stage equation is as follows:²¹

$$(4) \quad Y_{mct, <65} = \eta \widehat{DTCA}_{mct, <65} + \gamma_t + \mu_m + \alpha_c + \epsilon_{mct}$$

where $Y_{mct, <65}$ measures drug utilization for the non-elderly. By focusing on the non-elderly population, we can isolate the effects of advertising on drug utilization from the direct effects of Part D on utilization. Interpreting the 2SLS estimates as causal effects of advertising relies on the assumption that the exclusion restriction is satisfied, i.e., that our instrument does not operate through channels other than advertising. The potential threat to identification is that Part D could have spillovers on the non-elderly beyond advertising. There are three main possible channels we consider. First, our 2SLS estimates could be biased upwards if drug prices decreased differentially (e.g., due to changes in bargaining power or benefit design) across high and low elderly share areas after Part D. Second, detailing (which is unobserved in our study) could have increased along with DTCA, leading us to overstate the contribution of DTCA. Finally, if physicians changed their practice styles in response to elderly patients gaining insurance coverage, this could have spillovers on non-elderly patients. We conduct tests for each of these possible channels in Section 5.2.4. We show that there are likely minimal spillover effects from these alternative channels and construct bounds for the advertising estimates allowing for these potential multiple mechanisms.

²¹ It should be noted that the Nielsen data on advertising exposure for the non-elderly is only available to us for the 2-64 and 65+ age groups (which cannot be further disaggregated). In our main analyses, we use advertising exposure from the 2-64 age group while we select on ages 40-60 in the utilization data to reduce noise in our outcome variables since individuals under 40 are less likely to be using these five drug classes. Our advertising elasticity estimates will not be biased by this measurement error if the proportional increase in advertising is the same for the 40-60 age group and the 2-64 age group. As shown previously in Appendix Table B.3, individuals 2-64 and 65+ have about the same proportional change in advertising exposure after Part D in high elderly share vs. low elderly share areas (about a 6% increase), which implies that we would likely find a similar proportional change for the 40-60 age group as well. Although it seems unlikely, if the proportional change for the 40-60 age group is significantly larger than the 2-64 age group (which would mean it is also larger than the 65+ group), then we would overstate the elasticity estimates. Additionally, we will show the robustness of our 2SLS and elasticity estimates to taking the log of DTCA, which should be unaffected by this type of measurement error.

5. Results

Our analysis proceeds in three steps. We first provide evidence that our instrument predicts changes in advertising exposure. Second, we estimate the impact of advertising exposure on total drug utilization for the non-elderly using two-stage least squares. We conduct several robustness tests which show that pre-trends, composition changes and alternative mechanisms do not drive our results. Third, we investigate causal pathways along which advertising operates by decomposing the total utilization effect into intensive and extensive margins including adherence and initiation. We also investigate spillovers on non-advertised drugs. Finally, we discuss the welfare effects of these results.

5.1. First-Stage Effects of Part D on Advertising Exposure

5.1.1. Overall Sample of Drugs

We begin by showing graphically that the share of the population that is 65+ in an area is strongly predictive of differential changes in advertising exposure after Part D. Figure 2 plots mean annual views per person (rating points) of ads for the top 200 brand-name pharmaceuticals from 2001-2010, comparing DMAs with above-median and below-median elderly shares. The figures show views by the non-elderly under age 65. Prior to 2006, both the levels and trends in advertising exposure for the non-elderly are nearly identical across geographic areas. That is, a non-elderly person would view the same number of pharmaceutical ads whether they lived in a market with a high or low concentration of elderly. However, after Part D began in 2006, advertising exposure increases sharply for non-elderly living in areas with a high elderly share relative to those living in areas with a low elderly share.²² This difference persists through the

²² The small increase in advertising prior to Part D's implementation may be due to anticipatory responses since Part D was signed into law in December 2003. Firms may have increased advertising in advance given the long-lasting effects of advertising. We find that advertising begins to increase in 2005q1 with the largest increase in 2005q4.

end of the study period.²³ We note that since we are following a panel of brand-name drugs, there is a secular downward trend in overall advertising over this time period due to the “aging” of these drugs (i.e., some drugs lost patent protection over the study period).²⁴ Since off-patent drugs typically do not advertise (Dave, 2013), patent expirations reduce advertising expenditures. In Appendix Figure B.2, we exclude drugs that lost patent protection during the study period. We find an upward trend in overall advertising, but still find a similar divergence in trends after Part D’s introduction. This verifies that the advertising effects are not driven by patent expirations.

Computing the magnitudes of these effects, we find that Part D generated an additional 25 ads viewed per year in areas with high elderly share relative to low elderly share areas, or about one additional ad every other week.²⁵ This represents a 6 percent increase relative to the baseline mean. For the elderly, the effect of Part D on the number of ads viewed is much larger, since Medicare beneficiaries are likely the intended target for these ads. We find that Part D generated an additional 72 ads viewed per year, or an additional ad every 5 days (5.7 percent increase). These results confirm that the introduction of Part D is associated with a large relative increase in advertising exposure for the elderly in high elderly share areas and that there are substantial spillover effects for the non-elderly.

5.1.2. Chronic Drugs

We also assess the predictive power of the instrument for our primary analysis sample of chronic drugs for five conditions (depression, diabetes, hyperlipidemia, hypertension,

²³ The patterns in advertising exposure are similar for the elderly (see Appendix Figure B.1). Prior to Part D, the trends are parallel but there is less advertising exposure in high elderly share areas, perhaps due to the lower income in these areas. After Part D, the pattern flips and advertising immediately increases in high elderly share areas.

²⁴ This pattern mimics the decline in national advertising expenditures shown in Figure 1. There were 4 major patent expirations that occurred around 2006 among the top 200 drugs (Pravachol, Wellbutrin XL, Zocor, and Zoloft) and a wave of patent expirations in the late 2000s which has been termed the “patent cliff.”

²⁵ We estimate the magnitude of the differential change in advertising exposure using a difference-in-differences model where we interact the DMA elderly share with an indicator for post-Part D and include quarter and DMA fixed effects. In Appendix Table B.3, the coefficients from the binary instrument, imply an additional 25 ads viewed per year ($6.233*4$) for the non-elderly and 72 for the elderly ($18.055*4$).

osteoporosis) that are prevalent among Medicare beneficiaries²⁶ and account for a large share of advertising. We focus on these drugs since they are predominantly used by the elderly and would likely experience the largest increase in advertising from Part D.

We replicate the graphical evidence from above for the chronic drugs sample. In Panel A of Figure 3, we plot advertising exposure for the selected brand-name chronic drugs at the quarterly level for 2004-2010. Trends in non-elderly advertising exposure are nearly identical across high and low elderly share areas prior to Part D and then diverge sharply in 2006.

We estimate the analogous first-stage difference-in-differences regression model shown in Equation 2. Panel A of Table 2 presents the first-stage results using the continuous instrument (Post interacted with Share 65+), while Panel B uses the binary instrument (Post interacted with above-median elderly share) mirroring the graphical evidence. We find a strong positive relationship between the introduction of Part D and differential changes in advertising across areas.²⁷ Panel A shows that a geographic area with a one percentage point higher elderly share experiences an increase in quarterly advertising exposure of 0.064 views per person after Part D (significant at the 1% level). Panel B compares above-median to below-median elderly share areas and shows markedly similar results.²⁸ Advertising exposure for chronic drugs increased by 8.1 percent relative to the baseline mean.

5.2. Second-Stage Effects of Advertising Exposure on Drug Utilization

5.2.1. Baseline Estimates

²⁶ These five conditions are among the most common conditions for Medicare beneficiaries: 58% have hypertension, 45% hyperlipidemia, 28% diabetes, 14% depression, 7% osteoporosis (CMS, 2012).

²⁷ The F-statistics for the binary and continuous instruments are 30.86 and 32.69, respectively.

²⁸ Appendix Table B.4 shows the first stage results for each quartile of elderly share. Advertising exposure after Part D increases monotonically with elderly share, although the pattern appears to be non-linear as there are substantial jumps between the 1st and 2nd quartile and between the 3rd and 4th.

Having shown that Part D had a substantial differential impact on advertising exposure for high elderly share markets, we next analyze how drug utilization responded to this shock to advertising. First, we graph trends in chronic drug prescriptions purchased by the non-elderly for above-median vs. below-median elderly share areas in Panel B of Figure 3. Prior to Part D, drug utilization trends track each other very closely in high and low elderly share areas, but then diverge precisely in 2006 with a relative increase in utilization for non-elderly living in high elderly share markets. This graph mirrors the patterns in advertising exposure and provides visual evidence of strong effects of advertising on utilization.

In Column 2 of Table 2, we estimate the reduced form difference-in-differences specification (Equation 3) using the total number of chronic prescriptions purchased by the non-elderly as the outcome variable. The effect of Part D on non-elderly drug utilization is positive and statistically significant at the 1% level for both the continuous and binary measures of elderly share.²⁹ Drug utilization increased by 4.5 percent relative to the baseline mean.

We also assess the timing of the utilization effect as well as the common trends assumption, by estimating an event study regression. The event study replaces the $Share65_m \times Post_t$ variable in Equation 3 with a full set of quarter dummies interacted with the elderly share measure. Each coefficient estimate gives the difference in prescriptions purchased in high versus low elderly share areas relative to the omitted reference period: quarter 4 of 2005. These coefficients are shown in Figure 4 for both the continuous (Panel A) and binary (Panel B) measures of elderly share. High and low elderly share areas had the same pre-trends in prescriptions purchased, as reflected in the statistically insignificant (and close to zero)

²⁹ When we consider the mean difference between high and low elderly share areas (a difference of about 4 percentage points), the continuous instrument estimate in Panel A implies that moving from a low to high elderly share area would lead to an increase of 0.004 ($.04 \times 0.107$) prescriptions (t-statistic = $0.107/0.023=4.7$). This is similar to the estimate and t-statistic using the binary instrument in Panel B.

coefficients prior to 2006. The coefficients then immediately become positive and statistically significant in quarter 1 of 2006 when Part D begins. The effect persists through the end of the study period. These results show that there was an immediate differential utilization response to Part D. Thus, any alternative explanation for the utilization effect would need to coincide precisely with the introduction of Part D.

In Column 3 of Table 2, we present 2SLS estimates for the effect of advertising on prescriptions (Equation 4). Results are similar in both panels. In Panel A, we find that an additional ad viewed would lead to an increase of 0.017 prescriptions for a chronic condition among non-elderly individuals. In Panel B, when we use the above/below median instrument, an additional ad viewed per person leads to an increase of 0.014. Using the more conservative estimate in Panel B, if an ad were viewed by 71 individuals ($1/0.014$), it would result in one additional prescription being filled. Using the mean for prescriptions and ads viewed, the implied elasticity of demand with respect to advertising views for chronic drugs is 0.54.

5.2.2. Robustness Tests

In this section, we conduct a series of robustness tests which show that our baseline estimates are not sensitive to pre-trends, sample restrictions or sample composition changes. These tests are presented in Table 3. Each cell represents a separate regression where the reported estimate is the coefficient on the instrument (either the continuous or binary version).

The first row in Table 3 repeats the baseline estimates. The second row includes condition x market and condition x quarter fixed effects instead of the separate fixed effects shown in Equation 3. In the third row, we add ZIP code specific linear time trends. In the fourth row, we control for condition-specific linear trends. The results are almost identical to the baseline estimates across these specifications, suggesting that pre-existing trends by condition or

ZIP code are not influencing our results. In the fifth row, we restrict our sample to 2004-2007 to exclude the Great Recession. Workers remaining in our sample during the recession may be observationally different than those in the pre-recession sample. This would be a concern if such composition changes are differentially occurring across high versus low elderly share markets. However, our results are similar when we exclude the recession years. In the sixth row, we exclude the year 2005 to assess the impact of anticipatory effects of Part D. The results are slightly smaller due to a small anticipatory increase in advertising in 2005. In the seventh row, we include only employers that were continuously in the data for all years.³⁰ Differential composition changes due to employer churn could bias our estimates. The results are similar for this subset of firms, although the precision is reduced slightly due to the smaller sample size.

In a final test for sample composition changes, we examine whether demographic characteristics of enrollees change after Part D differentially across high and low elderly share markets (see Appendix Figure B.3). Given the lack of demographic detail in the claims data, we assign each person the average characteristics of their three-digit ZIP code using the 2000 Census. While there are small composition changes throughout the study period, we do not observe any large *differential* changes in the demographic characteristics of the sample around the introduction of Part D. Together, these tests show that sample composition changes and differential pre-trends cannot explain the observed patterns in drug utilization.

We also test the robustness of the results to functional form. In Appendix Table B.5, we include $\text{Log}(DTCA_{mct,<65})$ as the independent variable in the 2SLS equation. Since zeros are common, we show results for $\text{Log}(DTCA_{mct,<65} + 1)$ and $\text{IHS}(DTCA_{mct,<65})$ where $\text{IHS}(\cdot)$ is the inverse hyperbolic sine transformation (Bellemare and Wichman, 2020). Although the

³⁰ Out of the 41 firms that we observe in the claims data, 13 firms are observed in all seven years of the study period. These firms account for about 50 percent of drug claims. On average, we observe firms for five consecutive years.

elasticities are broadly similar to the main specification, they are generally larger than the level measure of DTCA which may reflect their sensitivity to zeros.

5.2.3. Validity of the Instrument

We implement a simple placebo test to test for other shocks to advertising incentives by estimating the effect of Part D on exposure to advertising for contraceptive drugs. Since contraceptives are unlikely to be used by the elderly, their advertising should be unaffected by Part D. In fact, Appendix Figure B.4 shows no differential effect of Part D on advertising for the non-elderly across high and low elderly share markets. We also showed in Section 5.1 that changes in advertising exposure after 2006 were larger for the elderly than the non-elderly, as would be expected if the advertising changes were due to Part D. This evidence provides reassurance that Part D, not another confounder, is driving the differential changes in advertising.

5.2.4 Tests for Alternative Mechanisms

The clear relationship between advertising views and prescription drug utilization by the non-elderly suggest large effects of advertising. However, there are three main alternative explanations (i.e., violations of the exclusion restriction) which should be considered that may lead us to overstate the advertising effects. First, reductions in drug prices (e.g., due to changes in bargaining power or benefit design) could also lead to increased drug utilization by the non-elderly. Second, firms may increase promotional activities after Part D in ways besides DTCA through promotion directed at physicians (“detailing”). Finally, Part D may impact physician practice styles which could spill over to non-elderly patients. If any of the above effects occur, then our estimates may reflect both advertising effects (for which we observed large first-stage changes) and other spillover effects of Part D on the non-elderly. In this section, we consider possible alternative explanations to assess their validity and quantify their contributions. Our

evidence will show that advertising is the dominant, though not sole, mechanism, and we will consider the possible bias of other confounding factors using a bounding exercise.

A. Changes in Prices?

We first examine whether pharmaceutical firms lowered drug prices more in areas with a higher elderly share after Part D, which could lead to a differential increase in drug utilization. Previous studies found that growth in national retail prices declined after Part D due to the increased bargaining power of insurers (Duggan and Scott Morton, 2010; Lakdawalla and Yin, 2015). However, it is not known: whether these retail price reductions were passed along to patients in the form of lower out-of-pocket prices, which is what determines consumer demand; whether out-of-pocket price reductions for the elderly spilled over to the non-elderly; and whether out-of-pocket prices declined more in areas with a high elderly share. Panels A and B of Appendix Figure B.5 plots trends in average out-of-pocket prices and total prices for the non-elderly across high and low elderly share areas.³¹ Appendix Table B.6 and Appendix Figure B.5 (Panels C and D) present the corresponding regression results from difference-in-differences regressions and event studies, respectively.³² In the regression results, we find no evidence of differential changes in out-of-pocket and total prices after 2006 across geographic areas. Additionally, for total prices, the event study coefficients are close to zero and statistically insignificant in every quarter before and after Part D. For out-of-pocket prices, we observe a sawtooth pattern in the event study due to non-linear insurance contracts, however, there are no

³¹ Since we do not observe rebates in the claims data, total prices are gross prices.

³² To address the possibility that price trends could reflect changes in the composition of drugs that individuals take, we restrict Appendix Figure B.5 and Table B.6 to a balanced panel of NDCs observed in every quarter so that product features are constant over time. We restrict the sample to 2004-2007 to maximize the number of NDCs included in the balanced panel. We compute average OOP and total price at the level of NDC and 3-digit ZIP code.

meaningful differences in trends across geographic areas.³³ This suggests that the observed drug utilization patterns following Part D cannot be explained by price changes.

B. Changes in Detailing?

Second, we assess whether pharmaceutical detailing may have also increased differentially across areas after the introduction of Part D. This could bias our findings upwards. There are reasons why we might not expect a sudden increase in detailing as was observed with DTCA. Additional detailing requires an increase in physicians' time allocated to sales calls and hiring sales representatives, so it may adjust more slowly.³⁴

Due to data limitations, we are unable to directly observe detailing data at the DMA level.³⁵ However, we conduct an indirect test for Part D's effect on detailing by exploiting within DMA variation in elderly shares. Direct-to-consumer advertising does not vary within a DMA, because local television station signals reach all households. Detailing, however, is more localized since pharmaceutical sales representatives can target individual practices. In other words, detailing efforts are not constrained by DMA boundaries and should respond to localized demand shocks in areas smaller than the DMA. If detailing responded to Part D, we would expect to observe a larger increase in detailing, and consequently, utilization, in localized areas (e.g. ZIP codes) with a higher share of elderly *within* a DMA. Thus, if our estimated utilization increases are due partly to detailing, we would expect changes in utilization within the DMA to be correlated with local elderly shares.

³³ Because of non-linear insurance contracts, patients face higher cost sharing at the start of the year and lower cost sharing at the end of the year. High elderly share areas seem to have more generous coverage in the earlier part of the year (i.e., lower deductibles) which creates a sawtooth pattern in the event study coefficients. In our main regressions, we include market fixed effects to account for these pre-existing differences in insurance generosity. Notably, the patterns in out-of-pocket prices do not mimic the patterns in utilization shown earlier in Figure 4 in which we observe an immediate discontinuous jump after Part D and persistent differences across areas.

³⁴ The fact that we observe an immediate utilization response after Part D suggests that direct-to-consumer advertising is the main driver of the effect, since detailing is more likely to adjust with a lag.

³⁵ We contacted leading data providers, but this data is unavailable at the sub-national level for this time period.

To test this hypothesis, we estimate the reduced form Equation 3 with elderly shares computed at the three-digit ZIP code level instead of the DMA level (three-digit ZIP codes are the only sub-DMA level of variation we can observe in our data³⁶) and include DMA x quarter fixed effects so that identification originates only from variation in elderly shares *within* DMAs. If within-DMA variation plays no role, then inclusion of the DMA x quarter fixed-effects should wipe out the estimated effects on utilization. This test is meaningful because within-DMA variation in elderly share is significant.³⁷ For example, in the Tampa-St. Petersburg (Sarasota) DMA, the three-digit ZIP code elderly share ranges from 11% to 27%.

The results of this test are presented in Table 4. Column 1 reproduces the baseline reduced form results (computing elderly share at the DMA level), but using only ZIP codes that can be uniquely matched to DMAs. The results are very similar to the main results in Table 2. Column 2 shows these results using elderly share computed at the three-digit ZIP code level instead of the DMA level. The effects of ZIP code-level elderly share on total prescriptions are of a roughly similar magnitude as the effects of DMA-level elderly share.³⁸ Since DMA and ZIP code elderly shares are correlated, the consistency of these results is not surprising. The main test is presented in Column 3, which adds DMA x quarter fixed effects. Here, the effect of the ZIP code-level elderly share goes to zero and becomes statistically insignificant.³⁹ This shows that utilization did not respond to Part D differentially by ZIP code elderly share within DMAs, which suggests that detailing responses were unrelated to elderly share. More generally, this

³⁶ Another measure of elderly share relevant to practices is Hospital Service Areas (HSAs). However, we are unable to collapse the data to this level since HSAs are smaller than 3-digit ZIP codes. Hospital Referral Regions (HRRs), another alternative, are larger than ZIP3s and comparable in size to DMAs, so they would not provide within DMA variation. It is difficult to pinpoint the exact geographic unit that would capture the elderly share of patients for a given practice, but it is likely somewhere between Dartmouth HSAs and HRRs, so ZIP3s represent a middle ground.

³⁷ On average, there are 12 three-digit zipcodes per DMA and a maximum of 40.

³⁸ Column 2 is using a noisier measure of the relevant elderly share variable and, indeed, we find that the estimate is attenuated in Panel A. The Panel B instrument is dichotomous so classical measurement error results do not apply.

³⁹ The standard errors in Columns 2 and 3 are nearly the same such that the absence of an estimated effect in Column 3 is not due to increased noise or a relative lack of intra-DMA variation.

result also provides evidence against other possible confounders of Part D that are correlated with elderly share at the sub-DMA level (e.g., changes in other promotional activities, physician behavior, or pharmacy behavior).

In addition to this test, we note that previous studies have also found limited evidence of a geographical correlation between detailing and DTCA. Shapiro (2018) shows that changes in DTCA at DMA borders are not correlated with changes in detailing. Additionally, Sinkinson and Starc (2019) show that detailing does not adjust during the political election cycle when DTCA is displaced by election ads. Finally, Lakdawalla, Sood, and Gu (2013) find that the introduction of Part D led to a five-times larger increase in DTCA than detailing for drugs with the highest Medicare market share. This study used national level data so it does not show that detailing increases were geographically correlated with DTCA. However, even if advertising responses were perfectly correlated, this suggests that the change in detailing would be small relative to the change in DTCA. Applying these estimates to our study, as much as one-sixth of our estimated DTCA elasticity could be driven by detailing.⁴⁰ Thus, although we cannot definitively rule out potential bias from detailing, which is a limitation of this study, the contribution to our advertising estimates is likely small.

C. Changes in Physician Practice Style?

Finally, we consider the possibility that there were other spillovers of Part D on the non-elderly, unrelated to advertising. For example, one leading possibility is changes to physician practice styles. Part D increased the volume of prescriptions written for the elderly, which may influence prescribing habits, leading physicians to write more prescriptions for their non-elderly

⁴⁰ Under the conservative assumption that DTCA and detailing responses were perfectly correlated geographically and that DTCA and detailing would have the same demand elasticity, this implies that about one-sixth of our estimated advertising elasticity could be attributed to detailing (reducing the elasticity from 0.54 to 0.45). This bias is likely an overestimate, however, since prior studies show minimal geographic correlation.

patients as well. We address this concern empirically with two tests. First, the results in the above section show that drug utilization responses by non-elderly individuals are not related to the zip code elderly share within a DMA. We would expect a correlation if practice style spillovers were driving our results since these responses would be more localized within a DMA.

Second, we conduct a placebo test examining whether there were differential effects of Part D on non-elderly drug utilization for drug classes that *do not* advertise.⁴¹ We would expect that other spillover effects from Part D (e.g., prescribing behavior changes) would affect utilization for all drug classes, whether or not they advertised. Figure 5 compares trends in non-elderly drug utilization across high and low elderly share markets for both advertised and non-advertised drug classes. For this test we draw from the full sample of drug classes, not only the five chronic conditions we analyzed previously. About half of all drug classes had zero advertising during the study period (typically related to the amount of generics in the class).⁴² We show that there was a large differential increase in drug use for advertised drug classes after Part D in high elderly share areas vs. low elderly share areas, but there was no statistically significant differential increase for non-advertised drug classes. This is consistent with a causal role for advertising, since the effect is found only for advertised drug classes.

Comparing the magnitudes in the analogous triple-difference regression in Panel B of Appendix Table B.7, we observe an increase of 0.012 prescriptions for advertised drug classes (an 8.2% increase relative to the baseline mean), which is statistically significant at the 1% level. For non-advertised drug classes, we observe an increase of 0.001 prescriptions (1.3% increase),

⁴¹A key assumption underlying this placebo test is that prescribing of non-advertised drug classes increased for the elderly after Part D. Non-advertised drug classes tend to have a higher share of generic drugs. For example, diuretics, which are nearly all generics, saw no advertising during the study period. On the other hand, anti-hyperlipdemia drugs, where only 26% of claims are for generics, had a positive amount of advertising. Thus, in order for this to be a meaningful placebo test, Part D must have affected both generic and branded drug use for the elderly. This has been shown in several studies (e.g., OIG, 2007; Zhang et al., 2011).

⁴² We restrict the sample to the top 10 advertised and non-advertised drug classes among individuals ages 40-60 in order to ensure that these drugs are relevant for our non-elderly sample.

which is not statistically significant. Using log prescriptions as the outcome in column 2 produces similar results (6.9% vs. 1.5% increase). When we use the continuous instrument in Panel A, the utilization effect for non-advertised classes becomes positive and significant, but the proportional effect for advertised classes is still substantially larger (6.2% vs. 2.3%).⁴³

We also examine the differential change in utilization across high and low elderly areas for each drug class separately in Appendix Figure B.6. The patterns appear strikingly different for advertised and non-advertised drug classes. We observe large differential increases in high elderly share areas for nearly all advertised drug classes. This demonstrates that our effects are not driven by a single class but appear consistently in all advertised classes. Meanwhile, for most non-advertised drug classes, we find no differential effect of Part D across high vs. low elderly share markets. For only 3 of the 10 non-advertised classes (diuretics, calcium channel blockers, and thyroid agents) do we observe any suggestive evidence of positive differential effects beginning after Part D. While the substantially larger utilization effects observed for advertised drug classes relative to non-advertised classes is strong evidence that advertising is the predominant driver of this increase, the small positive effects for some non-advertised drug classes suggest that other mechanisms may also play a role.

To quantify these other spillover effects, we use the placebo test to estimate the extent of the potential bias and provide bounds for our estimates. We leverage the results from the non-advertised drug classes in Appendix Table B.7 to identify other Part D spillovers, since these drug classes do not advertise but may be influenced by the broader effects of Part D in high elderly share areas. If we assume, conservatively, that the entire effect for non-advertised drugs

⁴³ To interpret the estimates in Panel A, we consider the mean difference between high and low elderly share areas. The continuous instrument estimate implies that moving from an average low to high elderly share area would lead to an increase of 0.009 [0.04*(.181+.044)] prescriptions for advertised drug classes. This is a 6.2% (0.009/.146) increase relative to the mean. The percentage increase for non-advertised drug classes is computed similarly.

is due to Part D spillovers unrelated to advertising and that the spillovers are the same magnitude for non-advertised and advertised drugs, then this implies that as much as 16% (using Panel B) to 37% (Panel A) of the utilization effect could be due to other Part D spillovers.⁴⁴ This suggests a lower bound on our main elasticity estimate with respect to advertising views of 0.34 to 0.45; this represents between 60 to 80 percent of our estimated elasticity.

In summary, we do not find a differential decrease in drug prices after Part D. We also find limited evidence that detailing explains utilization patterns, though we cannot observe detailing directly. Detailing is conducted at a more localized level than a DMA, and we find that elderly share is unrelated to utilization outcomes at the sub-DMA level. We further study non-advertised drugs to quantify possible spillovers and find suggestive evidence that these alternative spillovers account for a small share of the advertising effect. We conclude that advertising is the predominant driver of utilization changes.

5.3. Potential Welfare Implications

Given the substantial effect of advertising on total drug utilization, we decompose the utilization effect to quantify the various causal pathways from advertising to utilization and their welfare implications. First, we decompose the utilization effect into the extensive and intensive margins. Second, we examine drug adherence, a special case of the intensive margin effect. Third, we estimate whether there are spillovers of advertising on non-advertised generic and brand drugs in the same drug class.

5.3.1. Extensive vs. Intensive Margin Effects

In Appendix Table B.8, we present 2SLS estimates for extensive and intensive measures of prescription drug use for chronic drugs. The corresponding event studies are shown in

⁴⁴ This is computed by noting that utilization increases for advertised drug classes by 8.2% vs. 1.3% for non-advertised drug classes using Panel B and 6.2% vs. 2.3% using Panel A.

Appendix Figure B.7 and generally demonstrate that there are no pre-trends for the outcomes studied. We estimate three margins of adjustment: extensive margin effects (any prescription drug use), intensive margin effects (number of prescriptions or days supplied conditional on use), and total effects combining both margins. We find positive effects for all outcome variables that are statistically significant at the 5% level in all but two specifications.⁴⁵ We perform a decomposition exercise to compare the relative magnitude of intensive and extensive margin effects (see Appendix A.1), finding that about 70 percent of the total advertising effect is driven by extensive margin responses. Thus, a substantial proportion of the utilization effect comes from increased treatment initiation.

5.3.2. Effects on Drug Adherence

We extend the above analysis of intensive margin effects by examining drug adherence. Advertising may increase adherence if it serves as a reminder to take medication, makes the condition more salient, or increases the perceived benefits of treatment. It may also reduce adherence if it enhances awareness of harmful side effects.

We present the results for drug adherence graphically in Figure 6 for the subset of patients who have filled at least one prescription for the condition. The outcome is the proportion of non-elderly individuals with “high adherence” (defined as $MPR \geq 80\%$). Similar results for the continuous measure of MPR are in Appendix Figure B.8. Adherence is mechanically very high in the first few quarters of the study period, because we start following patients in the quarter of their first observed drug treatment. By construction, most individuals in these early quarters have just initiated treatment. However, this mechanical relationship is

⁴⁵ In Column 5, the extensive margin estimates for any prescription drug use are less precisely estimated; the event studies show a large increase in treatment initiation immediately after Part D is implemented but the effects become noisier in the later years of the study period. This may be partially due to churn in the sample composition of workers and firms during the Great Recession. When we exclude the recession years (Column 6), the estimates are similar in magnitude, but more precisely estimated.

uniform across geographic areas and should not impact our results. Once the adherence measure has stabilized in 2005, we find that the proportion of non-elderly with high adherence is nearly identical across high and low elderly share areas, but then immediately diverges in 2006.

To estimate the magnitudes, we present regression results for the reduced form and 2SLS estimates in Table 5 (see Appendix Table B.10 for continuous measure of MPR). The corresponding event studies are in Appendix Figure B.9.⁴⁶ In the regression results, we present results separately for the full sample, excluding the recession years, and excluding 2004 when adherence is mechanically high. The results are qualitatively similar across samples. In the full sample, Part D led to a 0.4 percentage point increase in the proportion of individuals with high adherence. Restricting the sample to 2005-07, the estimate increases to 1.2 percentage points. These estimates imply an adherence elasticity with respect to advertising ranging from 0.09 to 0.25 depending on the sample. Using the high end estimate, the number of ads viewed would need to increase by 40% in order to increase adherence by 10%.

Next, we present results from an alternative measure of MPR in Appendix Table B.11 which accounts for discontinuation of treatment. In our baseline results (Row 1), we computed the MPR between a person's first and last drug claim. However, in Row 2, we assume that the MPR equals zero after the last observed drug claim.⁴⁷ The 2SLS results using this alternative measure of MPR are slightly larger than the baseline results in most samples. This is suggestive that advertising also reduces treatment discontinuation.

⁴⁶ We present the event studies excluding 2004 when adherence is mechanically high. There is no pre-trend in the year before Part D. We observe an immediate jump in adherence after 2006. As with the other event studies, we find that the estimates are slightly noisier during the recession years when the composition of workers is changing.

⁴⁷ This assumption is most appropriate for chronic conditions that require lifetime treatment. Depression is an exception because guidelines recommend treatment until the symptoms have improved (Donahue et al., 2004). Our baseline MPR is conservative as it does not assume that lifetime treatment is needed.

We also estimate the adherence effect for only existing patients in Rows 3 and 4. In Rows 1 and 2, the changes in adherence represented a combination of effects from both existing and new drug users. The increase in advertising after Part D caused more people to initiate drug treatment, and these new entrants into the sample may have different underlying compliance behavior. To isolate the adherence responses of the existing patients from the new initiators, we replicate the previous results using only the sample of individuals who initiated drug treatment before Part D. When we exclude the new initiators in Rows 3 and 4, the results become larger for both measures of MPR. This suggests that the marginal person who initiates treatment because of advertising is on average less compliant. A back-of-the-envelope calculation suggests that those who initiate treatment due to advertising are about half as likely to have high adherence relative to existing patients.⁴⁸ There are a few possible reasons for this. The marginal person might have a less severe condition, or advertising may attract people who are less attached to treatment (e.g. impulsively trying a drug they saw on TV only to quickly discontinue its use) or less appropriate for treatment. Thus, while increasing adherence among existing users is likely to be welfare enhancing, the welfare effects of new initiation due to advertising are less clear.⁴⁹ Some of the additional drug spending due to advertising could be wasteful since patients initiating a chronic treatment without adhering to it will not experience improved health.

5.3.3. Spillover Effects to Non-Advertised Drugs

⁴⁸ The total effect of advertising on adherence is 0.017, which is a weighted average of the effect for new initiators and existing patients. The adherence effect for existing patients is 0.022. The probability of drug take-up increased by 0.001 after Part D in high elderly share areas from a baseline of 0.06. Using this estimate, combined with the adherence effects, we estimate that the proportion of non-elderly individuals with high adherence after Part D is 0.632 for existing patients and 0.319 for new initiators.

⁴⁹ These findings also suggest that informational frictions are not excluding the most clinically appropriate patients from treatment. New patients drawn in by advertising tend to be less likely to adhere to their medication and this suggests that patients at the margin of consumption are less likely to benefit from the therapy.

Finally, we analyze whether there were spillover effects of advertising on non-advertised drugs within the same drug class to test for market expansion versus substitution effects.⁵⁰ Substitution effects occur when a person who would have taken a competitor drug switches to an advertised drug after viewing an ad for the drug. Market expansion effects occur when a person requests an advertised drug from her doctor, but the doctor then prescribes another therapeutically similar drug instead. Insurance formularies could also induce such spillovers. For example, if brand-name Lipitor is excluded from the formulary, while generic Zocor is covered, advertising for Lipitor could increase generic Zocor use. We test for these types of spillover effects by estimating Equation 3 using as the outcome variable the total prescriptions filled for *non-advertised* drugs belonging to the same therapeutic drug classes as the advertised chronic drugs. We separately estimate the effects for non-advertised generic and brand drugs. This analysis differs from the previous test in Figure 5, because we are now comparing products within a drug class based on whether or not they advertise (as opposed to comparing products across drug classes that advertise or do not advertise).

Figure 7 shows the trends in average prescriptions purchased across high and low elderly share markets for advertised chronic drugs (repeated from Figure 3) and non-advertised chronic drugs (generics and brands separately) in the same classes as the advertised chronic drugs. For non-advertised generic drugs, we see an increase in utilization in high elderly share markets immediately after Part D. This provides strong evidence of a market expansion effect, since DTCA increases generic use and on net does not cause substitution away from lower-cost generics to higher-cost advertised drugs. Interestingly, we find no increase in the use of non-advertised brands after Part D using the binary instrument and only a small increase using the

⁵⁰ While a brand's advertising may also have spillovers on the use of other advertised brands, we cannot identify this effect since Part D impacts all advertised drugs simultaneously. Thus, we focus on identifying market expansion and substitution effects for non-advertised versus advertised drugs, which is one component of the spillover effects.

continuous instrument. There are likely greater spillovers for non-advertised generics than brands because they are cheaper.

The regression analogs are in Table 6 and show that these effects are all statistically significant, except for non-advertised brand drugs. The corresponding event studies are presented in Appendix Figure B.10 and display similar patterns of effects. Consistent with the previous advertising literature, we find large positive spillovers from advertising. We add to this by showing that spillover effects are concentrated among lower-cost generic drugs, which has important welfare implications. From the consumer perspective, spillovers may be welfare enhancing, since this suggests at least some role for informative, rather than market-stealing advertising. In contrast, had we found a complete shift from non-advertised to advertised drugs, this would have represented little welfare gain since advertised drugs may not be significantly superior to non-advertised drugs.⁵¹

5.3.4. Discussion of Potential Welfare Effects

We find that advertising increases the initiation of new prescriptions and adherence to existing ones. However, those who initiate due to advertising have lower adherence. We also find spillovers to non-advertised generic drugs. To estimate the aggregate effects of these responses, we compute a back-of-the-envelope estimate of the lifetime welfare gain or loss from one additional ad viewed by a non-elderly person. This calculation is discussed in detail in Appendix A.3 and briefly summarized here.

⁵¹ We find evidence of advertising spillovers of two types in this paper: 1) spillovers to non-advertised rival drugs (this is most analogous to the prior literature) and 2) spillovers to non-targeted populations (non-elderly). Both represent market expansion effects, however, the welfare consequences for firms are positive in the case of spillovers to the non-elderly because this generates additional revenue for the firm, whereas spillovers have negative consequences when they increase demand for rival drugs. Additionally, since spillovers on the non-targeted population are already factored into the pricing of advertising (advertising is priced per exposure), these spillovers would not affect firm decisions about how much to advertise. On the other hand, spillovers to rivals' demand is not factored into pricing and may reduce advertising.

For illustrative purposes, we focus on the case of statins because they are one of the most commonly used and advertised drug classes during our study period, and there is a credible body of clinical trial evidence on their short and long-term health effects. We combine published estimates on the health gains from statin initiation with our estimates of advertising effects on initiation, adherence, and spillovers to generics to parameterize our model. We find that one additional ad viewed leads to about \$106 in additional lifetime spending per capita and \$126 in lifetime benefits in net present value.⁵² This suggests that the benefits of advertising exceed the costs for this class of drugs and implies that advertising bans could be harmful to consumers in this case. Although advertising leads to some inappropriate use or low adherence by initiators, this is more than offset by the welfare gains from even a small amount of appropriate use since statins are highly cost-effective. Additionally, since generic substitutes are available in this drug class, advertising spillovers to low-cost generics mute the spending increase.

While we focus on a single drug class in this section, our calculation illustrates how the consumer welfare effects from advertising could vary by drug class. First, the welfare gain will be higher for drugs that are more cost-effective, because the welfare gain from advertising accrues from the benefits of increased use of drugs relative to the increased costs. Since statins are one of the most cost effective therapies available, our estimates likely provide an upper bound on the welfare benefits. Second, when low-cost generic drugs are available in the drug class and advertising leads to spillovers to these drugs, the costs are further reduced and the welfare gain will be larger. While critics worry that advertising causes substitution away from generics towards advertised brands (implying a reduction in generic use), we find that generic

⁵² Actual welfare gains could be larger because we do not include the cost savings from reduced medical utilization. We also likely underestimate the benefits of statins, because we assume that low adherence does not produce any health benefits. However, even under these conservative assumptions, our estimate suggests that the benefits of advertising exceed the costs for this class of drugs.

use actually *increases* because of advertising. All else equal, this implies larger consumer benefits of advertising for older classes of drugs, where there is more generic penetration. Finally, the welfare gain will be larger the more well targeted advertising is to initiators who are appropriate for treatment as they will have higher adherence. In our case example, the benefits of advertising were reduced by 38% because new initiators complied with therapy only about half the time. This implies that television advertising targeted to programs that appropriate patients are more likely to watch or online advertising where targeting can be even more precise could have greater consumer welfare benefits than advertising campaigns that reach a more diffuse audience. Overall, the welfare analysis suggests that for certain drug classes (where advertised drugs are cost-effective, generics are available, or a substantial fraction of the population has not initiated treatment) advertising can have significant positive externalities and thus private markets might undersupply advertising. In these markets, increasing public information about the benefits of treatment will likely enhance welfare.

5.4 Comparison of Elasticity Estimates with Prior Literature

We conclude by comparing our estimates to the prior literature. Our main 2SLS regressions show an elasticity of demand with respect to advertising views of 0.54. Although this estimate exceeds prior estimates from the literature, this difference can be largely reconciled by accounting for the different ways that advertising has been measured.⁵³ We measure advertising in terms of views, whereas most other studies use expenditures. We cannot directly estimate an elasticity with respect to expenditures for comparison with prior studies since the

⁵³ Recent pharmaceutical advertising studies have generally found advertising elasticities below 0.14. For example, Sinkinson and Starc (2019) estimate an advertising elasticity of revenue with respect to the quantity of ads ranging from 0.0761 to 0.1395. Shapiro (2018) finds a class level elasticity of sales with respect to advertising expenditures of 0.0496 and Shapiro (2022) finds an elasticity of treatment initiation with respect to GRPs of 0.031. Rosenthal et al. (2003) estimate a class level elasticity with respect to advertising expenditures ranging from 0.096 to 0.114.

Nielsen data does not contain information on advertising expenditures.⁵⁴ However, Sethuraman et al. (2011)—in a meta-analysis of advertising elasticities across consumer products—show that elasticities with respect to views are larger than expenditure elasticities on average, suggesting that firms are operating in a region where increases in advertising dollars lead to less than proportional increases in views.⁵⁵ Specifically, Sethuraman et al. (2011) find that elasticities measured with respect to advertising views are on average 2.3 times larger than elasticities measured with respect to advertising expenditures. If we apply this “conversion factor” to our estimate, this would produce an advertising expenditure elasticity of 0.23 ($0.54/2.3$), which is closer to previous studies. Finally, as discussed in Section 5.2.4, we can further adjust our estimates to account for potential alternative mechanisms, outside of advertising. Accounting for these mechanisms and using the above conversion factor implies a lower bound on our main elasticity estimate with respect to advertising expenditures of 0.14 to 0.19 (or 0.34 to 0.45 with respect to advertising views). These estimates are still slightly larger than prior studies which is potentially due to the comparatively large and permanent shock to advertising that we use as our source of variation as well as our data on advertising views which may more strongly predict utilization than indirect advertising measures used in prior studies.

⁵⁴ Nielsen collects limited data on advertising expenditures at the local level. Expenditures are not available for local commercials shown during network or syndicated programming, which comprise the majority of local commercials. Only commercials shown during “local television” programs (e.g. local news) have expenditure data.

⁵⁵ In theory, the exposure elasticity could be higher or lower than the expenditure elasticity, depending on whether exposures have decreasing or increasing returns to advertising spending. Advertising could have increasing returns due to quantity discounting or decreasing returns due to saturation or competitive responses. For example, the cost per exposure may increase as spending increases, because scarce television time becomes saturated and firms must shift from buying ads for the most well targeted to less well targeted programming. Moreover, as firms spend more on advertising this could trigger stronger competitor advertising responses, which would also increase costs per exposure at the margin and lead to decreasing returns. Sethuraman et al. (2011) offer the following framework for comparing the relative size of the advertising exposure—or Gross Rating Point (GRP)—elasticity versus the advertising expenditure elasticity: “Let a 1% increase in advertising dollars increase GRPs by $v\%$ and sales by $w\%$. Then, by definition, dollar advertising elasticity = w , and GRP advertising elasticity = w/v . It follows that, all else being equal, dollar elasticity is greater than GRP elasticity if $v > 1$, and GRP elasticity is greater than dollar elasticity if $v < 1$.” In other words, we should expect that exposure elasticities are larger than the expenditure elasticities if there are decreasing returns to advertising dollars ($v < 1$).

6. Conclusion

This paper provides a new natural experiment approach for estimating the impact of DTCA on drug utilization and sheds light on the welfare effects of advertising. We find that non-elderly living in high elderly share areas were exposed to relatively more pharmaceutical advertising after Part D. This led to substantial increases in the utilization of chronic drugs. Our results suggest a demand elasticity with respect to advertising views of 0.54, or expenditure elasticity of 0.23.⁵⁶ While we find limited evidence to suggest that this estimate is driven by changes in prices, detailing, or physician prescribing behavior differentially occurring in high elderly share areas, we consider the possible bias that these alternative mechanisms could contribute to the utilization response. We conducted several tests of these mechanisms which are suggestive that the effects are small relative to the effects of advertising. Accounting for these multiple mechanisms suggests a conservative lower bound on the advertising expenditure elasticity ranging from 0.14 to 0.19. These values are close to the main estimates of the paper. However, due to data limitations—particularly the inability to observe detailing data—we cannot rule out all other alternative mechanisms. This is a limitation of our analysis.

Applying our advertising elasticity estimate to the national trend in DTCA expenditures, we estimate that about 31% of the rise in drug spending since 1997 (when the FDA relaxed its advertising restrictions) can be attributed to DTCA.⁵⁷ While one must exercise caution in

⁵⁶ While the literature on prescription drug demand has focused heavily on the importance of prices and insurance status in explaining utilization patterns, we generate estimates of the responsiveness of demand to a non-monetary factor and find economically important effects. Using the range of price elasticities in the literature (-0.2 to -0.6) from Goldman, Joyce, and Zheng (2007) combined with our main results, our estimates imply that a 10 percent increase in advertising produces the same increase in utilization as a 9 to 27 percent reduction in out-of-pocket price.

⁵⁷ To recover this parameter, we begin by estimating the advertising elasticity for total drug spending (including spillover effects on non-advertised drugs) in Appendix Table B.12. Our results imply that a 10 percent increase in advertising views increases total drug spending by about 4 percent (using the binary instrument). We deflate this estimate by a factor of 2.3 to account for the relationship between advertising expenditures and advertising views (Sethuraman et al., 2011), since we only observe time series growth in advertising expenditures. Based on this estimate, we predict that drug spending would increase by 59% in response to the increase in national DTCA

extrapolating our estimates to the national trend, our results are suggestive that DTCA is a significant, though not primary, contributor to the rapid rise in drug spending in the U.S.

Our work also informs the ongoing debate on the welfare impacts of DTCA. The estimates provide a rich picture of the utilization responses to a large and permanent market-wide shock to advertising.⁵⁸ Our results show that the majority of the utilization response to advertising is driven by treatment initiation. Given that the conditions we study are generally considered to be under-treated and under-diagnosed (e.g. Hirschfeld et al., 1997; Majumdar et al., 1999), increased initiation is likely to lead to improved health, representing a welfare gain for consumers. However, we find that individuals who initiate therapy due to advertising have lower rates of treatment compliance, which mitigates some of these health gains. On the other hand, patients already using advertised drugs increase their adherence to treatment in response to advertising, which has positive welfare effects. If advertising served primarily to persuade, rather than to inform, we would observe distortions in use towards the newest, most expensive drugs, irrespective of their quality. Instead, our evidence on spillover effects suggests that a significant share of the increase in utilization comes from lower-cost generic drugs. Summarizing these effects in the case of statins, our back-of-the envelope calculation shows that the lifetime benefits of statin use from viewing one additional ad exceeds the costs, suggesting a net welfare gain from advertising for cost-effective drugs.

Our estimates are also relevant to understanding the broader consequences of insurance expansions and the operation of such spillovers through advertising. Part D led to unintended

expenditures from 1997-2010. Comparing this to the actual increase in national drug spending (193%), DTCA accounts for about 31% of the growth in drug spending since the FDA relaxed restrictions in 1997.

⁵⁸ Since Part D affected advertising incentives for all drugs, it does not serve as an appropriate instrument to test for market stealing between one advertised brand name drug and another.

large and economically important increases in demand among individuals not enrolled in the Medicare program.

Acknowledgements

We are grateful for helpful comments from numerous colleagues, discussants Tina Marsh Dalton, Eric Helland, Sean Nicholson, Julian Reif, Kosali Simon, and seminar and conference participants at the Conference of the American Society of Health Economists, Annual Health Economics Conference, Midwest Health Economics Conference, NBER Summer Institute Health Care Meeting, Southern California Conference in Applied Microeconomics, Stanford University, Harvard University, University of California-Irvine, University of Maryland, University of Southern California, and Wharton. We thank Sarah Dykstra, Laura Gascue, Karina Hermawan, Sushant Joshi, and Chia-Wei Lin for excellent research assistance and programming. This work was supported by NIA P01AG033559, NIA R01 AG062277, and AHRQ R01HS025983.

References

- Alpert, A. 2016. The anticipatory effects of Medicare Part D on drug utilization. *Journal of Health Economics*. 49: 28-45.
- Andrade, S., et al. 2006. Methods for evaluation of medication adherence and persistence using automated databases. *Pharmacoepidemiology and Drug Safety*. 15(8): 565-574.
- Avery, R., Eisenberg, M. and K. Simon. 2012. The impact of direct-to-consumer television and magazine advertising on antidepressant use. *Journal of Health Economics*. 31: 705-718.
- Baigent, C., et al. 2005. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet*. 366(9493): 1267-78.
- Bellemare M. and C. Wichman. 2020. Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics*. 82(1): 50-61.

Berndt, E., Bui, L., Reiley, D., and Urban, G. 1995. Information, marketing, and pricing in the US antiulcer drug market. *American Economic Review*. 85(2): 100-105.

Berndt, E. 2005. To inform or persuade? Direct-to-consumer advertising of prescription drugs. *New England Journal of Medicine*. 352 (4): 325-328

Blume-Kohout, M. and N. Sood. 2013. Market size and innovation: Effects of Medicare Part D on pharmaceutical research and development. *Journal of Public Economics*. 97: 327-336.

Bradford, W., et al. 2006. How direct-to-consumer television advertising for osteoarthritis drugs affect physicians' prescribing behavior. *Health Affairs*. 25(5): 1371-1377.

Centers for Medicare and Medicaid Services. 2012. Chronic conditions among Medicare beneficiaries. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Chronic-Conditions/2012ChartBook> (accessed 22 February 2023).

Dave, D. 2013. Effects of pharmaceutical promotion: A review and assessment. *NBER Working Paper No. 18830*.

Deng, Y. and C. Mela. 2018. TV viewing and advertising targeting. *Journal of Marketing Research*. LV: 99-118.

Donohue, J., Berndt, E., Rosenthal, M., Epstein, A., and R. Frank. 2004. Effects of pharmaceutical promotion on adherence to the treatment guidelines for depression. *Medical Care*. 42(12): 1176–1185.

Dubois, P., Griffith, R. and M. O'Connell. 2018. The effects of banning advertising in junk food markets. *Review of Economic Studies*. 85: 396-436.

Duggan, M. and Scott Morton, F. 2010. The impact of Medicare Part D on pharmaceutical prices and utilization. *American Economic Review*. 100(1): 590-607.

Goldman, D., Joyce, G., and Zheng, Y. 2007. Prescription drug cost sharing: Associations with medication and medical utilization and spending and health. *JAMA*. 298(1): 61-69.

Goldman, D., et al. 2004. Pharmacy benefits and the use of drugs by the chronically ill. *JAMA*. 291(19): 2344-2350.

Hirschfeld, R, et al. 1997. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. *JAMA*. 277(4): 333-340.

Iizuka, T. and G. Jin. 2005. The effect of prescription drug advertising on doctor visits. *Journal of Economics and Management Strategy*. 14(3): 701–727.

Imbens, G. and J. Angrist. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica*. 62(2): 467-475.

- Kanazawa, M. and J. Funk. 2001. Racial discrimination in professional basketball: Evidence from Nielsen ratings. *Economic Inquiry*. 39(4): 599–608.
- Kearney, M. and P. Levine. 2015. Media influences on social outcomes: The impact of MTV’s 16 and Pregnant on teen childbearing. *American Economic Review*. 105(12): 3597-3632.
- Ketcham, J. and K. Simon. 2008. Medicare Part D's effects on elderly drug costs and utilization. *American Journal of Managed Care*. 11: SP14-22.
- Lakdawalla, D., Sood, N., and Q. Gu. 2013. Pharmaceutical advertising and Medicare Part D. *Journal of Health Economics*. 32: 1356-1367.
- Lakdawalla, D., and W. Yin. 2015. Insurers' negotiating leverage and the external effects of Medicare Part D. *Review of Economics and Statistics*. 97(2): 314-331.
- Lichtenberg, F. and Sun, S. 2007. The impact of Medicare Part D on prescription drug use by the elderly. *Health Affairs*. 26(6): 1735-1744.
- Majumdar, S., Gurwitz, J., and S. Soumerai. 1999. Undertreatment of hyperlipidemia in the secondary prevention of coronary artery disease. *Journal of General Internal Medicine*. 14(12): 711-717.
- Nielsen. 2014. The total audience report. December Issue.
- Office of Inspector General. 2007. Generic drug utilization in the Medicare Part D program. <https://oig.hhs.gov/oei/reports/oei-05-07-00130.pdf> (accessed 22 February 2023).
- Pandya, A., et al. 2015. Cost-effectiveness of 10-year risk thresholds for initiation of statin therapy for primary prevention of cardiovascular disease. *JAMA*. 314(2): 142-50.
- Rosenthal, M., Berndt, E., Donohue, J., Epstein, A., and R. Frank. 2003. Demand effects of recent changes in prescription drug promotions. *Frontiers in health policy research*. 6: 1-26.
- Schwartz, L. and S. Woloshin. 2019. Medical marketing in the U.S., 1997-2016. *JAMA*. 321(1): 80-96.
- Sethuraman, R., Tellis, G., and R. Briesch. 2011. How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research*. 48(3): 457-471.
- Shapiro, B. 2018. Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy*. 126(1): 381-437.
- Shapiro, B. 2022. Promoting Wellness or Waste? Evidence from Antidepressant Advertising. *American Economic Journal: Microeconomics*. 14 (2): 439-77.

Sinkinson, M. and A. Starc. 2019. Ask your doctor? Direct-to-consumer advertising of pharmaceuticals. *Review of Economic Studies*. 96: 836-881.

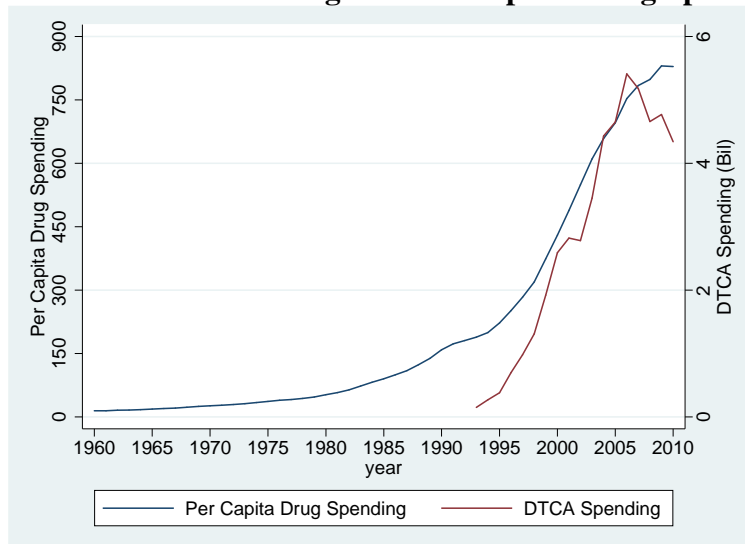
Wosinska, M. 2002. Just what the patient ordered? Direct-to-consumer advertising and the demand for pharmaceutical products. *HBS Marketing Research Papers No. 02-04*.

Wosinska, M. 2005. Direct-to-consumer advertising and drug therapy compliance. *Journal of Marketing Research*. 42(3): 323-332.

Yin, W., Basu, A., Zhang, J., Rabbani, A., Meltzer, D., and G. Alexander. 2008. The effect of the Medicare Part D prescription benefit on drug utilization and expenditures. *Annals of Internal Medicine*. 148(3): 169-177.

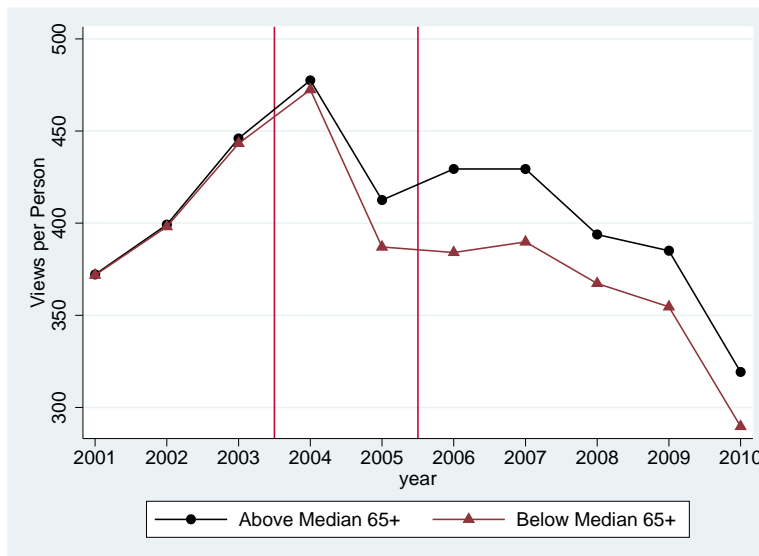
Zhang, Y., Donohue, J., Lave, J. and W. Gellad. 2011. The impact of Medicare Part D on medication treatment of hypertension. *Health services research*. 46: 185-198.

Figure 1 –Direct-to-Consumer Advertising and Prescription Drug Spending, 1960-2010



Sources: Dave (2013), National Health Expenditure Accounts (2015). The data are presented in nominal values.

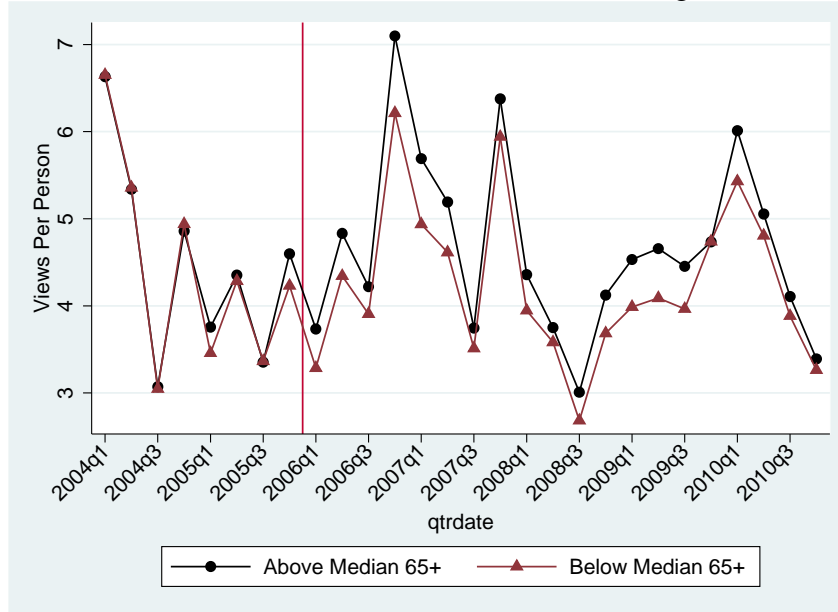
Figure 2 – Annual Views per Person of TV Ads for Top 200 Brand Name Drugs, for Non-Elderly



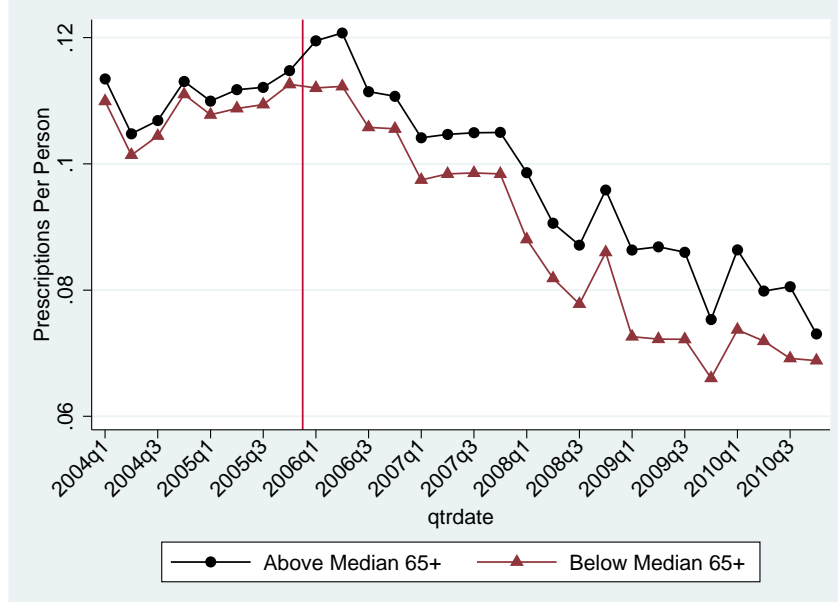
Notes: Sample means from Nielsen Ad*Views in above median elderly share markets relative to below median elderly share markets. The vertical lines represent the dates when Part D was signed into law (December 2003) and was implemented (January 2006). Means are plotted for a balanced panel of the top 200 advertised brand-name drugs. There is a secular downward trend in overall views per person due to patent expirations of several of these drugs over this time period (in particular, four of the top 200 drugs went off patent around 2006: Pravachol, Wellbutrin XL, Zocor, and Zolof). The downward trend in views matches the pattern in national advertising expenditures shown in Figure 1. In Appendix Figure B.2, we exclude all drugs that went off-patent during the study period.

Figure 3 – Quarterly Views per Person of TV Ads and Mean Utilization of Chronic Drugs, for Non-Elderly

Panel A: Views Per Person for Chronic Drug Ads



Panel B: Average Number of Prescriptions Purchased for Chronic Drugs

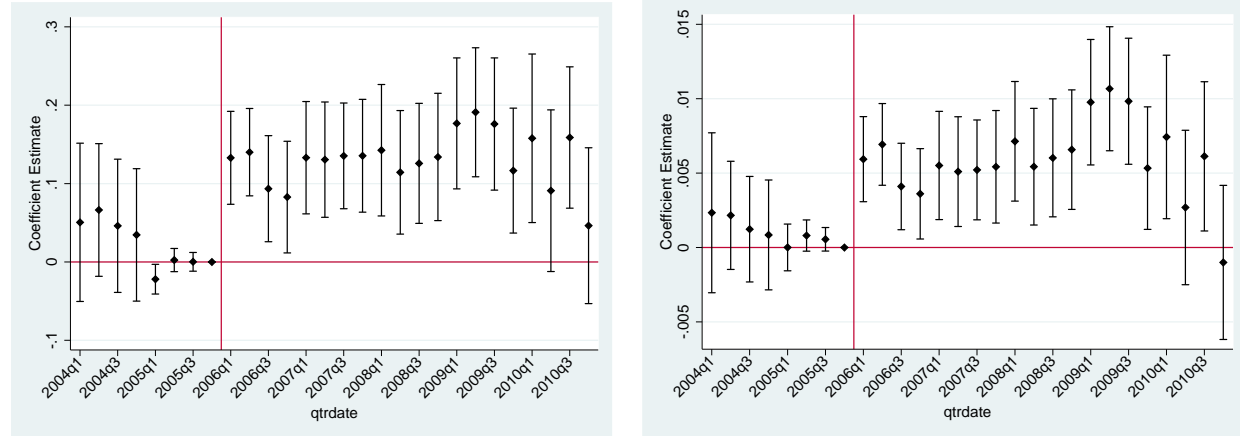


Notes: Sample means from Nielsen Ad*Views (views per capita for non-elderly) and claims (mean number of total prescriptions purchased for non-elderly) in above median elderly share markets relative to below median elderly share markets. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depressions, diabetes, hyperlipidemia, hypertension, and osteoporosis. The vertical line represents the implementation date of Medicare Part D.

Figure 4 – Event Study: Total Utilization of Chronic Drugs, Non-Elderly

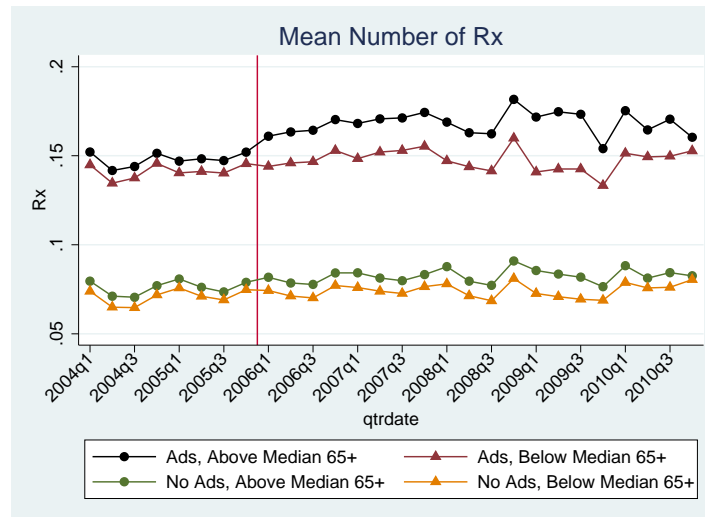
Panel A: Instrument=Share65+

Panel B: Instrument=High Elderly Share



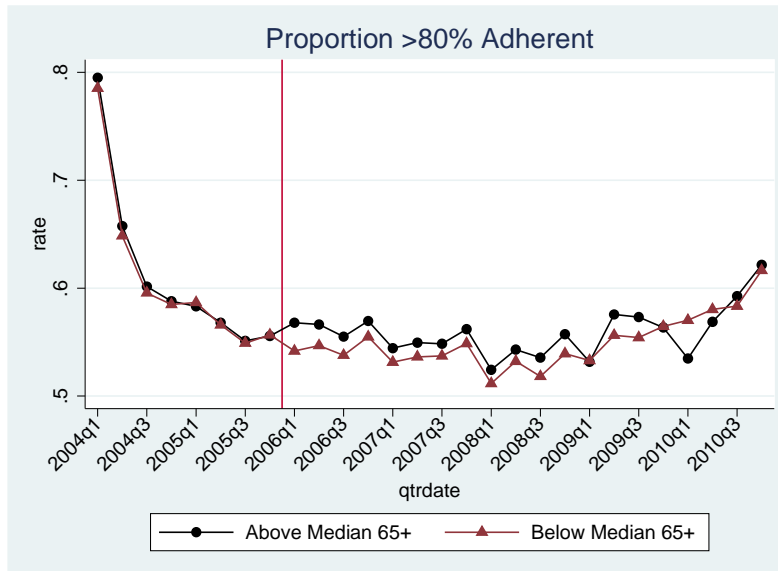
Notes: Event study coefficients and 95% confidence intervals. Clustered standard errors at the 3-digit ZIP code level. The outcome variable is the number of prescriptions. The event study is similar to Equation 3, but Panel A interacts quarter fixed effects with the Share65+ variable (continuous instrument) and Panel B interacts quarter fixed effects with the High Elderly Share indicator (binary instrument). All specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Estimates are normalized to zero in quarter 4 of 2005. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Figure 5 – Mean Utilization for Advertised Drug Classes vs. Non-Advertised Drug Classes, for Non-Elderly



Notes: Sample means from claims (mean number of total prescriptions purchased, ages 40-60) in above median elderly share markets relative to below median elderly share markets. The top two lines (black and red) are for the top 10 advertised drug classes and the bottom two lines (green and orange) are for the top 10 non-advertised drug classes (see Appendix Figure B.6 for full list of drug classes included). The vertical line represents the implementation date of Medicare Part D. We use the first two digits of the GPI code (available from IMS Health) to identify major classes of drugs.

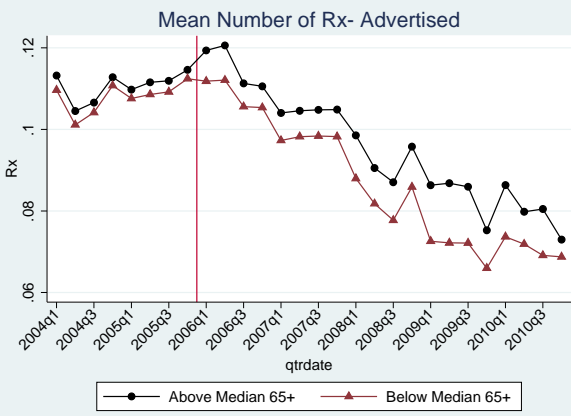
Figure 6 –Proportion with High Adherence of Chronic Drugs, for Non-Elderly



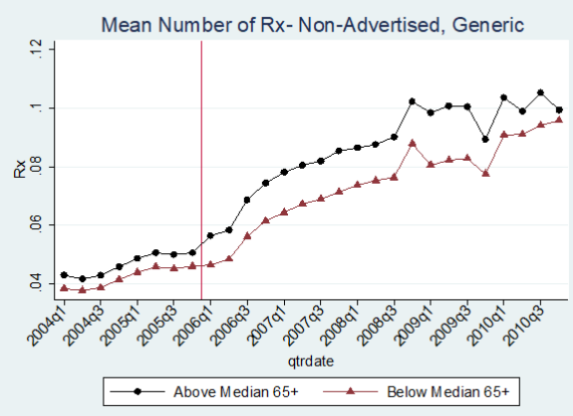
Notes: Sample means from claims (proportion of individuals with $MPR \geq 80\%$, ages 40-60) in above median elderly share markets relative to below median elderly share markets. Includes the 50 drugs that advertised during the study period and the drugs that did not advertise for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. The vertical line represents the implementation date of Medicare Part D.

Figure 7 – Quarterly Mean Utilization of Chronic Drugs: Spillover Effects

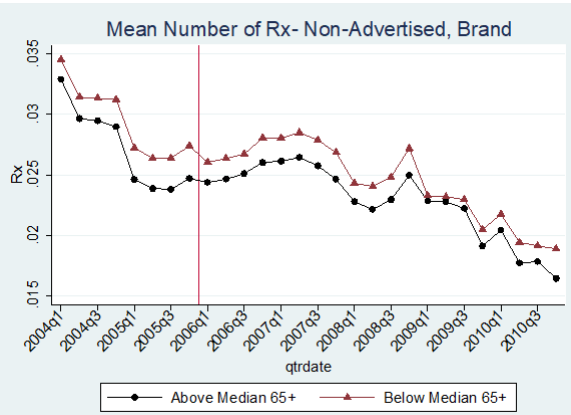
Panel A: Advertised Chronic Drugs



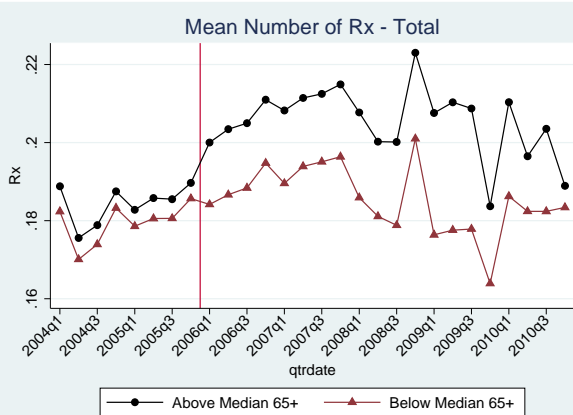
Panel B: Non-Advertised Chronic Drugs - Generic



Panel C: Non-Advertised Chronic Drugs – Brand



Panel D: Total Chronic Drugs



Notes: Sample means from claims (mean number of total prescriptions purchased, ages 40-60) in above median elderly share markets relative to below median elderly share markets. Panel A includes the 50 chronic drugs that advertised during the study period (repeated from Figure 3); Panel B includes generic drugs that did not advertise, but are in the same classes as the 50 advertised chronic drugs; Panel C includes brand drugs that did not advertise, but are in the same classes as the 50 advertised chronic drugs; Panel D includes both the advertised and non-advertised chronic drugs combined. The vertical line represents the implementation date of Medicare Part D.

Table 1 – Heterogeneity in Elderly Share Across Local TV Markets

TV Market	Share 65+	Pop 65+ (Census 2000)	Total Pop (Census 2000)	TV Market Ranking (Size)
<i>Top 8 High Elderly Share Markets</i>				
FT. MYERS-NAPLES	0.257	234,535	912,887	62
WEST PALM BEACH-FT. PIERCE	0.238	380,814	1,598,528	38
TAMPA-ST. PETE (SARASOTA)	0.213	787,553	3,702,269	14
WILKES BARRE-SCRANTON-HZTN	0.175	259,761	1,481,798	54
PITTSBURGH	0.173	503,077	2,901,329	23
ORLANDO-DAYTONA BCH-MELBRN	0.167	488,991	2,926,227	18
PADUCAH-CAPE GIRARD-HARSBG	0.158	156,329	987,215	81
SPRINGFIELD, MO	0.158	148,844	942,604	75
<i>Top 8 Low Elderly Share Markets</i>				
HOUSTON	0.082	410,910	5,020,575	10
SALT LAKE CITY	0.085	204,008	2,387,354	33
AUSTIN	0.085	116,640	1,371,385	40
ATLANTA	0.085	437,654	5,149,717	9
DALLAS-FT. WORTH	0.087	503,232	5,761,057	5
DENVER	0.093	320,372	3,451,529	17
WASHINGTON, DC (HAGRSTWN)	0.096	501,141	5,232,970	8
LOS ANGELES	0.098	1,578,642	16,144,245	2

Table 2 – Baseline Regressions for Total Utilization of Chronic Drugs, for Non-Elderly

Dependent Variable:	First Stage	Reduced Form	2SLS
	Views per Person (Non-Elderly) (1)	# of Prescriptions (2)	# of Prescriptions (3)
<i>A. Instrument=Share65+*Post</i>			
Share65+*Post	6.358*** (1.116)	0.107*** (0.023)	
Views per Person (Non-Elderly)			0.017*** (0.004)
F-statistic	32.69		
<i>B. Instrument=High Elderly Share*Post</i>			
High Elderly Share*Post	0.348*** (0.063)	0.005*** (0.001)	
Views per Person (Non-Elderly)			0.014*** (0.005)
F-statistic	30.86		
Mean of Dep. Var. (pre- Part D)	4.28	0.11	0.11
Zipcode x Condition x Quarter Obs	107,345	107,345	107,345

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. Data is from 2004-2010.

Table 3 – Total Utilization of Chronic Drugs – Alternative Specifications

	Reduced Form		2SLS	
	Instrument= Share65+*Post	Instrument= High Elderly Share*Post	Instrument= Share65+*Post	Instrument= High Elderly Share*Post
Dependent Variable: # of Prescriptions				
	(1)	(2)	(3)	(4)
1. Baseline Specification	0.107*** (0.023)	0.005*** (0.001)	0.017*** (0.004)	0.014*** (0.005)
2. Using condition x market and condition x quarter fixed effects	0.107*** (0.023)	0.005*** (0.001)	0.017*** (0.004)	0.014*** (0.005)
3. Adding zipcode-specific linear trends	0.102*** (0.021)	0.005*** (0.001)	0.010*** (0.003)	0.008*** (0.002)
4. Adding condition-specific linear trends	0.107*** (0.023)	0.005*** (0.001)	0.017*** (0.004)	0.014*** (0.005)
5. Excluding 2008-2010	0.097*** (0.018)	0.004*** (0.001)	0.011*** (0.003)	0.008*** (0.003)
6. Excluding 2005	0.084*** (0.027)	0.004*** (0.002)	0.012*** (0.004)	0.010** (0.004)
7. Including only continuously enrolled firms	0.072*** (0.027)	0.004** (0.002)	0.012** (0.006)	0.015* (0.008)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects (unless otherwise specified). Each cell shows the coefficient on Instrument x Post from a separate regression. The specifications are: 1) same as Table 2, 2) adds condition x market and condition x quarter fixed effects (instead of condition, market and quarter fixed effects separately), 3) adds 3-digit ZIP code specific linear trends, 4) adds condition-specific linear trends, 5) excludes the years 2008-2010, 6) excludes the year 2005, 7) includes only firms that were continuously in the claims sample from 2004-2010. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Table 4 – Detailing Test: Within-DMA Total Utilization of Chronic Drugs

Dependent Variable:	# of Prescriptions		
	Baseline	ZIP3 level	DMA x Qtr FE
	(1)	(2)	(3)
A. Instrument=Share65+*Post			
Share65+*Post (DMA level)	0.111*** (0.033)		
Share65+*Post (ZIP3 level)		0.087*** (0.027)	0.015 (0.026)
B. Instrument=High Elderly Share*Post			
High Elderly Share*Post (DMA level)	0.003** (0.002)		
High Elderly Share*Post (ZIP3 level)		0.006*** (0.002)	0.002 (0.002)
Mean of Dep. Var. (pre- Part D)	0.10	0.10	0.10
Zipcode x Condition x Quarter Obs	67,495	67,495	67,495

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include qtr fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Col 1: same as Table 2, but for sample of ZIP codes that are uniquely matched to one DMA, elderly share computed at the DMA-level; Col 2: elderly share computed at 3-digit ZIP code level; Col 3: adds DMA x quarter fixed effects, elderly share computed at 3-digit ZIP code level. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Table 5 – Adherence of Chronic Drugs

Dependent Variable: I(High Adherence)	Reduced Form			2SLS		
	Full Sample	2004-2007	2005-2007	Full Sample	2004-2007	2005-2007
	(1)	(2)	(3)	(4)	(5)	(6)
A. Instrument=Share65+*Post						
Post*Share65+	0.184*** (0.057)	0.234*** (0.056)	0.404*** (0.114)			
Views per Person (Non-Elderly)				0.017** (0.007)	0.017*** (0.005)	0.033*** (0.013)
B. Instrument=High Elderly Share*Post						
High Elderly Share*Post	0.004* (0.003)	0.008*** (0.003)	0.012*** (0.004)			
Views per Person (Non-Elderly)				0.008 (0.005)	0.011*** (0.004)	0.021** (0.008)
Mean of Dep. Var. (pre- Part D)	0.61	0.61	0.56			
Zipcode x Condition x Quarter Obs	102,477	59,252	44,519	102,477	59,252	44,519

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. The outcome variable is the proportion of individuals with MPR>=80%. Includes the 50 drugs that advertised during the study period and the drugs that did not advertise for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Table 6 – Spillover Effects on Non-Advertised Chronic Drugs

Dependent Variable:	# of Prescriptions	Reduced Form				2SLS			
		Advertised Drugs	Non-Advertised Drugs: Generic	Non-Advertised Drugs: Brand	Total	Advertised Drugs	Non-Advertised Drugs: Generic	Non-Advertised Drugs: Brand	Total
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Instrument=Share65+*Post									
Post*Share65+		0.107*** (0.023)	0.100*** (0.027)	0.025*** (0.007)	0.233*** (0.038)				
Views per Person (Non-Elderly)						0.017*** (0.004)	0.016*** (0.005)	0.004*** (0.002)	0.037*** (0.008)
B. Instrument=High Elderly Share*Post									
High Elderly Share*Post		0.005*** (0.001)	0.006*** (0.002)	0.001 0.000	0.011*** (0.002)				
Views per Person (Non-Elderly)						0.014*** (0.005)	0.016*** (0.005)	0.002 (0.002)	0.032*** (0.008)
Mean of Dep. Var (pre- Part D)		0.11	0.04	0.03	0.18	0.11	0.04	0.03	0.18
Zipcode x Condition x Quarter Obs		107,345	107,345	107,345	107,345	107,345	107,345	107,345	107,345

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Includes the 50 chronic drugs that advertised during the study period and the drugs in the same classes that did not advertise. Non-advertised drugs are separated into generic and brand products. Data is from 2004-2010.

FOR ONLINE PUBLICATION

Appendix A

A.1 Decomposition of Extensive and Intensive Margin Effects

The increase in total drug utilization we observe due to advertising is driven both by increased take-up of treatment (extensive margin) as well as increased use among existing drug users (intensive margin). In this section we decompose the overall effect of advertising into the extensive and intensive margins. We re-estimate the main 2SLS regression for total prescriptions shutting down the extensive margin effect (i.e. we hold the take-up rate constant at pre-Part D levels). This allows us to estimate the proportion of the total effect that is due to intensive margin changes. The remaining proportion of the total effect is then due to extensive margin changes. We estimate:

$$(A1) \text{ AnyUse}_{mct,<65} = \sigma_m + \tau_c + \epsilon_{mct}$$

Where $\text{AnyUse}_{mct,<65}$ is the proportion of non-elderly individuals with any use of prescription drugs for condition c in market m and quarter t . We estimate this equation for the pre-Part D period from 2004-2005. We then use the estimated parameters to predict $\widehat{\text{AnyUse}}_{mct,<65}$ for the entire sample from 2004-2010. Since $\widehat{\text{AnyUse}}_{mct,<65}$ varies only by condition and market, the proportion that takes-up prescription drugs (i.e. extensive margin) is held constant over time.

Next we construct a counterfactual measure of total prescriptions purchased which holds the extensive margin constant: $\hat{Y}_{mct,<65} = \widehat{\text{AnyUse}}_{mct,<65} * Y_{mct,<65}$, where $Y_{mct,<65}$ is the average number of prescriptions purchased conditional on use. We use this constructed outcome variable to estimate the 2SLS model as before. The 2SLS results are presented in Columns 3 and 4 of Appendix Table B.9. The estimates using the counterfactual outcome, representing the intensive margin, are less than one-third the size of the baseline total effect (Columns 1 and 2). Using the preferred continuous instrument specification, this implies that extensive margin changes explain 71% to 73% of the total utilization effect. A straightforward back-of-the-envelope calculation using the pre-Part D take-up rate and estimated intensive margin effect from Appendix Table B.8 produces similar estimates of the extensive margin effect.⁵⁹

A.2 Construction of Adherence Measures

⁵⁹ Using the continuous instrument, the change in prescriptions purchased is 0.017 for one additional ad viewed. The change in prescriptions purchased along the intensive margin is predicted to be the fraction of the sample that used chronic drugs prior to Part D (0.063) x the estimated change in prescriptions purchased among users (0.057). Subtracting this from the total effect, we get the predicted extensive margin effect: 0.013 (or 76% of the total effect).

We measure adherence between the individual's first drug claim through their last drug claim for that condition.⁶⁰ We consider adherence as receiving continuous treatment for a condition, rather than for a specific brand since this is the more relevant dimension from a welfare perspective. Our measure of adherence is the medication possession ratio (MPR). The MPR is calculated as the number of days with drug on-hand (i.e. days supplied) divided by the number of days in the quarter. We adjust the numerator of the MPR to account for claims with overlapping days supplied. For example, it is typical to refill a prescription before finishing the days supplied for the initial prescription. If overlapping claims have the same active ingredient,⁶¹ we assume that the person finishes the days supplied in the first claim before starting the days supplied in the second claim. For overlapping claims with different active ingredient names (for the same condition), we assume that patients start using the days supplied for the second claim on the fill date and discard remaining days supplied for the first claim. This case likely represents a drug switch. Since advertising may lead to more drug switching, it is especially important to account for this case to avoid overstating the effect of advertising on adherence. Days in the hospital were assumed to be fully compliant and patients resumed their prescriptions after they were discharged.

After constructing the quarterly MPR for each individual by condition, we also create a binary indicator for individuals who had $MPR \geq 80\%$, which is considered high adherence and is the threshold most commonly reported in the pharmaceutical literature (Andrade, et al., 2006). As before, we collapse the data by three-digit ZIP code, condition, and quarter, computing the mean MPR and the proportion of individuals with $MPR \geq 80\%$ in each cell.

A.3 Potential Welfare Effects

In this section, we compute a back-of-the-envelope estimate of the lifetime welfare gain or loss from one additional ad viewed by a non-elderly person. The welfare effects flow from three sources. First, the additional exposure to advertising encourages some to initiate therapy and have high adherence (defined as $MPR \geq 80\%$) with therapy. This population experiences a welfare gain from an improvement in health but needs to pay for medication costs. We do not include savings from reduced hospitalizations and other healthcare utilization, so this estimate will understate the welfare gain. The net welfare gain/loss for this population is computed as:

⁶⁰ In some specifications, we account for discontinuation by computing an alternative measure of adherence where we assume that a person is non-adherent after their last observed drug claim.

⁶¹ Combination drugs are viewed as a unique combination of two or more active ingredients.

$$(A1) \quad E(QALY * Value(QALY) - MC|Initiate, HighAdherence) \\ * Pr(HighAdherence|Initiate) * Pr(Initiate)$$

Where $E(QALY * Value(QALY) - MC|Initiate, HighAdherence)$ measures the value of Quality-Adjusted Life Years (QALYs) gained net of incremental costs, among the population of patients who initiate therapy and adhere to it. $Pr(Initiate)$ is the increase in the probability of initiating therapy due to one additional ad viewed and $Pr(HighAdherence|Initiate)$ is the probability of high adherence to therapy among the marginal initiators.

Second, the additional exposure to advertising encourages some to initiate therapy but have low adherence to therapy (MPR < 80%). These marginal initiators are potentially less appropriate for therapy and, consequently, it produces an insufficient clinical benefit or adverse side effects. We assume this population has no health benefit but pays higher treatment costs, therefore experiencing a welfare loss. This assumption is conservative and may understate the benefits for this population. We compute the welfare loss for this population as:

$$(A2) \quad E(MC|Initiate, LowAdherence) * (1 - Pr(HighAdherence|Initiate)) * Pr(Initiate)$$

Where $E(MC|Initiate, LowAdherence)$ is the marginal cost of treatment for patients who initiate but have low adherence.

Third, additional exposure to advertising encourages some who have previously initiated treatment to better adhere to their treatment (i.e., shift from low to high adherence). The welfare gain for this population is given by:

$$(A3) \quad E(\Delta QALY * Value(QALY) - \Delta MC|HighAdherence, Rx use) \\ * \Delta Pr(HighAdherence| Rx use) * Pr(Rx use)$$

Where $E(\Delta QALY * Value(QALY) - \Delta MC|HighAdherence, Rx use)$ measures the value of QALYs gained net of incremental costs, among the population of existing drug users who attain high adherence as a result of viewing one additional ad. This term is multiplied by the increase in the probability of high adherence among existing prescription drug users, and the baseline proportion of individuals in our sample using prescription drugs, $Pr(Rx use)$.

As one can see from the above equations, the consumer welfare effect from exposure to advertising might vary by drug class. The welfare gain will be higher for drugs that are more cost-effective, because the welfare gain from advertising accrues from the benefits of increased use of drugs relative to the increased costs. Also, when low-cost generic drugs are available in the drug class and advertising leads to spillovers to these drugs, the costs are further reduced and

the welfare gain will be larger. Finally, the welfare gain will be larger the more well targeted advertising is to initiators who are appropriate for treatment as they will have higher adherence.⁶²

We illustrate the welfare calculation in the case of statins. There are two main advantages of focusing on this drug class. First, statins are one of the most commonly used and advertised drug classes during our study period. Second, statins have been evaluated in numerous long-term clinical trials that provide credible evidence on the short-term and long-term health effects of statin use. We use a widely cited study estimating the net present value of lifetime benefits of initiating statin use (which has been discounted at 3%) to parameterize: $E(QALY|Initiate, HighAdherence) = 0.1275 QALYs$.⁶³ We set the value of one QALY equal to the conventional threshold, $Value(QALY) = \$150,000$.

The remaining parameters are estimated from our data. As before, non-elderly views per person are instrumented by the elderly share interacted with the introduction of Part D. Our 2SLS results are shown below in Appendix Table A.1 for hyperlipidemia drugs (the vast majority of hyperlipidemia drugs are statins). We use estimates from Panel A in the welfare calculation, but Panel B produces similar results. The outcome variables combine both advertised and non-advertised drugs in the drug class to account for advertising spillovers. Column 1 shows that one additional ad viewed by a non-elderly person increases the probability of using a hyperlipidemia drug by 0.009. Therefore, we set $Pr(Initiate) = 0.009$.

We estimate $\Delta Pr(HighAdherence| Rx use)$ as the marginal effect of advertising on the probability that an existing drug user has high adherence. Column 3 demonstrates that each additional non-elderly view among existing users increases the probability of high adherence by 0.014. Thus, we set $\Delta Pr(HighAdherence| Rx use) = 0.014$. The baseline proportion of individuals in our sample using hyperlipidemia drugs is $Pr(Rx use) = 0.10$.

We also estimate $Pr(HighAdherence|Initiate)$ or the probability of high adherence among the marginal initiators (those who initiate because of advertising). The average probability of adherence in our sample at baseline is 61 percent. However, marginal initiators

⁶² For example, targeting advertising to television programs or markets with more appropriate consumers could have higher welfare gains than targeting to a broader set of consumers. This may also apply to online advertising where targeting can be made even more precise.

⁶³ The QALY estimate comes from the cost-effectiveness analysis from Pandya et al. (2015) which uses as an input Baigent et al. (2005) to parameterize the health benefits from statins. Specifically, we use Table 2 to compute the per person lifetime QALY gain from a one percentage point increase in the proportion of adults receiving statins: $(17.327-17.276)/(.48-.08)=0.1275 QALYs$. The time horizon for this health benefit is 20 years on average (life expectancy minus the median aged person in the study's sample).

exhibit lower adherence, because excluding them from the sample increases the marginal effect of advertising on adherence from 0.011 to 0.014. These two coefficient estimates, along with the estimate of the reduced form probability of initiating due to advertising allows us to calculate that $Pr(\text{HighAdherence}|\text{Initiate}) = 0.578$.

Combining all of these effects – the welfare benefit of initiation among patients who adhere to treatment and welfare benefit of increased adherence among existing users – the lifetime benefits of one additional ad viewed is \$126.26 per capita for the 40 to 60 year old population (including individuals who do not receive treatment) in net present value.

We also estimate the marginal costs per quarter from one additional ad viewed from Column 4. This estimate encompasses the costs of initiation and adherence responses as well as spillovers to lower cost generics. Using this estimate, we project costs over the same time horizon used to calculate lifetime benefits, discounted at 3% (the same discount rate used to compute lifetime benefits). Following Pandya et al. (2015), we assume that the rate of high adherence falls to 50% (from our estimated 62.1% in year 1) in the second and all subsequent years.⁶⁴ We add to this the median cost to the firm of one advertising view, which is about five cents (Deng and Mela, 2018).⁶⁵ The lifetime costs stemming from one additional ad viewed is \$105.63 per capita in net present value.

Overall, our exercise yields a total estimated lifetime welfare gain of \$20.63 per capita from one additional statin ad viewed in the US for the 40 to 60 year old population. We interpret our estimate as a lower bound of the potential welfare gain for this drug class, because we do not include the cost savings from reduced medical utilization or account for reduced costs associated with future generic entry in this class during the lifetime of these individuals (e.g., cost-reductions from Lipitor’s 2011 patent expiration). We also likely underestimate the benefits of statins because we assume that low adherence does not produce any health benefits. However, even under these conservative assumptions, our estimate suggests that the benefits of advertising exceed the costs for this class of drugs.

⁶⁴ We multiply the year 1 annual cost by 0.5/0.621 in the second and subsequent years.

⁶⁵ In Table 2, Deng and Mela (2018) report that the median costs of one ad exposure is one cent. The chance that exposure leads to one view is about 20% (see Figure 2). Thus, the cost per view is five cents.

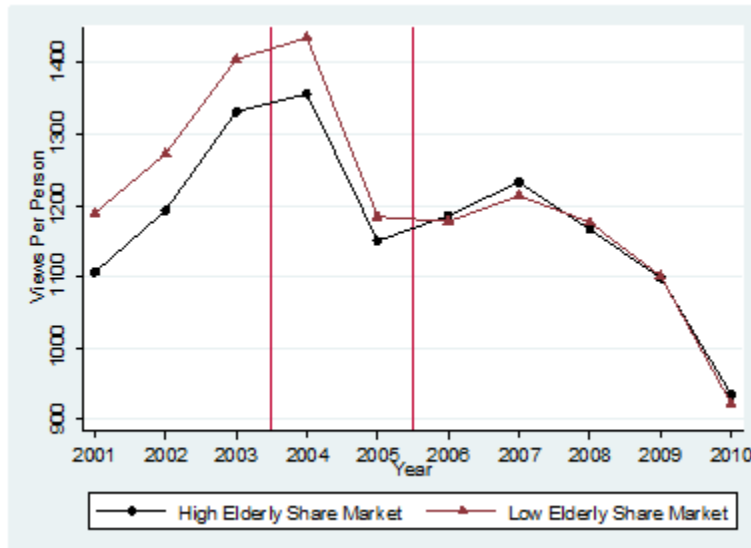
Appendix Table A.1 – Welfare Analysis: 2SLS Estimates for Hyperlipidemia Drugs

Dependent Variable :	Any Use	High Adherence (full sample)	High Adherence (excl. initiators)	Expenditures
	(1)	(2)	(3)	(4)
A. Instrument=Share65+*Post				
Views per Person (Non-Elderly)	0.009*** (0.002)	0.011*** (0.004)	0.014*** (0.004)	2.106*** (0.603)
B. Instrument=High Elderly Share*Post				
Views per Person (Non-Elderly)	0.007*** (0.002)	0.007** (0.003)	0.011*** (0.003)	1.219** (0.555)
Mean of Dep. Var (pre- Part D)	0.10	0.61	0.61	19.36
Zipcode x Quarter Obs	21,469	21,271	20,766	21,469

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the ZIP code level; the outcome variable is aggregated by zipcode and equal to: (1) indicator for any use, (2) indicator for high adherence (MPR>=80%) for the full sample; (3) indicator for high adherence (MPR>=80%) for the sample of existing users; (4) total expenditures; all specifications include quarter fixed effects and 3-digit ZIP code fixed effects. Outcomes include all advertised and non-advertised drugs in the hyperlipidemia drug class during the study period. Data is from 2004-2010.

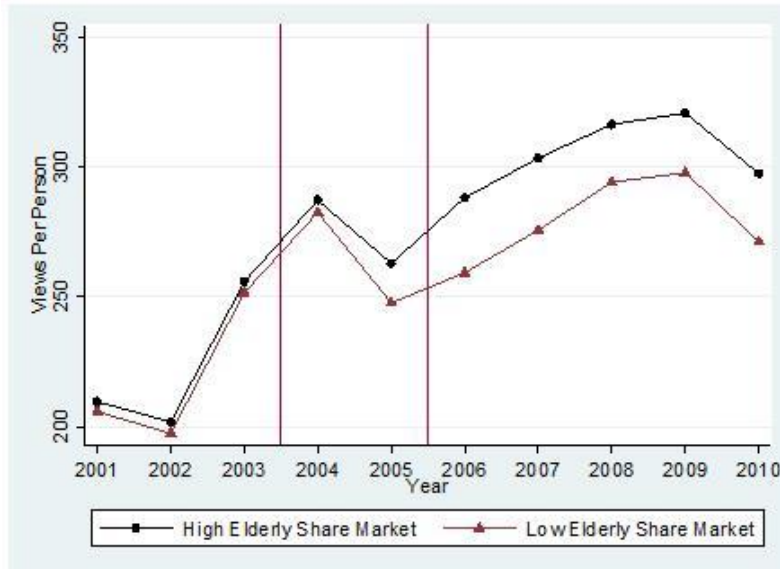
Appendix B: Additional Figures and Tables

Appendix Figure B.1 – Annual Views per Person of Ads for Top 200 Brand Name Drugs, for Elderly 65+



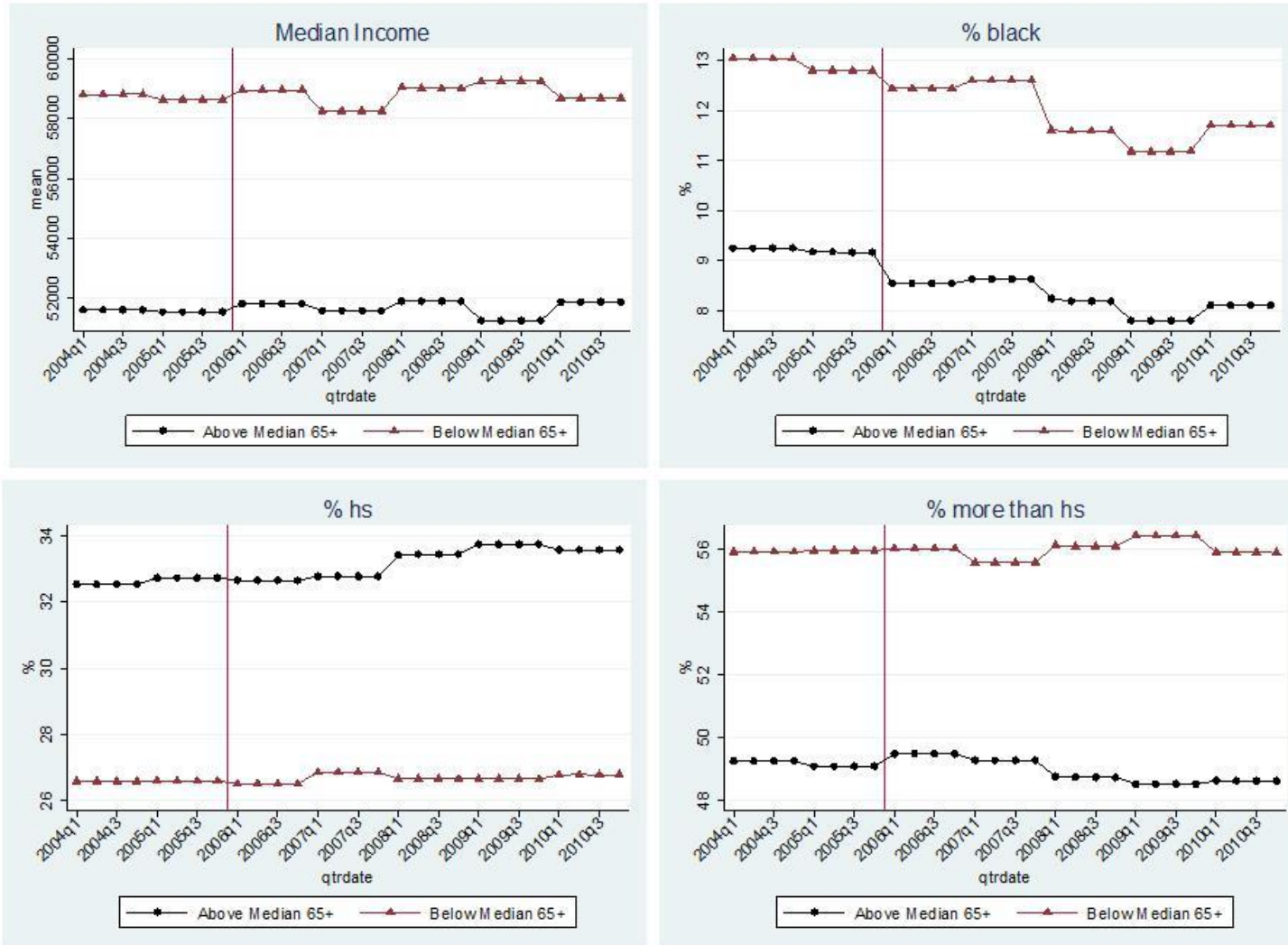
Notes: Sample means from Nielsen Ad*Views in above median elderly share markets relative to below median elderly share markets. The vertical lines represent the dates when Part D was signed into law (December 2003) and was implemented (January 2006). Means are plotted for a balanced panel of the top 200 advertised brand-name drugs. There is a secular downward trend in overall views per person due to patent expirations of several of these drugs over this time period (in particular, four of the top 200 drugs went off patent around 2006: Pravachol, Wellbutrin XL, Zocor, and Zoloft). The downward trend in views matches the pattern in national advertising expenditures shown in Figure 1.

**Appendix Figure B.2 – Annual Views per Person of Ads for Top Brand Name Drugs:
Excluding Drugs that went Off-patent from 2001-2010, for Non-Elderly**



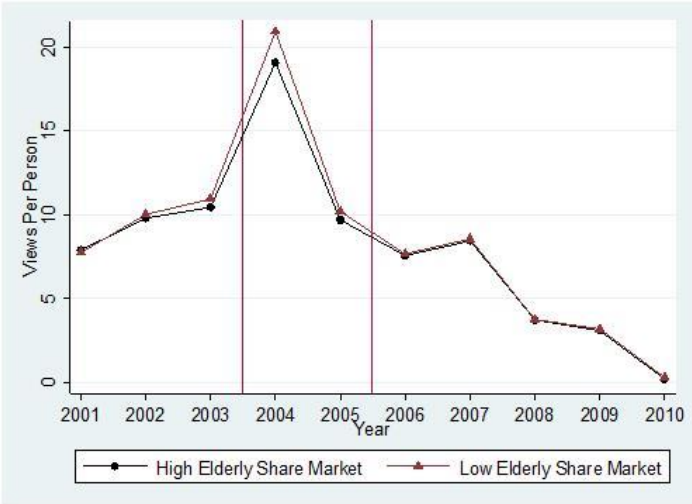
Notes: Sample means from Nielsen Ad*Views in above median elderly share markets relative to below median elderly share markets. The vertical lines represent the dates when Part D was signed into law (December 2003) and was implemented (January 2006). Means are plotted for a balanced panel of the top advertised brand-name drugs. We exclude all drugs that went off-patent during the study period.

Appendix Figure B.3 – Trends in Composition of Claims Data Sample



Notes: Sample means of Census 2000 characteristics linked to individuals in claims sample by 3-digit ZIP code.

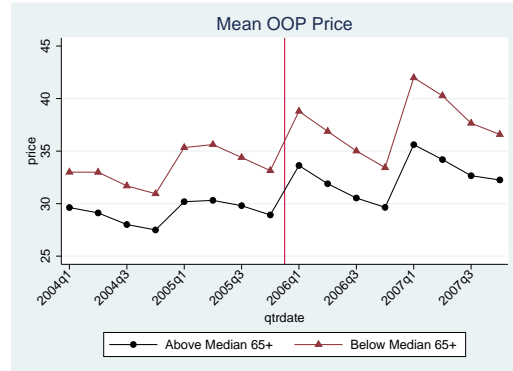
Appendix Figure B.4 – Placebo Test: Annual Views per Person of TV Ads for Contraceptive Drugs, for Non-Elderly



Notes: Sample means for contraceptive drugs from Nielsen Ad*Views in above median elderly share markets relative to below median elderly share markets. The vertical lines represent the dates when Part D was signed into law (December 2003) and was implemented (January 2006).

Appendix Figure B.5 – Prices for Balanced Panel of Chronic Drugs, for Non-Elderly

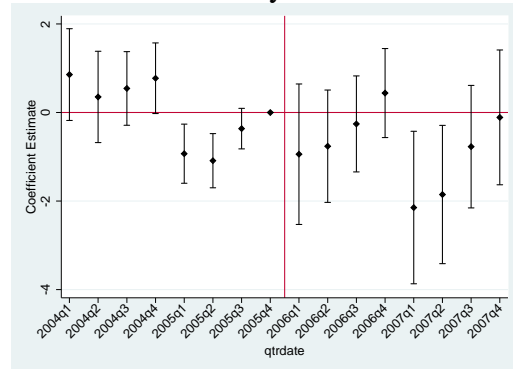
Panel A: Mean OOP Price



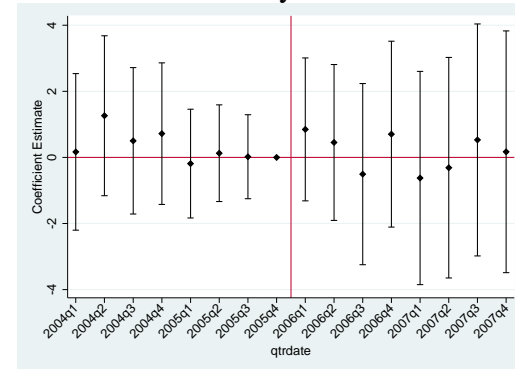
Panel B: Mean Total Price



Panel C: Event Study - OOP Price



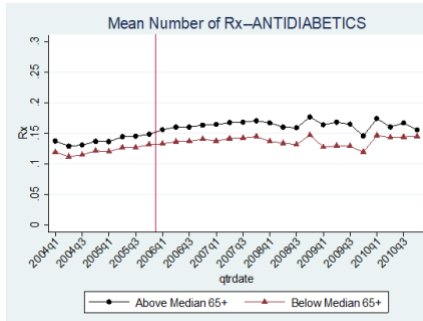
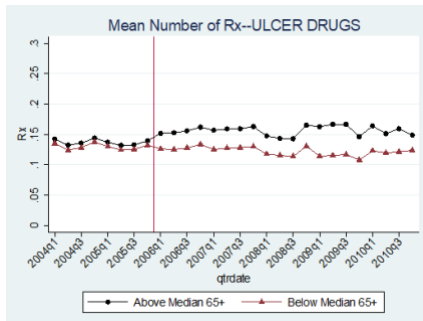
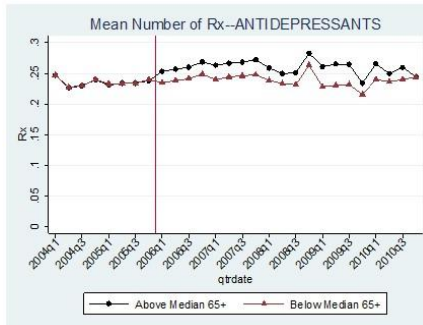
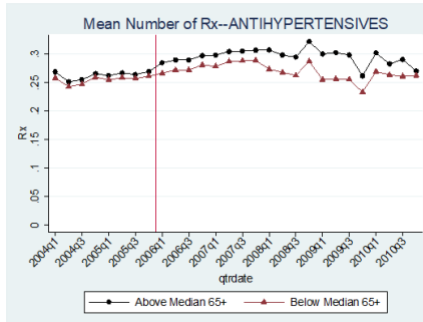
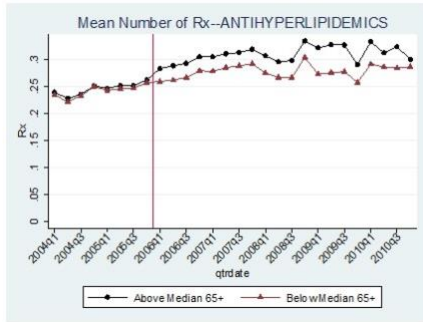
Panel D: Event Study - Total Price



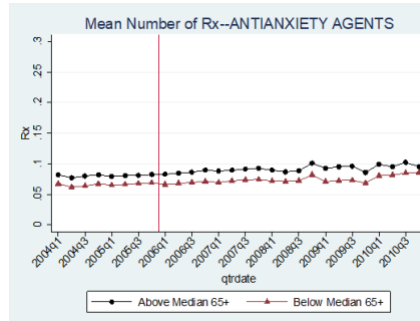
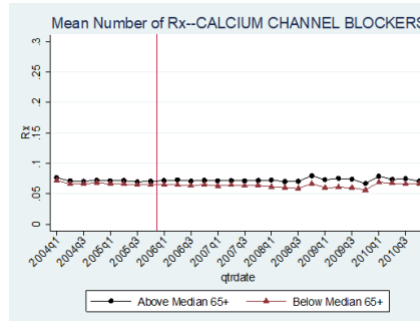
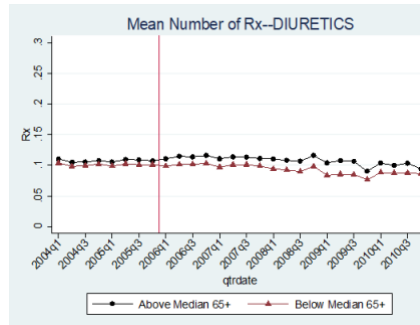
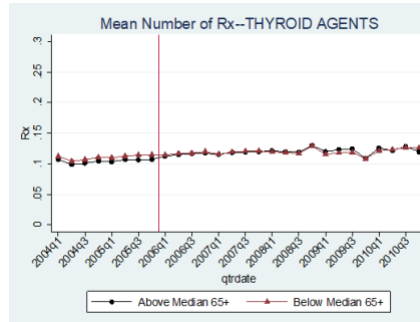
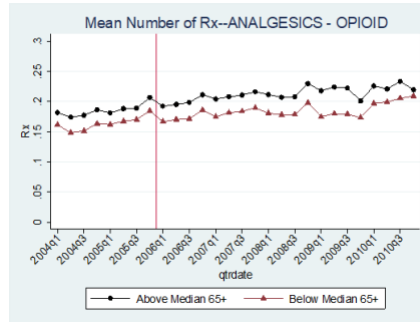
Notes: Panels A and B show raw sample means of out-of-pocket price and total price from claims (ages 40-60) in above median elderly share markets relative to below median elderly share markets. Panels C and D show event study coefficients from a regression at the NDC-market level similar to Equation 3 that replaces the interaction term with a full set of quarter dummies interacted with the high elderly share indicator variable. We also replace the condition fixed effects with NDC fixed effects. Estimates are normalized to zero in quarter 4 of 2005. In all figures, we use a balanced panel of NDCs from 2004-2007 for the 5 chronic conditions (i.e. each NDC has a non-missing price in all quarters). The vertical line represents the implementation date of Medicare Part D.

Appendix Figure B.6 - Effects for Advertised vs. Non-Advertised Drug Classes, by Class

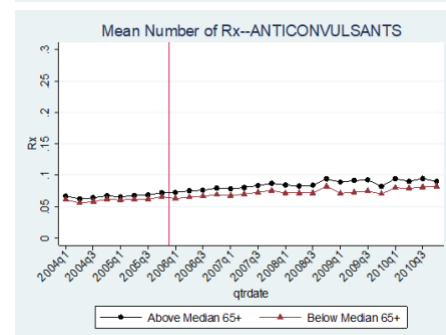
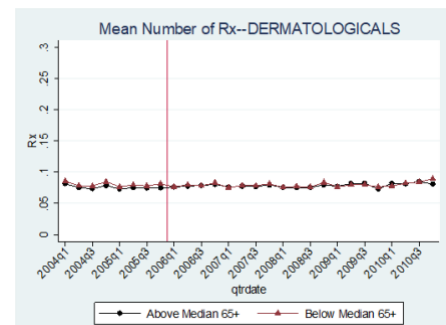
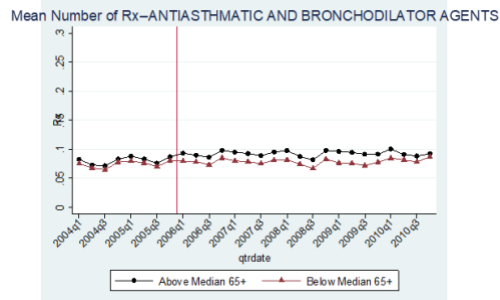
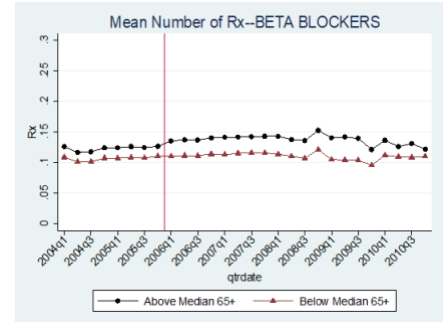
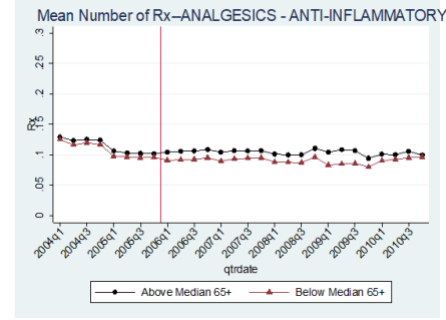
Advertised Drug Classes



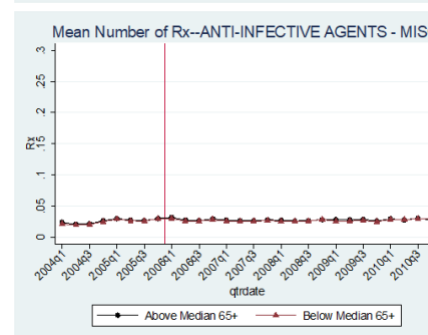
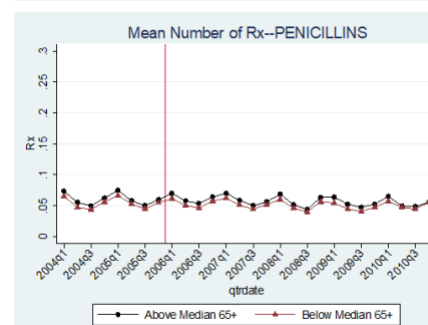
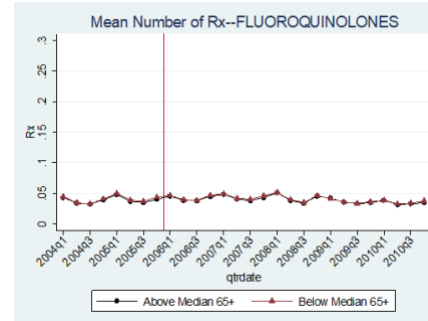
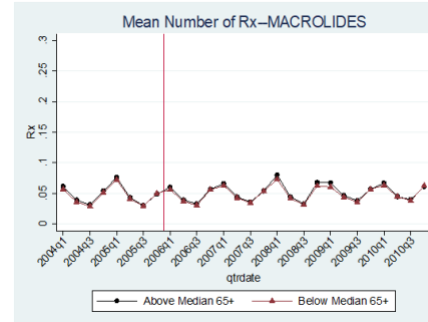
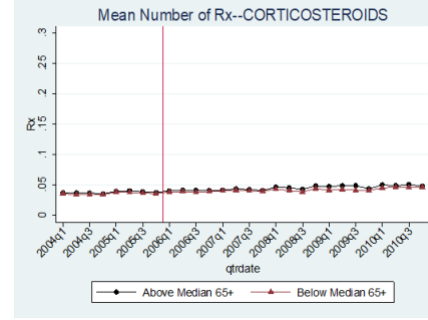
Non-Advertised Drug Classes



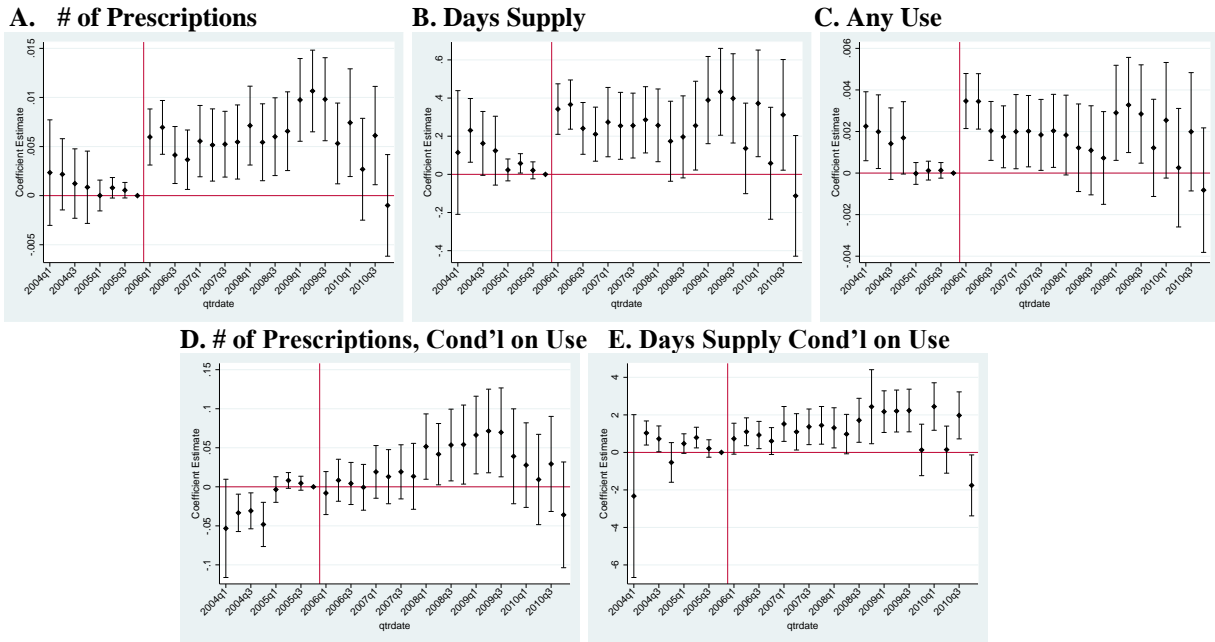
Advertised Drug Classes (Cont'd)



Non-Advertised Drug Classes (Cont'd)

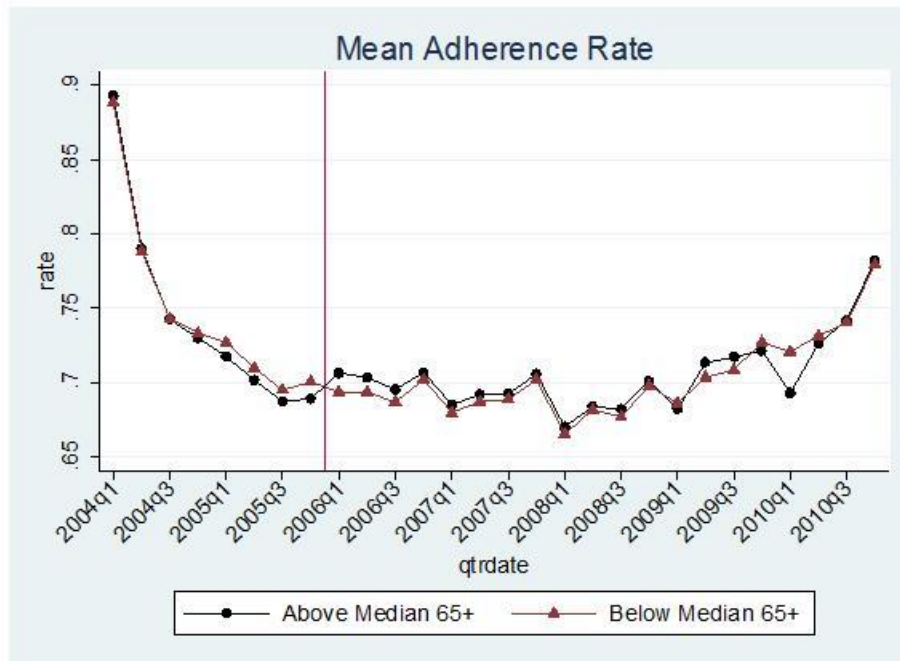


Appendix Figure B.7 –Event Study: Total Utilization of Chronic Drugs, Extensive and Intensive Margins



Notes: Event study coefficients and 95% confidence intervals. Clustered standard errors at the 3-digit ZIP code level. The outcome variable is the number of prescriptions (Panel A), days supply (Panel B), indicator for any prescription drug use (Panel C), number of prescriptions conditional on use (Panel D), days supply conditional on use (Panel E). Event study coefficients are from a regression similar to Equation 3 that replaces the interaction term with a full set of quarter dummies interacted with the high elderly share indicator variable. All specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Estimates are normalized to zero in quarter 4 of 2005. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

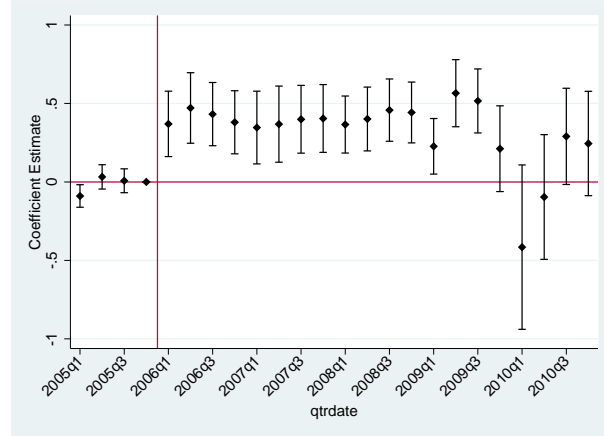
Appendix Figure B.8 – Adherence of Chronic Drugs: Medication Possession Ratio



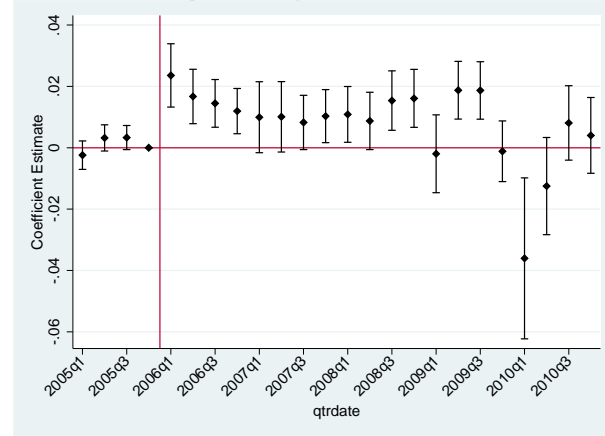
Notes: Sample means from claims (mean MPR, ages 40-60) in above median elderly share markets relative to below median elderly share markets. Includes the 50 drugs that advertised during the study period and the drugs that did not advertise for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. The vertical line represents the implementation date of Medicare Part D.

Appendix Figure B.9 –Event Study: Adherence of Chronic Drugs
Panel A: Proportion with High Adherence

Instrument=Share65+

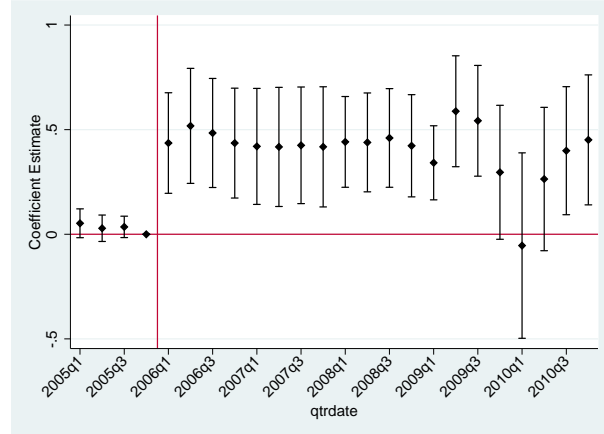


Instrument=High Elderly Share

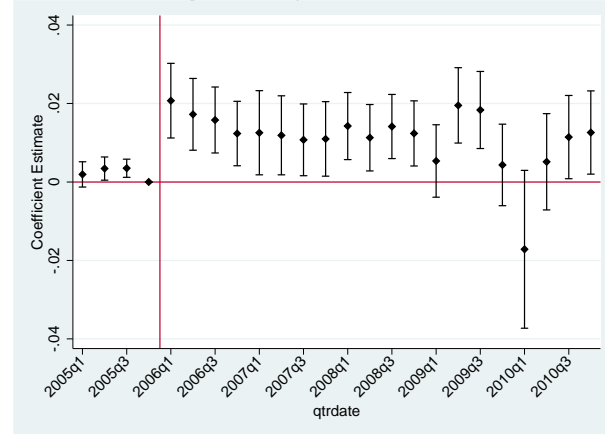


Panel B: Medication Possession Ratio

Instrument=Share65+

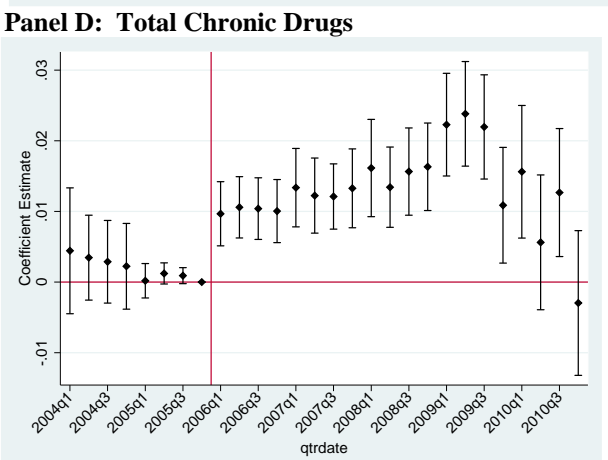
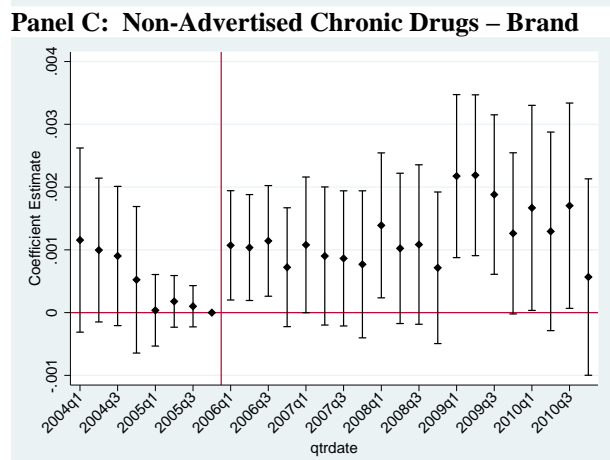
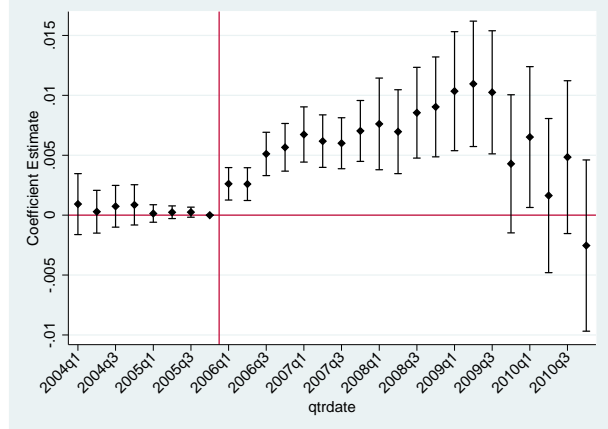
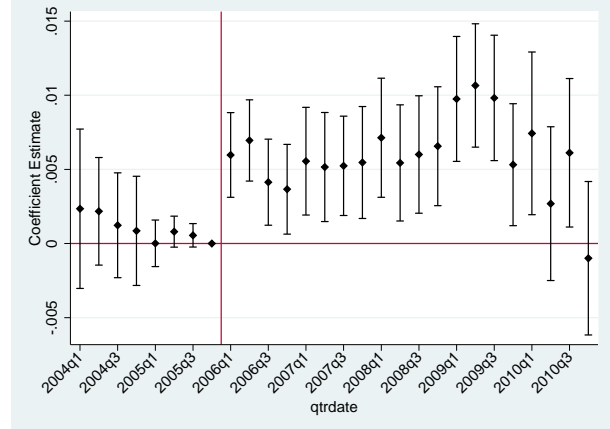


Instrument=High Elderly Share



Notes: Event study coefficients and 95% confidence intervals. Clustered standard errors at the 3-digit ZIP code level. The outcome variable in Panel A is an indicator for high adherence ($MPR \geq 80\%$) and Panel B is the medication possession ratio (MPR). The event study is similar to Equation 3, but interacts quarter fixed effects with either the Share65+ variable (continuous instrument) or High Elderly Share indicator (binary instrument). All specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Estimates are normalized to zero in quarter 4 of 2005. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Appendix Figure B.10 –Event Study: Spillover Effects on Non-Advertised Chronic Drugs



Notes: Event study coefficients and 95% confidence intervals. Clustered standard errors at the 3-digit ZIP code level. The outcome variable is the number of prescriptions. Panel A includes the 50 chronic drugs that advertised during the study period (repeated from Figure 4); Panel B includes generic drugs that did not advertise, but are in the same classes as the 50 advertised chronic drugs; Panel C includes brand drugs that did not advertise, but are in the same classes as the 50 advertised chronic drugs; Panel D includes both the advertised and non-advertised chronic drugs combined. Event study coefficients are from a regression similar to Equation 3 that replaces the interaction term with a full set of quarter dummies interacted with the high elderly share indicator variable. All specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Estimates are normalized to zero in quarter 4 of 2005.

Appendix Table B.1—Top Advertised Chronic Drugs, 2001-2010

Condition	Brand-Name Drug
DEPRESSION	CYMBALTA
DEPRESSION	EFFEXOR
DEPRESSION	PAXIL
DEPRESSION	PRISTIQ
DEPRESSION	PROZAC
DEPRESSION	SARAFEM
DEPRESSION	WELLBUTRIN
DEPRESSION	ZOLOFT
DIABETES	ACTOS
DIABETES	AVANDIA
DIABETES	BYETTA
DIABETES	EXUBERA
DIABETES	HUMALOG
DIABETES	JANUVIA
DIABETES	LANTUS
DIABETES	LEVEMIR
DIABETES	METAGLIP
DIABETES	ONGLYZA
DIABETES	NOVOLIN
DIABETES	NOVOLOG
HYPERLIPIDEMIA	ALTOCOR
HYPERLIPIDEMIA	BIDIL
HYPERLIPIDEMIA	CADUET
HYPERLIPIDEMIA	CRESTOR
HYPERLIPIDEMIA	LESCOL
HYPERLIPIDEMIA	LIPITOR
HYPERLIPIDEMIA	LOVAZA
HYPERLIPIDEMIA	NIASPAN
HYPERLIPIDEMIA	PRAVACHOL
HYPERLIPIDEMIA	TRILIPIX
HYPERLIPIDEMIA	VYTORIN
HYPERLIPIDEMIA	WELCHOL
HYPERLIPIDEMIA	ZETIA
HYPERLIPIDEMIA	ZOCOR
HYPERTENSION	ALTACE
HYPERTENSION	AVAPRO
HYPERTENSION	COREG
HYPERTENSION	DIOVAN
HYPERTENSION	INNOPRAN
HYPERTENSION	TEKTURNIA
HYPERTENSION	TOPROL
OSTEOPOROSIS	ACTIVELLA
OSTEOPOROSIS	ACTONEL
OSTEOPOROSIS	BONIVA
OSTEOPOROSIS	EVISTA
OSTEOPOROSIS	FORTEO
OSTEOPOROSIS	FOSAMAX
OSTEOPOROSIS	PREMARIN
OSTEOPOROSIS	PREMPRO
OSTEOPOROSIS	RECLAST

Appendix Table B.2 – Sample Means of Nielsen Advertising Variables by Elderly Share

Variable (Mean)	2005		2007		2005-07 Change	
	Low Elderly Share	High Elderly Share	Low Elderly Share	High Elderly Share	Low Elderly Share	High Elderly Share
Proportion 65+ (2000)	0.110	0.146	0.110	0.146	-	-
Population 65+ (2000)	333,864	256,288	333,864	256,288	-	-
Total Population (2000)	3,070,123	1,748,112	3,070,123	1,748,112	-	-
Views per Person (ages 2-64)	387	413	390	429	3	17
Views per Person (ages 65+)	1,184	1,150	1,214	1,233	30	82
Year x Market observations	50	50	50	50	50	50

Notes: Means are computed across DMAs by year for the top 200 advertised brand-name drugs. Views per Person (rating points) are from the Nielsen data. Elderly share and population counts are from the 2000 Census.

Appendix Table B.3— Effect of Part D on Views Per Person for Top 200 Drugs for Non-Elderly and Elderly Viewers

Dependent Variable: Views Per Person	Views per Person	Views per Person
	(Non-Elderly)	(Elderly)
	(1)	(2)
<i>A. Instrument=Share65+*Post</i>		
Post*Share65+	64.379 (50.69)	263.830* (138.34)
<i>B. Instrument=High Elderly Share*Post</i>		
High Elderly Share*Post	6.233*** (1.63)	18.055*** (5.04)
Mean of Dep. Var. (pre-Part D)	104.56	315.70
DMA x Quarter Obs	3,991	3,991

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the DMA level; the outcome variable is total advertising views at the DMA level; all specifications include quarter fixed effects and DMA fixed effects. Data is from 2001-2010.

Appendix Table B.4 – Baseline Regressions using Quartile of Elderly Share Instrument

Dependent Variable:	First Stage	Reduced Form	2SLS
	Views per Person (Non-Elderly)	# of Prescriptions	# of Prescriptions
	(1)	(2)	(3)
Quartile2 *Post	0.316*** (0.074)	0.001 (0.002)	
Quartile3 *Post	0.368*** (0.078)	0.001 (0.002)	
Quartile4 *Post	0.574*** (0.081)	0.009*** (0.002)	
Views per Person (Non-Elderly)			0.013*** (0.004)
F-statistic	20.72		
Mean of Dep. Var. (pre- Part D)	4.28	0.11	0.11
Zipcode x Condition x Quarter Obs	107,345	107,345	107,345

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the ZIP code level; the outcome variable is total advertising views or utilization aggregated by condition and zipcode; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. The excluded category is Quartile 1 (the lowest quartile of elderly share). Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. Data is from 2004-2010.

Appendix Table B.5 – Baseline Regressions using Alternative Functional Forms for Advertising Views

Specification:	Baseline: Views per Person	Log(Views per Person)	Log(Views per Person + 1)	IHS(Views per Person)
	(1)	(2)	(3)	(4)
<i>1. First Stage</i>				
High Elderly Share*Post	0.348*** (0.063)	0.070*** (0.012)	0.044*** (0.008)	0.052*** (0.009)
<i>2. 2SLS</i>				
f(Views per Person)	0.014*** (0.005)	0.069*** (0.024)	0.110*** (0.036)	0.093*** (0.031)
<i>3. Exposure Elasticity</i>				
<i>4. Expenditure Elasticity (Estimated)</i>	0.54	0.63	1.00	0.85
	0.23	0.27	0.43	0.37
Mean Views per Person (pre- Part D)	4.28			
Mean # of Prescriptions (pre- Part D)	0.11			
Zipcode x Condition x Quarter Obs	107,345	77,847	107,345	107,345

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the ZIP code level. Row 1 shows first stage results where the outcome variable is the specified functional form of advertising views: the level of DTCA, Log(DTCA) for non-zero observations, Log(DTCA+1), IHS(DTCA) which is the inverse hyperbolic sine transformation. Row 2 shows 2SLS results where the outcome is utilization and the independent variable is the specified functional form of advertising views. All specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Row 3 calculates the demand elasticity with respect to advertising views using the 2SLS estimates and means reported in this table. Row 4 converts Row 3 to an elasticity with respect to advertising expenditures; this calculation is discussed in Section 5.4. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. Data is from 2004-2010.

Appendix Table B.6 –Reduced Form Price Effects for Balanced Sample of Chronic Drugs, 2004-2007

Dependent Variable:	OOP Price	Ln(OOP Price+1)	Total Price	Ln(Total Price)
	(1)	(2)	(3)	(4)
<i>A. Instrument=Share65+*Post</i>				
Share65+*Post	-2.264 (12.254)	-0.005 (0.368)	19.251 (23.492)	-0.014 (0.130)
<i>B. Instrument=High Elderly Share*Post</i>				
High Elderly Share*Post	-0.818 (0.543)	0.009 (0.029)	-0.169 (1.192)	-0.008 (0.007)
Mean of Dep. Var (pre- Part D)	31.42	3.13	136.86	4.62
Zipcode x NDC x Quarter Obs	229,024	229,024	229,024	229,024

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the ZIP code level; the outcome variable is average out-of-pocket price or total price (or log transformations of these outcomes) by NDC and zipcode; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, NDC fixed effects. The sample is a balanced panel of NDCs from 2004-2007 for the 5 chronic conditions (i.e. each NDC has a non-missing observation in all quarters).

Appendix Table B.7 – Reduced Form Effects for Advertised Drug Classes vs. Non-Advertised Drug Classes

Dependent Variable:	# of Prescriptions	Log(# of Prescriptions)
A. Instrument=Share65+*Post		
Post*Share65+*Advertise	0.181*** (0.025)	0.761*** (0.130)
Post*Share65+	0.044*** (0.012)	0.679*** (0.183)
B. Instrument=High Elderly Share*Post		
Post*High Elderly Share*Advertise	0.011*** (0.002)	0.054*** (0.008)
Post*High Elderly Share	0.001 (0.001)	0.015 (0.009)
Mean of Dep. Var. (pre-Part D): Advertise=1	0.146	
Mean of Dep. Var. (pre-Part D): Advertise=0	0.075	
Zipcode x Drug Class x Quarter Obs	429,380	408,165

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all reduced form specifications include quarter fixed effects, 3-digit ZIP code fixed effects, an Advertise indicator (equals 1 when the drug class has any positive amount of advertising during the study period, zero otherwise), Post x Advertise, and Advertise x Share 65+ (or Advertise x High Elderly Share). These variables have been omitted from the table to conserve space. The sample includes the top 10 advertised drugs classes and top 10 non-advertised drug classes (see Appendix Figure B.6 for the full list of drug classes) and is at the Zipcode x Drug Class x Quarter level of observation. This is a different set of drug classes than what is shown in our main results in Table 2 for the 5 chronic conditions. Note that the estimated effect for non-advertised classes is Post*HighElderlyShare and the effect for advertised classes is the linear combination of Post*HighElderlyShare + Post*HighElderly Share*Advertise (Panel A is defined similarly). For example, in Panel B, the differential effect on utilization for non-advertised drug classes is 0.001, while advertised classes is 0.012 [0.001+0.011]. Data is from 2004-2010.

**Appendix Table B.8 – Total Utilization of Chronic Drugs, 2SLS—
Extensive and Intensive Margins**

Dependent Variable:	# of Prescriptions		Days Supply		Any Use	
	Full Sample	2004-2007	Full Sample	2004-2007	Full Sample	2004-2007
	(1)	(2)	(3)	(4)	(5)	(6)
A. Instrument=Share65+*Post						
Views per Person (Non-Elderly)	0.017*** (0.004)	0.011*** (0.003)	0.750*** (0.217)	0.560*** (0.131)	0.006** (0.002)	0.004*** (0.001)
B. Instrument=High Elderly Share*Post						
Views per Person (Non-Elderly)	0.014*** (0.005)	0.008*** (0.003)	0.480** (0.223)	0.400*** (0.127)	0.003 (0.002)	0.003** (0.001)
Mean of Dep. Var (pre- Part D)	0.11		5.68		0.06	
Zipcode x Condition x Quarter Obs	107,345	61,440	107,345	61,440	107,345	61,440

Dependent Variable:	# of Prescriptions		Days Supply	
	Conditional on Use		Conditional on Use	
	Full Sample	2004-2007	Full Sample	2004-2007
	(7)	(8)	(9)	(10)
A. Instrument=Share65+*Post				
Views per Person (Non-Elderly)	0.057*** (0.017)	0.035*** (0.012)	2.151*** (0.533)	1.428*** (0.406)
B. Instrument=High Elderly Share*Post				
Views per Person (Non-Elderly)	0.068*** (0.026)	0.028* (0.016)	1.861*** (0.681)	1.366*** (0.523)
Mean of Dep. Var (pre- Part D)	1.81		90.87	
Zipcode x Condition x Quarter Obs	100,427	58,624	100,427	58,624

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. We estimate each specification separately for the full sample and the pre-recession years 2004-2007. The total effects (which are presented in Cols 1-4) include zeros for those who do not purchase any chronic drugs. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Appendix Table B.9— Decomposition of Extensive and Intensive Margin Effects, 2SLS

Dependent Variable:	# of Prescriptions		Counterfactual # of Prescriptions		% of Total Effect is Extensive Margin	
			<i> Holding Extensive Margin Constant </i>		<i> 1-(3)/(1) 1-(4)/(2) </i>	
	Full Sample	2004-2007	Full Sample	2004-2007	Full Sample	2004-2007
	(1)	(2)	(3)	(4)	(5)	(6)
A. Instrument=Share65+*Post						
Views per Person (Non-Elderly)	0.017*** (0.004)	0.011*** (0.003)	0.005*** (0.001)	0.003*** (0.001)	71%	73%
B. Instrument=High Elderly Share*Post						
Views per Person (Non-Elderly)	0.014*** (0.005)	0.008*** (0.003)	0.006*** (0.002)	0.003* (0.001)	57%	63%

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Columns 1 and 2 repeat the main baseline results for total number of prescriptions purchased; Columns 3 and 4 show the effect on total prescriptions purchased coming from intensive margin changes (i.e. assuming that there are no changes in the extensive margin) as described in Appendix A; Columns 5 and 6 compute the percentage of the effect on total prescriptions purchased due to extensive margin effects. Includes the 50 drugs that advertised during the study period for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Appendix Table B.10— Adherence of Chronic Drugs – Mean Medication Possession Ratio

Dependent Variable:	Medication Possession Ratio	Reduced Form			2SLS		
		Full Sample	2004-2007	2005-2007	Full Sample	2004-2007	2005-2007
		(1)	(2)	(3)	(4)	(5)	(6)
A. Instrument=Share65+*Post							
	Post*Share65+	0.195*** (0.067)	0.219*** (0.073)	0.414*** (0.140)			
	Views per Person (Non-Elderly)				0.018** (0.008)	0.016** (0.007)	0.034** (0.015)
B. Instrument=High Elderly Share*Post							
	High Elderly Share*Post	0.005* (0.003)	0.007** (0.003)	0.012*** (0.004)			
	Views per Person (Non-Elderly)				0.009* (0.005)	0.009** (0.004)	0.020** (0.009)
	Mean of Dep. Var (pre- Part D)	0.75	0.75	0.71			
	Zipcode x Condition x Quarter Obs	102,477	59,252	44,519	102,477	59,252	44,519

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. The outcome variable is the medication possession ratio (MPR). Includes the 50 drugs that advertised during the study period and the drugs that did not advertise for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Appendix Table B.11 – Adherence of Chronic Drugs: Alternative Specifications, 2SLS

Dependent Variable: I(High Adherence)	Instrument= Share65+*Post			Instrument= High Elderly Share*Post		
	Full			Full		
	Sample	2004-2007	2005-2007	Sample	2004-2007	2005-2007
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. All Individuals ages 40-60</i>						
1. Baseline Specification	0.017** (0.007)	0.017*** (0.005)	0.033*** (0.013)	0.008 (0.005)	0.011*** (0.004)	0.021** (0.008)
2. Alternative MPR (Including Discontinuation)	0.016* (0.008)	0.021*** (0.006)	0.037*** (0.013)	0.006 (0.007)	0.016*** (0.005)	0.026*** (0.009)
<i>B. Excluding Individuals who Initiated Treatment after Part D</i>						
3. Baseline Specification	0.022*** (0.006)	0.020*** (0.005)	0.036*** (0.012)	0.014*** (0.004)	0.013*** (0.004)	0.023*** (0.008)
4. Alternative MPR (Including Discontinuation)	0.030*** (0.008)	0.025*** (0.005)	0.041*** (0.013)	0.023*** (0.006)	0.019*** (0.005)	0.031*** (0.009)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. The outcome variable is the proportion of individuals with MPR>=80%. Each cell represents a separate regression with the coefficient on “views per person (non-elderly)” reported. Includes the 50 drugs that advertised during the study period and the drugs that did not advertise for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis.

Appendix Table B.12 – Effects on Total Chronic Drug Spending

Dependent Variable:	Total Expenditures	Reduced Form			2SLS		
		Advertised	Non-	Total	Advertised	Non-	Total
		Drugs	Drugs		Drugs	Drugs	
		(1)	(2)	(3)	(4)	(5)	(6)
A. Instrument=Share65+*Post							
Post*Share65+		12.488*** (3.399)	5.878*** (2.033)	18.367*** (3.801)			
Views per Person (Non-Elderly)					1.964*** (0.537)	0.925** (0.424)	2.889*** (0.779)
B. Instrument=High Elderly Share*Post							
High Elderly Share*Post		0.532*** (0.205)	0.092 (0.121)	0.624** (0.248)			
Views per Person (Non-Elderly)					1.526*** (0.576)	0.265 (0.350)	1.790** (0.716)
Mean of Dep. Var (pre- Part D)		14.61	4.75	19.36			
Zipcode x Condition x Quarter Obs		107,345	107,345	107,345	107,345	107,345	107,345

Notes: *** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors at the 3-digit ZIP code level; all specifications include quarter fixed effects, 3-digit ZIP code fixed effects, condition fixed effects. Includes the 50 drugs that advertised during the study period and the drugs that did not advertise for 5 chronic conditions: depression, diabetes, hyperlipidemia, hypertension, and osteoporosis. Data is from 2004-2010.