NBER WORKING PAPER SERIES

THE EFFECTS OF OFF-LABEL DRUG USE ON DISABILITY AND MEDICAL EXPENDITURE

Katharina E. Blankart
Frank R. Lichtenberg

Working Paper 30440
http://www.nber.org/papers/w30440

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2022

We acknowledge funding by the Information Technology and Innovation Foundation (itif.org) and the Leibniz Science Campus Ruhr. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Effects of Off-label Drug Use on Disability and Medical Expenditure
Katharina E. Blankart and Frank R. Lichtenberg
NBER Working Paper No. 30440
September 2022
JEL No. I1,L65

ABSTRACT

Does using prescription drugs off-label increase disability and medical expenditure? This paper uses a unique dataset to evaluate off-label vs. on-label drug use in the US non-institutionalized population. Patients using drugs off-label have on average $515 higher medical expenditure and work-loss cost. Pharmaceutical innovation has direct and indirect effects on off-label drug use. Market size is indicative of the fraction of treatments used off-label. Our findings have implications for regulation and welfare. We address endogeneity issues by demonstrating that patients with higher disease severity do not experience higher off-label uses and by controlling for unobserved individual and condition effects.

Katharina E. Blankart
University of Duisburg-Essen
Weststadttuerme, Berliner Platz 6-8
45127 Essen
Germany
katharina.blankart@uni-due.de

Frank R. Lichtenberg
Columbia University
Kravis Hall 522
665 West 130th Street
New York, NY 10027
and NBER
frl1@columbia.edu

The effects of off-label drug use on disability and medical expenditure

By KATHARINA BLANKART AND FRANK R. LICHTENBERG *

Does using prescription drugs off-label increase disability and medical expenditure? This paper uses a unique dataset to evaluate off-label vs. on-label drug use in the US non-institutionalized population. Patients using drugs off-label have on average $515 higher medical expenditure and work-loss cost. Pharmaceutical innovation has direct and indirect effects on off-label drug use. Market size is indicative of the fraction of treatments used off-label. Our findings have implications for regulation and welfare. We address endogeneity issues by demonstrating that patients with higher disease severity do not experience higher off-label uses and by controlling for unobserved individual and condition effects.

* Blankart: Leibniz Science Campus Ruhr and CINCH Health Economics Research Center and Faculty of Business Administration and Economics, University of Duisburg-Essen, Berliner Platz 6-8, 45127 Essen, Germany. Email: (katharina.blankart@uni-due.de). Lichtenberg: Graduate School of Business, Columbia University in the City of New York, Kravis Hall 522, 665 West 130th Street, New York, NY 10027USA. National Bureau of Economic Research and CESifo, Email: (frl1@gsb.columbia.edu). We would like to thank Pierre Koning, Helge Liebert, Marisa Miraldo, Carol Propper, David Bradford and the participants of the 2022 editions of the Essen Health Conference, the American European Health Economics Workshop in Barcelona, the German and Swiss Health Economics Society meetings, the Imperial College Centre for Health Economics & Policy Innovation seminar, the European Health Economics Association conference in Oslo for comments. We acknowledge funding by the Information Technology and Innovation Foundation (itif.org) and the Leibniz Science Campus Ruhr.

Off-label use of prescription drugs is an integral part of modern medicine today and accounts for a substantial proportion of pharmaceutical expenditure (Dresser and Frader 2009). 17–40% of prescription drugs in ambulatory care are not prescribed for approved indications, i.e. health conditions for which the FDA or another regulatory agency have approved their use (Bradford, Turner,

and Williams 2018; McKibbin 2020; Radley, Finkelstein, and Stafford 2006). One third of prescriptions for antibiotics have been attributed to inappropriate use in the United States, contributing to a rise in antibiotic resistance (Fleming-Dutra et al. 2016). If a drug is not used for conditions for which it is considered safe and effective through regulatory approval, such off-label use may be inappropriate to adequately treat a condition. It could be even harmful to patients (Dresser and Frader 2009). Although off-label use is sometimes unavoidable when it is the best available option for a patient, off-label use frequently occurs with little or no scientific support (Lat et al. 2011; McKibbin 2020; Radley, Finkelstein, and Stafford 2006).

Reasons why off-label use exists are tied to economic factors and the role of technological progress. Clinical trials are sometimes difficult or costly and incentives to provide solid evidence for an already licensed drug may be low (Dresser and Frader 2009). Some discoveries about a mechanism of action of an active pharmaceutical ingredient (API) are made post-approval when incentives to license a new indication have decreased. Finally, there may be strategic licensing of drugs, and drug prices typically do not differ by indication. These factors raise questions about the overall value of off-label drug use compared to using drugs according to their designated label. Patients and physicians may face tradeoffs when making decisions about prescription drug use in terms of health care use that results in disability.

To contribute to our understanding of the consequences of off-label use, the role of technological progress, and why off-label use is more frequent in some drug markets than others, we combine and evaluate a novel data set in which we observe on- and off-label uses of prescription drugs using a nationally representative sample of 201,489 American patients and 247 conditions in the United States, primarily derived from the Medical Expenditure Panel Survey (MEPS), 1996-2015. We use comprehensive marketing authorization data that links products defined by an API to their on-label use case for a medical condition relying on a large pharmaceutical reference database covering the

universe of approved products in France. We link those data to the universe of APIs available as authorized by the US FDA to investigate the fraction of drugs used off-label at condition-by-individual level in the ambulatory prescription drug market. The data allow us to evaluate the consequences and costs of off-label use and to examine the economic drivers of off-label use. Multiple conditions per individual observed is essential to our analysis, because it allows us to isolate the influence of off-label use of prescription drugs compared to on-label uses within individuals and conditions.

In our empirical approach, we estimate two-way fixed effects regressions by individual and condition to identify the effect of off-label use and drug vintage (FDA approval year) on health care use and disability. The evidence to date relies on characterizing determinants of off-label uses and capturing their associations. The main outcomes that have been studied are adverse drug events and price differentials between off- and on-label drugs of the same class. Our approach enables us to control for the effects of attributes of an individual that are invariant across his or her medical conditions such as age, sex, race, education, income, region, overall health status or behavioral factors[1]. In addition, we control for the unobserved effects related to attributes of a medical condition that are invariant across individuals and related to condition severity. We control for the mean vintage (FDA approval year) of the drugs used by a person to treat his or her medical condition to reflect variation in average quality by approval status within a condition.[2] We also investigate the effects of medical condition prevalence and drug approval year on the rate of off-label use. We perform heterogeneity analysis for population groups that have experienced a larger exposure to off-label use and about whom there is little scientific evidence from clinical trials.

[1] Most medical conditions are borne by, and most medical care is provided to, people with multiple medical conditions (comorbidities).

[2] Using aggregate (longitudinal disease-level) data, Lichtenberg (2014) showed that increases in mean drug vintage were associated with fewer disability days and hospitalizations.

In general, a potential pitfall of using observational data is unobserved heterogeneity that may be correlated with the observed treatment variable. However, we can control for unobserved individual and condition heterogeneity. Another potential pitfall is that our measure of off-label use may be subject to measurement error. Notably, our data rely on self-reports of individuals that suffer from multiple conditions and may not correctly recall off-label use of drugs. In addition, there is concern that off-label uses may serve as last resort of treatment. Off-label drug may then be considered in patients of higher disease severity. We assess these concerns and show that our estimates withstand several specification tests.

This paper relates to the literature that studies the consequences of off-label uses. We use comprehensive population-based repeated cross-sectional survey data, not observational data from administrative sources that has been used in the medical literature to primarily study adverse drug events resulting from off-label uses in selected diseases, settings or populations (Eguale et al. 2016; Yackey et al. 2019). The combined dataset we analyze provides condition-specific information about each person's disability days and use of six types of medical care (inpatient events, emergency room events, etc.) linked to a condition, which allows us to study consequences beyond adverse drug events. We contribute to a smaller set of studies that has dealt with economic consequences of off-label use considering price differentials and prescribers (Bradford, Turner, and Williams 2018; McKibbin 2020; Molitor 2012; Tunçel 2020). In contrast to previous evidence that studied off-label use in selected conditions or patient sub-groups (Bradford et al. (2018) and McKibbin (2020)), we consider a wide array of health conditions and outcomes to quantify and determine drivers of off-label use related to market size and innovative activity by condition. We further account for differences in quantity in prescription volumes of on- compared to off-label drugs to more accurately reflect differences in medical expenditure and subsequent welfare considerations.

We advance the study of productivity and welfare losses in health care markets caused by off-label compared to on-label use of prescription drugs by. To be most effective, a technology should be used by consumers for what it is designated for and has a proven quality. Productive inefficiency may be due to use of the wrong combination of inputs or the wrong technology (Glied and Sacarny 2018). Off-label use may be one reason why treatments do not work and why some patients experience no change in health status after having received a certain treatment. Murphy et al. (2020) demonstrated that medication used inappropriately by Alzheimer's patients was associated with a greater risk of mild to severe adverse events, unscheduled hospital visits and general practitioner visits. For prescription drugs, biopharmaceutical innovation has contributed to increasing economic growth and longevity substantially (Bryan and Williams 2021; Lichtenberg 2022; 2020; 2014) and, likely to a much larger extent than other non-pharmaceutical interventions (Buxbaum et al. 2020).

Our paper is closely related to work aimed to evaluate the role of market size and innovation in pharmaceutical markets. Previous studies have demonstrated that research and development of new drugs is more intensive in larger markets (Acemoglu and Linn 2004; Blume-Kohout and Sood 2013). Blume-Kohout and Sood (2013) suggest that research and development activities increased substantially in disease areas where markets expanded through Medicare Part D insurance coverage extension.

I. Background

A. Off versus on-label use of drugs

Off-label use refers to any intentional use of an authorized product not covered by the terms of regulatory approval of the US Food and Drug Administration (FDA) or similar regulators. To qualify for on-label use, manufacturers need to demonstrate that a drug is safe and effective for a certain

indication[3]. Such evidence is typically based on randomized controlled trials that compare patients receiving the on-label drug to the best available therapeutic option or a placebo treatment. Although drugs may be legally prescribed outside their designated label, advertisement of off-label uses is banned. Manufacturers therefore have incentives to perform clinical trials to demonstrate the use case of a drug for a condition (Richardson 2016), although incentives may differ by condition and time horizon to demonstrate effectiveness (Budish, Roin, and Williams 2015). If a drug is authorized, but used outside the scope of its designated label, such prescriptions are defined as off-label use. Among health care providers, using the right drug for the right patient is embedded in the five rights of a patient in administering a drug besides the right time, the right dose, and the right route (Grissinger 2010). Courts have reinforced that off-label use by physicians is a legally viable option if considered suitable.[4] Of note is that we do not consider any illicit or illegal use of drugs but focus on prescribed medicines.

Off-label use of drugs is widespread and varies between 0 and 100% by condition and API. A report of the European Union documents that among prescribers, prevalence of off-label use is higher than 55% of physicians (Weda et al. 2017). For the United States, Bradford et al. (2018) estimate off-label use to account for 29-38% according to the prescriber-based National Ambulatory Medical Care Survey. For cancer drugs, Molitor (2012) estimates off-label use rates at 20%, and McKibbin (2020) at 17.4% of prescriptions. Tunçel (2020) estimated that 21% of depression drugs were prescribed off-label in France.

[3] In medicine, an indication is a valid reason to use a certain test, medication, procedure, or surgery. "Indication - Medical Definition and More from Merriam-Webster". Archived from the original on 14 July 2011. Retrieved 14 December 2010 ("Indication - Medical Definition and More from Merriam-Webster" 2011).

[4] In contrast to off-label use, inappropriate use is broader as it refers to uses of drugs in (sub-)groups of patients where clinical trial and additional evidence does not support its use for that subgroup within its designated marketing authorization. Inappropriate use is often defined after marketing authorization in clinical guidelines or similar recommendations by physician professional groups. Inappropriate use is widespread in the elderly population. Whether use of a drug is appropriate depends on factors such as target population, dosing, mode of administration and counter-indications (Awad and Hanna 2019). As our data will essentially reflect use of APIs according to their designated label, we will concentrate on off-label compared to on-label uses, although the concept of inappropriate use often includes and expands information from marketing authorization.

Similar to previous approaches we define off-label use as any use of an API approved for one condition, but used to treat any other condition (Bradford, Turner, and Williams 2018; 2020; Molitor 2012). We capture off-label use by relying on data about APIs labelled for a condition. Drug label use is different as it restricts the use of an API for a certain approved indication. Drugs with identical API but labeled for different conditions can often be used interchangeably.[5] However, using the identical active-ingredient in another condition may be thus anticipated (Bradford, Turner, and Williams 2018). We aim to deal with unanticipated off-label use when an API is used outside its designated label and has not been labelled for another condition.

The previous evidence on the factors that relate to economic incentives and availability of scientific evidence suggests that it use off-label use is higher when there are few other options available to treat a condition and when there are less restrictive reimbursement regulations that may financially discourage off-label uses (Bradford, Turner, and Williams 2018). One critical factor is the unavailability of scientific evidence. Molitor (2012) suggested that off-label uses of cancer drugs are higher in regions where physicians are closer to a clinical trial unit. McKibbin (2020) suggested that the off-label use increases by 85% if there are any randomized controlled trials showing that an off-label use of a cancer drug improves patient survival. Regarding welfare effects, Bradford et al. (2018) and Tunçel (2020) suggest that off-label use could be welfare enhancing by using less expensive off-label drugs when there are similar on-label drugs. Tunçel (2020) suggests that expenditures would increase by 15% without significant changes in health outcomes when replacing on-label by off-label uses in depression treatments. Howevere, these studies do not consider any non-pharmaceutical medical expenditure or disability in their welfare considerations.

[5] For example, the beta-blocker timolol has different ATC codes when used as a cardiovascular drug (C07AA06) and as a treatment for glaucoma (S01ED01) (Bodenreider and Rodriguez 2014).

There may be desirable off-label uses where manufacturers simply have not requested marketing authorization from the FDA beyond the primary label, but for which there may be supportive scientific from clinical trials or daily clinical routine. However, the extent and amount of evidence available for off-label uses is typically not on par with the extensive data required for regulatory approval by the FDA (Dresser and Frader 2009). Studies that capture the scientific evidence base of off-label drugs used in practice suggest that a minority of uses are supported by scientific evidence. Analyses of prescription patterns suggest that 61–84% of off-label uses in ambulatory care (Radley, Finkelstein, and Stafford 2006) and 48% of choices in intensive care unit settings (Lat et al. 2011) lack scientific support. For cancer treatments, McKibbin (2020) shows that for 12 unique drugs used off-label, of nine drugs covering eight unique diseases, 75% were ever FDA approved and showed positive trial results. Another ten drugs for ten diseases had negative trial results, of which 13.6% were ever FDA approved. For the majority of drugs, the efficacy of the additional uses compared to the first approved use was worse compared to initial uses. Once there is scientific support of positive efficacy of an off-label drug, demand increases by 85%, on average.

B. Health care use, disability and drug approval

Studies of the inefficiencies in supply and use of health care suggest that there are substantial proportions of medical expenditure that can be attributed to low value, or unwarranted forms of health care (Chandra and Skinner 2012). An extreme form of low-value care is a therapy that is not appropriate as treatment for the condition according to its designated label. Here, off-label use of drugs, for which frequently there is no supportive scientific evidence may contribute to the discussed productivity losses.

To account for productivity losses, off-label use may increase medical costs, and make patients less healthy, even when it does not cause adverse drug events (ADEs). A drug may have either a positive effect on a patient's health, no effect,

or a negative effect (Figure 1). We expect on-label drug use to have a positive effect on the patient's health, because the drug's safety and effectiveness has been demonstrated. Regarding off-label use, most previous evidence like Eguale et al. (2016) consider only one disadvantage of off-label use: ADEs. But using off-label drugs can have undesirable consequences even if they do not harm you. Off-label use is likely to have either no effect on the patient's health, or a negative effect, for example, due to an ADE. Even if most off-label use has no effect, use of off-label drugs in lieu of on-label drugs may be undesirable—there may be an opportunity cost if individuals substitute safe but ineffective treatments for safe and effective ones. This is the counterfactual we study. Taking an off-label drug may be like taking a placebo instead of a safe and effective drug for the same condition.

–Figure 1–

Eguale et al (2016) provide evidence that off-label prescribing may lead to adverse drug events (ADEs), defined as discontinuations of drug use made by physicians owing to an adverse drug an allergic reactions. The study of 46,021 patients at primary care clinics in Québec, Canada, estimated that the ADE rate was 44% higher for off-label than it was for on-label use, controlling for drug class, drug age, patient age and sex, measures of comorbidity, polypharmacy, and continuity of care. The cost of ADEs was not measured, but using previously published figures, the authors estimated that the mean cost of ADEs per patient was between $51 and $77. Hence, off-label use may have increased the mean cost of ADEs per patient in Québec by between $22 (=44%*$51) and $34 (=44%*$77).[6] Average off-label use (versus no off-label use) increased mean cost of ADEs per patient by $2-$4, the only monetary value identified in the literature.

[6] U.S. per capita health spending is about twice as high as Canadian per capita health spending (10,949 USD vs. 5370 USD), so off-label use may increase ADE-related costs per patient in the U.S. by about twice as much.

C. Market size, innovation and off-label use

To study why off-label use varies across conditions, we will investigate two determinants: an attribute of the medical condition (market size) and an attribute of the drug (vintage captured by FDA approval year). One reason for variation in the number of drugs available with a designated label for a condition is market size. Incentives to earn profits for biopharmaceutical development are higher in larger markets and thus are the driving force for innovation. Acemoglu and Linn (2004) demonstrated for the pharmaceutical industry that exogenous changes in demographics lead to a higher frequency of product entries. Blume-Kohout and Sood (2013) suggest that market size expansion through implementation of Medicare Part D has increased innovation in therapeutic classes with higher Medicare market shares.

Previous medical and epidemiological studies have identified several groups that appear to be more vulnerable to off-label use: children and adolescents, individuals suffering from rare diseases, (pregnant) women and, prisoners (Weda et al. 2017). All of these groups have in common that the market size may be relatively low. Moreover, there are federal regulations that require additional measures for performing clinical trials in these populations and which has discouraged the testing of medication in children (Yackey et al. 2019).

Similarly, certain population groups have been underrepresented in clinical trials. The strength of the evidence base for these groups may be lower in demonstrating efficacy. Especially, race, sex and gender differences are known as sources of variability in how drugs work given differences for example in metabolism and hormonal status (Mauvais-Jarvis et al. 2021). Investments in clinical trials vary by market size such that off-label use may be more pronounced and may have stronger consequences on patient health when population groups are small and differences in health outcomes are costly to demonstrate (Budish, Roin, and Williams 2015). If a market is too small, the

cost of an additional clinical trial to demonstrate efficacy in the conditions that are off-label use may exceed the expected value (McKibbin 2020).

The extent to which a drug is used off-label may also depend on its vintage (FDA approval year) and on how long it has been on the market (Figure 2). For newer drugs, there may be more intensive promotional activity compared to older drugs based on the timing of approval (Kremer, Moritz, and Siemsen 2011). For example, Duflos and Lichtenberg (2012) demonstrate that promotional activity that educates physicians about the on-label uses of drugs is highest around years 8 to 12 after approval with a strong decline after generic competitors have entered the market. Controls like prior authorization may be stronger by health insurances in newer drugs, which are typically more expensive (Bradford, Turner, and Williams 2018).

II. Data and Descriptive Evidence

To evaluate off-label use of prescription drugs across many conditions and individuals, measures of off-label use and approval year to reflect average quality are required based on regulatory data along with linked health care use and disability data. An important requirement is that reported prescription drug uses and other health care uses link to the same condition where off-label use takes place. We introduce each of the data sources used below, our measures of off-label use and vintage, and provide descriptive evidence in the US non-institutionalized population.

A. Individual level prescription drug, health care use and disability data

Medical Expenditure Panel Survey. We use data obtained from a comprehensive household survey of the US non-institutionalized population, the Medical Expenditure Panel Survey (MEPS) (J. W. Cohen, Cohen, and Banthin 2009; S. B. Cohen, Ezzati-Rice, and Yu 2006). It is a distinctive feature of MEPS that it provides data on disability and prescription drugs and other medical care use, *by person and by medical condition*. We can capture what

type of prescription medicine an individual used to treat a condition. We use the 1996-2015 prescription drug and medical condition files, including the linkages between conditions and prescription drugs. MEPS provides data on the individual-reported medical condition that each prescription was used to treat. We obtained data regarding prescription drugs through the household questionnaire and a pharmacy follow-back component within the Medical Provider Component. MEPS primarily contains information reported by survey respondents, so we cannot observe or infer any decision-making by health care providers. MEPS is unique as the vast majority of datasets on prescription drug use such as insurance claims do not provide information about drug indications.

We used all reported medical conditions that had links to a prescription drug using the condition-event link files provided by MEPS.[7] In terms of health care use these are the number of: inpatient events (variable IPNUM in the MEPS), emergency room visits (ERNUM), home health events (HHNUM), office-based visits (OBNUM), outpatient events (OPNUM), and prescription medications (RXNUM). Regarding disability associated with a condition, we capture whether the individual spent any days in bed (INBEDFLG), missed any school days (MISSSCHL), and missed any workdays (MISSWORK).

Pharmaceutical reference data of on-label drug uses. To study off-label use in prescription drugs, we rely on data that combines product information by API and the labelled conditions for which the API is designated across the universe of products available. Regulatory data typically specify the condition(s) in which clinical trial evidence has shown the API to be effective for approval. As the FDA does not provide linked data in a useful format, we rely on an independent data source from France, Thériaque which is edited by the Centre National Hospitalier d'information sur le Médicament (CNHIM) (Husson

[7] 62% of condition events were linked to prescription medicines and other health related events whereas 29% of condition events were linked to a prescription drug only but no other events. The remaining 9% of condition events were linked to other medical events besides prescriptions that we did not consider. Our analyses therefore concentrate on individuals that report a condition for which they used prescription drugs as a treatment.

2008). Besides other information, Thériaque includes data on all pharmaceutical products marketed and used in France, coded by the World Health Organization's Anatomical Chemical Therapeutic Classification (ATC classification) and International Statistical Classification of Diseases and Related Health Problems (ICD-10). This source identifies the approved indications of every product. We used the 2017 version of Thériaque, which includes 1,463 5th-level ATC codes and 1,030 conditions by ICD-10 level, resulting in 6,183 combinations of APIs and conditions by ICD-10 level 3. That way our classification of off-label uses reflects the status of labelled ATC and conditions in 2017 such that we do not capture effects on newly introduced on-label drugs.

Linking conditions and prescription drug files. To link the classification of off-label use by API and condition based on Thériaque to prescription data provided by MEPS in the United States, we used several auxiliary data sets. Thériaque includes WHO ATC codes and ICD disease codes that are not readily available in US-based prescription data at the product level. For this reason, we extracted historical (for the years 1996 to 2009) and current (2010-2015) product information based on the FDA's National Drug Code (NDC) directory, which includes information on the initial marketing authorization application of each pharmaceutical product. To obtain historical information, we used a search engine of archived web pages to extract versions of the NDC directory from previous versions of the FDA's website. For current information, we relied on data on all products by NDC documented in Drugbank (Wishart et al. 2006). To link data of regulatory approvals to ATC codes and the earliest year of approval by API, we relied on DrugCentral (https://drugcentral.org/). This approach led to 87% of products linked to ATC codes in the American data.

To link conditions based on ICD-10 to the classification of conditions based on diagnoses of the Clinical Classification System (CCS) as provided in MEPS,

we relied on the crosswalks provided of the Agency for Healthcare Research and Quality[8].

For 1996-2015, we extracted data on 45,479 products by NDC that correspond with 1,768 5th-level ATC codes and 247 conditions by the CCs classifications. We excluded CCS codes not related to medical conditions, for example referring to "Administrative/social admission" and "Residual Codes Unclassified". We aggregated observations to the condition by individual level. The final analysis dataset contains 553,302 observations for 201,489 individuals using prescription medications. Linking Thériaque to MEPS covers 80% of prescriptions documented in the MEPS with coverage at 89% in 2015 and at 76% in 1996.

Data preparation and analyses were performed using SAS version 9.4. Two-way fixed effects regressions were performed using the reghdfe package in Stata /MP 16.0.

B. Measuring off-label use based on pharmaceutical reference data

We define an API to be on-label to treat a condition if this unique ATC/ICD-10 combination is listed in Thériaque. Off-label use is present when an API – medical condition combination is used that is not listed in Thériaque. It is the residual of all on-label uses and captures the extent to which a potentially inappropriate technology is used (Glied and Sacarny 2018). We measure the fraction of prescriptions used by individual i for medical condition c whose APIs have been approved for the treatment to measure off-label use as follows:

$$\text{off-label}\%_{ic} = 1 - \text{on-label}\% = 1 - \frac{\sum_d \text{N\_Rx}_{\text{idc}} \text{INDIC}_{\text{dc}}}{\sum_d \text{N\_Rx}_{\text{idc}}} \qquad (1)$$

N_Rx_{dic} equals the number of prescriptions for API (5th-level WHO ATC)[9] d used by individual i for medical condition c. The parameter INDIC_{dc} is based

[8] https://www.hcup-us.ahrq.gov/toolssoftware/ccsr/ccsr_archive.jsp#return

[9] In case there are two or more WHO ATC codes at level 5, we assign them to all eligible conditions. A prominent example for off-label use of the same API with different dosage form and dose is.

on the data obtained in Thériaque and takes the value of 1 if API d has been approved for the treatment of medical condition c and 0 if API d has not been approved for the treatment of medical condition c. A share of 100 percent of off-label$\%_{ic}$ indicates that all of the prescriptions used by individual i to treat condition c were off-label. A share of 0% indicates that all of prescriptions used to treat condition c were on-label which refers to the counterfactual we analyze.

The approach to capture off-label use may be subject to measurement error. All sources of error we identify for off-label% would make our measure conservative (Wooldridge 2002). A disadvantage is that we rely on patient-reported information regarding the type of condition. Individuals may confuse conditions and medications. We will provide descriptive evidence about the relationship between off-label use and the number of conditions an individual reports.[10]

Measurement error may arise as we apply the French classification of drug labels to medication usage in the United States. Labels as defined in Thériaque may deviate from drug approvals by the FDA. We assume that the approved condition-drug pairs within the scope of Thériaque are equally relevant for drug uses in the United States. Using the 2017 version of Thériaque, we assume that all documented on-label uses were fully known during our study period. That means that a drug that was prescribed off-label in 1996 that has subsequently received approval for that condition is considered as on-label use.

To ensure that our sample of ATC/ICD combinations represents a significant amount of similar pairs approved by the FDA, we validated overlap using the subset of cancer drugs and their approved indications based the National Cancer Institute's data[11]. 53% of drug-condition pairs were equally present in

[10] Medical conditions reported by the Household Component respondent were recorded by the interviewer as verbatim text, which was then coded by professional coders to fully-specified ICD-9-CM codes. ICD-9-CM condition codes were then aggregated into clinically meaningful categories that group similar conditions (CCS). CCS was generated using Clinical Classification Software, which aggregates conditions and V-codes into mutually exclusive categories, most of which are clinically homogeneous (Agency for Healthcare Research and Quality (2022a))

[11] https://www.cancer.gov/about-cancer/treatment/drugs/cancer-type, last accessed March 10, 2022

Thériaque (Lichtenberg 2020). Discordance in approval decisions is often due to rejection of efficacy based on the same clinical trial evidence but diverging indications which may explain difference in ATCs and conditions covered (Kashoki et al. 2020). The extent of discordance in approvals between the US and France we find is similar drug label comparisons between the FDA and the European European Medicines Agency, which is responsible for drug approvals in France (Tafuri 2013).[12]

Finally, a source of measurement error may originate from use of the (CCS) categorization of 247 mutually exclusive diagnosis categories that is broader than the conditions specified in the label of an approved drug. We capture an upper bound of on-label use as we consider a prescription drug to be on-label if it has an indication for any ICD-10 code within a CCS medical condition. There may be additional clinical criteria (e.g. age, sex, clinical parameters) that may rule out the use of the API in sub-populations. Our measurement of off-label use is conservative as it captures on-label uses of drugs for a condition when in fact it is an off-label sub-categories of a condition.[13]

C. Defining vintage of prescription medication

The second variable of interest is the mean vintage of drugs used. Combining data about the earliest approval year of the API based on FDA data, we can determine the mean vintage of the APIs used by individual i for medical condition c, weighted by the number of prescriptions used of that drug N_Rx_{dic}, as:

$$vintage_{ic} = \frac{\sum_d \text{N\_Rx}_{\text{dic}}\, approval\_year_d}{\sum_d \text{N\_Rx}_{\text{dic}}} \qquad (2)$$

[12] Comprehensive validation of drug-indication pairs across a large number of drugs and conditions is difficult as we could not identify a source that structurally links FDA labels to APIs and conditions.

[13] Guidance of appropriateness criteria beyond the condition is provided by clinical guidelines or additional classifications such as the Beer's criteria for pharmaceutical treatments in the elderly, which we do not consider (Beers et al. 1991; Hanlon and Schmader 2013). Thus, an API having a designated label is a necessary, but not a sufficient condition for a treatment to be on-label.

The vintage measure may reflect the average quality of the drugs used for treating condition c. It also reflects the approval status of the drug within the US. Comparisons of drug vintages have been used earlier to account for the possibility that the average quality of a newer drug may be superior compared to an older drug within the same class (Jovanovic and Yatsenko 2012; Lichtenberg 2014). For relatively new drugs, vintage is indicative of whether promotional activity takes place and of whether insurers implement controls such as step-therapy or prior authorization, which most often apply to newer, more expensive, patent-protected drugs.

D. Descriptive evidence of off-label use

Between 1996 and 2015, average off-label use decreased from 59.8% to 43.7%. In terms of the level, the fraction of off-label use identified by Bradford et al. (2018) is somewhat lower, but that approach linked any condition listed to an eligible on-label API. In our approach, we relied on linked API-condition pairs reported by individuals.

Across the same period, the average vintage of drugs used increased from 1973 to 1980. The left panel of Figure 3 shows the mean fraction of drugs used on-label by condition in 1996 and 2015, ranked by the average fraction of on-label use across all individuals diagnosed with that condition. Between 1996 and 2015, the number of conditions where 100% of prescription drugs use is off-label has reduced from 71 to 48 conditions. This provides first evidence that the number of conditions where individuals had no other choice than using a drug off-label has decreased substantially, potentially due to pharmaceutical innovation and manufacturers newly licensing drug-condition pairs. The right panel of Figure 3 displays the mean vintage by condition, independent of their off- or on-label status.

–Figure 3–

Table 1 presents descriptive statistics of measures of health care use and disability including average cost. Across the two-year survey period of MEPS, the average number of inpatient events for a medical condition was less frequent (0.0373) compared to emergency events (0.0809) and home health events (0.0963). Office-based visits were most frequent. Per condition, individuals received on average 2 prescription medicines. In 12.6% of observations, individuals reported bed days and in 5.6% of observations, they reported missed school days. Missed work days were reported in 10.3% of observations.

–Table 1–

Figure 4 shows the distribution of the number of conditions per individual of the persons included in MEPS. The mean number of individuals per condition reported is 2,240 (sd 5,526). 56% of individuals report two or more conditions. Besides, the majority of conditions are borne by individuals with at least two conditions. In 2015, 83% of medical conditions are borne by individuals with two or more conditions. This descriptive evidence lets us conclude that we are able to account for individual-level fixed effects in our empirical approach, because many individuals report multiple conditions.

–Figure 4–

To assess how off-label use evolves by the number of conditions, Figure 5 shows the mean deviation from the population average off-label use by the number of conditions reported. Positive deviations indicate that the individuals who use more drugs off-label than the average are largest in individuals with one or two conditions. When individuals report 3-9 conditions, deviations from the mean are negative. Deviations from the mean become variable in terms of direction when individuals report 10 or more conditions.

–Figure 5–

III. Empirical strategy

Our hypothesis about the the impact of off-label use of drugs on health care use and disability is based on theories of productivity and human capital formation (Böhm, Grossmann, and Strulik 2021; M. Grossman 2000; Strulik 2015). We account for endogenous growth of technology to motivate why newer vintage drugs are of higher quality compared to older vintage drugs and why the fraction of drugs used off-label is lower for newer drugs compared to older drugs (G. M. Grossman and Helpman 1991; Solow 1960). We provide a comprehensive structural framework in the Appendix. Figure 2 shows the structural relationships.

We assume that human capital formation is dependent on the use of medical technology. We account for the fact that individuals accumulate health deficits during their life course that reduces their human capital stocks. Use of effective technology reduces health deficit accumulation. When a technology is used according to label, it reduces health deficit accumulation, with higher quality treaments being more effective compared to treatments not designed to treating a condition. Thus, off-label use and the average quality of a drug are attributes of medical technology, and disability and utilization of health services (e.g. inpatient and emergency room visits) reflect the extent to which health deficits are accumulated.

–Figure 2–

We assume that a prescription drug d that has received regulatory approval for condition c has demonstrated effectiveness. This is the information included in the FDA's label or in the French regulatory data. In contrast, we assume that the drug has zero or negative effects in the set of conditions c^{-} for which drug d has not received approval. Not requesting marketing authorization does not mean that there is no scientific evidence that supports the effective use of a drug outside its designated label, that a regulator disapproves this use or, that the drug

use is ineffective. We assume that the evidence base is less solid leading to at least lower or zero effectiveness of off-label compared to on-label uses to treat the same condition.

Technologies that do not receive a label to treat a medical condition increase the health deficit accumulation when effectiveness is not demonstrated through regulatory approval. Subsequently, we should observe increases in health care use and disability compared to when using a drug on-label.

The theoretical relationships allow specifying two hypotheses regarding off- compared to on-label use and the approval status of a drug. The first is the hypothesis that off-label use has no effect or increases disability and health care use due to adverse drug events. If anything, as suggested by empirical evidence, off-label use is either not supported by solid evidence or this evidence was not screened as part of a regulatory approval process that typically is stricter and more comprehensive, making effectiveness of off-label drugs smaller and more uncertain (McKibbin 2020). The second hypothesis concerns the spectrum of quality of drugs used in the treatment of condition c that have proven their effectiveness. Here, we hypothesize that the average quality of newer drugs is larger within the same condition.

We estimate two-way fixed effect regressions at the individual by condition level to estimate the average treatment effect of off-label use, controlling for drug vintage (Angrist and Pischke 2008). That way, we account for any unobserved individual- and condition-level effects on our outcomes of interest. Corresponding with Figure 2, we specify the following general linear model.

$$y_{ic}^{k} = \beta_1 \cdot \text{off-label}\%_{ic} + \beta_2 \cdot vintage_{ic} + \alpha_i + \pi_c + \varepsilon_{ic} \qquad (3)$$

For each outcome y_{ic}^{k} for individual i and condition c that measures health care use and disability, we estimate separate models, where the superscript k indicates a dimension of health care use or disability.

Our primary variable of interest (off-label%$_{ic}$) is the fraction of prescriptions used by individual i for condition c that have not been approved to treat condition c as described in equation (1). The estimates of β_1 indicate the

average treatment effect of off-label use, i.e. the effect of the patient using a drug for a condition that *was not* approved to treat the condition, compared to using a drug for a condition that *was* approved to treat the condition. Given that there are potential errors in our measure of off-label%, any such biases are likely to bias estimates of β_1 towards zero.

We include $vintage_{ic}$ as a regressor in equation (3) to test the hypothesis that newer drugs are typically of higher average quality than older drugs and therefore reduce health care use and disability. The estimates of β_2 indicate the effect of using newer technology captured by its vintage in the production of outcome y_{ic}^{k}. Besides, mean drug vintage likely influences off-label use and our outcomes simultaneously; as noted above, the amount of promotional activity, which can only be for labeled indications, declines sharply after patent expiration, and newer on-label drugs are subject to stronger control by payors to avoid highly priced treatments.

To analyze the effect of off-label compared to on-label use, we need to consider variability in effectiveness of pharmaceutical treatment at individual and condition level. α_i reflect unobserved individual fixed effects of individual i that account for person specific disease severity related to individual-level attributes such age, gender, race, (epi-)genetics, metabolism and hormonal influences, factors related to socioeconomic status, insurance, location, income, and mean date of interview coverage. Tunçel (2020) suggests that patient and physician characteristics substantially contribute to treatment choice of off-label drugs. α_i may include behavioral factors that some individuals are more risk-averse towards using off-label drugs or that their prescribing doctor is recommending off-label drugs. However, the primary goal is to account for the fact that some drugs may work better in one person than another and to account for any unobserved factors at the individual level that may be correlated with off-label use.

π_c reflect unobserved medical condition fixed effects, which control for the fact that some medical conditions tend to be more disabling and require more

medical care than others. ε_{ic} is the error term. Regressions were weighted by the number of prescriptions individual i received for treatment of condition c. We clustered standard errors at the individual level.

To assess the robustness of our estimates of the effect using a prescription drug off-label (β_1), we re-estimate the model stated in eq. (1) to account for factors that may potentially bias our estimates. To assess biases that might arise because some individuals report only one condition, we exclude individuals reporting one condition only. To account for the role of weighting observations by the number of prescriptions used, we calculated unweighted estimates. To account for the fact that the stock of medical knowledge in 2015 was different from what it was in 1996, we include a time fixed effect δ_t for the timing of the treatment episode. Since we observe a given individual during a maximum of two years in MEPS, the individual fixed effects essentially control for this. Finally, we estimated model variants that excluded the mean vintage of the drugs used and in interaction with off-label%$_{ic}$.

IV. Main Results

A. Effect of off-label use and innovation on health care use and disability

We start by evaluating the relationship between off-label use of individual i for condition c and the number of inpatient visits of that same individual for that condition. Table 2 shows the estimates of a one standard deviation increase in the fraction of drugs used off-label (off-label%) use and vintage from the two-way fixed effects regressions by medical condition and individual. Columns (1)–(5) indicate that the estimates of $\hat{\beta}_1$ are positive: the larger the fraction of drugs consumed off-label by individual i for medical condition c, the higher the number of inpatient visits. Column (1) reports the baseline estimate, indicating that if off-label use increases by one standard deviation, the number of inpatient visits increases by on average 0.0029. Column (2) confirms that our results are identical for $\hat{\beta}_1$ when we exclude individuals who report only one condition.

Column (3) suggests that when we do not weight our regression model by the number or prescription drugs consumed the estimate of $\hat{\beta}_1$ increases to 0.0039. In column (4), we report results when accounting for the year of the MEPS panel. Our baseline estimates are confirmed.

Column (5) shows our estimates of a one standard-deviation increase of off-label% when we excluded the vintage variable. The estimate is 0.0003 standard deviations larger and suggests that vintage indirectly influences the number of prescription drugs consumed off-label. Our results are confirmed by column (6) that shows interaction effects between off-label use and vintage. When individual i primarily uses more than 50% of prescriptions were off-label, we find a significant negative effect of vintage.

Regarding the contribution of innovation that we approximate through a drug's approval year, based on our estimates of $\hat{\beta}_2$ from equation (3), in individual i for medical condition c, we find that the average use of later vintage drugs within the same condition significantly decreases the number of inpatient visits, with estimates of one standard deviation increase in vintage between -0.0019 and -0.0023 (Columns (1)-(4)).

–Table 2–

We find consistent positive effects to suggest that generally, off-label use increasesl disability and health care use across all specified measures. We find significant negative effects of mean drug vintage on the majority of outcomes studied. Figure 6 shows the range of estimates for all outcomes from the same set of regressions shown in Table 2. The points indicate the baseline specification of equation (3), the gray bars indicate the range of estimates, expressed as percentage of the mean value of each outcome.

Comparing effect estimates of one standard deviation increase in off-label use as percentage of the mean use (Figure 6 and Table 2) suggests that effects of off-label was attributed to about 30% of mean home health events. The lowest contribution was attributed to the mean emergency events of about 5%. Effects

are smaller for disability. A one standard deviation increase in off-label use was attributed to about 10% of the mean numbers of days spend in bed and 5% of any work days. We did not find significant effects for individuals reporting missed school days and estimates were attributed to 1% of the mean number of missed school days.

Comparing effect estimates of a one standard deviation increase in the vintage of drugs used relative to the mean use levels of the different types of health care suggests that one standard deviation increase in using newer compared to older drugs contributed most strongly to reducing the number of outpatient and emergency events (both by about 6.1%). We also find that using one standard deviation newer compared to older drugs significantly contributed to a reduction in the probability of having to spend any days in bed (5.25%).

–Figure 6–

C. Cost of off-label use

To assess the cost of off-label compared, we perform counterfactual evaluations to consider reductions in medical expenditure and work loss cost if off-label use was fully substituted by on-label use of prescription drugs. To arrive at the aggregated value of medical expenditure and work loss cost, we assumed that 44% of the off-label use in 2015 is replaced by on-label options. We use estimates of the models that include person-level and condition fixed effects as reported in column (1) in Table 2 (and equation (3)). Column (1) in Table 3 shows effect estimates of a one unit increase in off-label use (i.e. changing from 0% to 100% to off-label use); Column (2) indicates the estimates as ratio to 2015 sample means. Column (3) presents the estimated monetary cost of off-label use.

Considering that all of the 44% of drugs that were administered off-label would be replaced by on-label uses 2015, medical expenditures would

effectively be reduced by on average $515 per individual. These costs represent 12.9% of the total cost that we could attribute to health care use and disability.

–Table 3–

These additional expenditures may be justified if the prescription cost of off-label uses are lower than their on-label counterparts in the same condition. Related evidence suggests potential welfare losses when banning or reducing use of off-label drugs (Bradford, Turner, and Williams 2018; Tunçel 2020). We calculated the average cost difference between on- and off-label treatments in addition to the cost attributed to medical expenditure and work loss cost that we identified in Table 3. Performing two-way fixed-effects regressions (equation (3)) for the mean number of prescriptions in 2015, we find that prescription cost of off-label uses are significantly higher by on average $680, or 22%. When people are using off-label drugs, their total expenditures are higher compared to on-label drug uses because of higher prices, higher quantities or both: the total cost attributed to off-label use amounts to $1,195 in 2015, on average.

B. Market size and innovation as causes of off-label use

To provide evidence about how off-label use differs by market size, we calculate disease prevalence based on the number of individuals reporting condition c across the study period and define five equally distributed groups. We categorize the average vintage of drugs used by individual i in condition c in five groups and estimate cross sectional regressions:

$$\text{off-label}\%_{ic} = \delta_1 \cdot market\ size_c + \delta_2 \cdot vintage_{ic} + \varepsilon_{ic} \qquad (4)$$

δ_1 captures the association of market size and off-label use whereas δ_2 captures the association of vintage and off-label use. We find that larger markets receive more innovation targeted to treat a condition such that off-label use is lower the higher the disease prevalence. Column (1) of Table 4 shows the marginal effect estimates by five categories of market size, suggesting a strong negative correlation. The association of conditions with the smallest market size

(estimate: 0.93, 1–106 MEPS sample individuals between 1996 and 2015) and the fraction of drugs used off-labels was about 2.14 times higher compared to conditions with the largest number of individuals (marginal effect –49.78 percentage points in prescription drug markets with 4,614–7,383 individuals).

Columns (2) and (3) of Table 4 suggest that drugs of later vintages were associated with a lower fraction of drugs used off-label compared to drugs of older vintages. Column (2) reports the marginal effect by one additional approval year which means that 10 years of innovation would reduce the average fraction of drugs used off-label by 6.9 percentage points. Estimates by groups of vintages as reported in column (3) suggests that drugs approved between 1896 and 1965 that were still in use during 1996-2015 had the highest fraction of off-label use. In contrast to our study which compares the average vintage of a drug, Eguale et al. (2016) found that newer drugs were more likely to be used off-label. However, Eguale et al. analyzed the age of the drug after marketing approval; they found that off-label use is highest at early stages of marketing across the life cycle.

When we include market size and vintage simultaneously, estimates do not change substantially (Column (4)). Column (5) reports estimates including interaction effects between market size and vintage. Innovation reduces off-label use more strongly in large markets, although at small extent (marginal effect -0.0057 in markets with the highest prevalence compared to a marginal effect of -0.0055 in markets with the lowest prevalence).

–Table 4–

V. Interpreting the Results

Our estimate of the cost impact of off-label use ($515) is much higher than previous estimates (Eguale), but we consider a broad set of health care use and disability, not merely the cost of ADEs. A similar counterfactual calculation

based on the results provided Eguale et al. (2016) suggests that the mean cost per patient in Québec, Canada, are $2-$4 when considering ADE only.

A concern to validly identify the effect of off-label use on health care use and disability ($\hat{\beta}_1$) is that off-label use of drugs may not be a random choice, especially when there are no other treatment options available or when scientific evidence suggests a benefit of off-label use. Similar to Eguale et al (2016) and other studies, we implicitly assume that off-label use is exogenous, or random, conditional on the covariates. Our estimates could be biased if this assumption were violated, e.g. if off-label use tends to be greater for people with higher disease severity. Off-label therapies can be a choice of last resort when other treatment options have been exploited (US Food and Drug Administration 2019). However, we are not aware of any evidence that supports the hypothesis that more severely ill patients are more likely to use drugs off-label. Even if off-label use is correlated with severity, including vintage in the model may control, at least to some extent, for the severity of the condition, because some payors or state policies may engage in step therapy (Tharp and Rothblatt 2022). Step therapy involves that patients first try less expensive options before 'stepping up' to drugs that cost more.

Off-label use is also prominent when there is no approved treatment available, as is the case in 71 out of the 247 conditions in 1996. It would likely not be feasible to randomize off-label use for any individual in a controlled study as based on the Declaration of Helsinki as there must be scientific and compelling arguments to test an effective on-label use against a potentially less effective off-label use of a drug.[14] In addition, there are considerations such as dosing and mode of administration when using a drug off-label that equally need scientific investigation.[15]

[14] Use of Placebo as based on https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

[15] One option would be a physician education program or randomizing physicians in using electronic health records to evaluate programs that can distinguish off- and on-label use that we are not aware of.

In the two-way fixed effects regressions, capturing person-level fixed-effects should account for the fact that some individuals are more severely ill than others given the individual's specific biologic and socio-economic profile (Mauvais-Jarvis et al. 2021). A study of off-label use in intensive care units, which typically treat the most severely ill patients, suggests that off-label use is frequently due to extrapolation of on-label uses from care settings outside the designated label, for example insulin treatments in medical compared to surgical intensive care (Lat et al. 2011). Such uses are unrelated to severity but owed to the setting in which a drug is provided supporting the argument that off-label use in the more severely ill is systematically related to that condition.

To substantiate the interpretation of the effect of off-label use, we explore the role of disease severity by number of conditions at individual level. If unobserved condition and individual level heterogeneity captures disease severity well, the number of conditions an individual suffers from may become irrelevant in the use of off-label drugs. We investigate the relationship between the number of conditions reported and any observable and unobservable individual variation as follows, considering the average fraction of off-label% used by individual i as outcome:

$$\text{off-label}\%_i = \alpha_1 \cdot N\_cond_i + \alpha_2 \cdot \hat{\pi}_i + \alpha_3 \cdot \boldsymbol{X}_i + \alpha_4 \cdot \hat{e}_i + \alpha_5 \cdot \overline{vintage_i} + \varepsilon_i, \qquad (5)$$

Estimates of α_1 indicate how much of the across-individual variation in off-label use can be explained by the number of conditions (N_cond_i). $\hat{\pi}_i$ is the aggregate disease severity by the individual using the fixed effects from equation (3). We use the condition-level fixed effects estimates by outcome k to calculate a severity measure summing up the estimated condition-specific fixed effects by individual i as $\hat{\pi}_i = \sum_{c=1}^{C} \hat{\pi}_c$ for each condition the individual reports. $\hat{e}_i = \sum_{c=1}^{C} \hat{e}_c$ specifies the sum of residuals of individual i across all conditions. That measures captures remaining unobserved individual variation. $\boldsymbol{X}_i$ is a vector of observable characteristics including age, sex, whether the individual is black compared to white, and years of education. $\overline{vintage_i}$ is the

average vintage of prescription drugs used by individual i across all prescription drugs and conditions. Robust standard errors are reported to account for heteroscedasticity.

Columns (1)–(6) of Table 5 report results of estimates of α_1 and off-label% of person i in terms of mean deviations which suggest that the number of conditions a person reports is a predictor of off-label use (estimate: 0.1067, $p<0.0001$, Column (1)). When we control for additional observable factors, the effect of the number of conditions becomes very small and close to zero (estimate: 0.02 to –0.007, $p<0.0001$). Column (6) shows that especially for individuals that report 3 to 9 conditions, off-label use is small but significantly smaller (estimate: -0.0096) compared to reporting one or two conditions, and is not significant when individuals report ten or more conditions. We conclude that the two-way fixed effect approach largely captures disease severity by individual and condition and that the number of conditions does not strongly bias our estimates of off-label use.

–Table 5–

In contrast to previous studies that investigated off-label use, we use patient-reported data of conditions, which may be less reliable and individuals may not correctly assign medications to conditions. In Eguale et al's (2016) study, medical conditions were reported by physicians in electronic health records. In the MEPS, conditions are reported by patients, not providers.[16] Individuals may not correctly assign the medical treatments they receive to a condition. Systematic biases in medical condition reporting by individuals were controlled for by including individual fixed effects. Random misreporting of medical conditions is likely to introduce measurement error into our measure of off-label use, and therefore make our estimates of the effect of off-label use conservative. Comparing our unadjusted estimate of 44% off-label use in 2015, we find that this value is higher compared to disease unspecific estimates of Bradford et al.

[16] Also, unlike Eguale et al. (2016), we are unable to distinguish between off-label use with and without strong scientific evidence. They estimated that 81% of off-label prescriptions lacked strong scientific evidence.

(2018), or disease-specific estimates by Tunçel (2020) and Molitor (2012), but using prescriber data. For our data, Figure 5 suggests that off-label use is higher compared to average in individuals reporting one or two conditions, and lower when reporting 3-9 conditions. The MEPS data do not suggest a clear pattern that individuals with a high number of conditions systematically assign their medication wrongly as off-label.

VI. Population Heterogeneity

We derived separate estimates by selected observable characteristics to study heterogeneity of the cost of off-label use. Unadjusted results suggest that off-label use is lower in adults compared to children and adolescents up to age 18 (50.55% vs. 42.87%), and about equal in females compared to males (43.92% vs. 43.40%), and in blacks compared to whites (42.79% vs. 43.39%). We estimated two-way fixed effects regressions (equation (3)) by these groups to arrive at group-specific counterfactual analyses of medical expenditure and work loss cost in 2015.

Panels (A)–(C) of Figure 7 report cost of off-label use by population groups. Aggregated medical expenditure and work-loss cost attributed to off-label use is smaller in children and adolescents, $288 compared to $587 adults (Panel (A)). These cost account for a larger proportion of the average expenditure of children (17% in children and adolescents compared to 10% in adults). Panel A highlights differences in cost composition across age groups. In children and adolescents, for example, a higher proportion of expenditures related to inpatient events are used. Panel (B) of Figure 7 suggests that there are no substantial differences between blacks and whites ($518 vs. $512). Of note is that our results are for individuals using prescription drugs which may disregard the possibility that blacks have worse access to receiving a prescription drug overall. Panel C shows cost differences attributed to off-label use between females and males suggesting that the absolute value of medical expenditures

and work loss cost attributed to off-label use is higher for women compared to men ($580 vs. $444) with simlar relative shares similar (11.8% vs. 10.9%).

VII. Conclusion

In this study, we empirically investigated the relationship between two attributes of prescription drugs used to treat a medical condition—whether the drug was approved to treat the condition, and when the drug was approved. Using unique data that include combinations of prescription drug labels and indicated conditions for over 200,000 individuals in the US non-institutionalized population, we find that off-label use generally increases medical expenditures and work loss cost. Our empirical estimates of two-way fixed effects regressions and counterfactual analyses suggest that using only on-label drugs compared to the average 2015 off-label drug use would lead to savings of $515 per individual in terms of health care use and work-loss cost. We demonstrate that off-label use is inversely related to both market size and drug vintage. Pharmaceutical innovation may therefore have both direct and indirect (via the propensity to use drugs off-label) effects on health care use and disability. These estimates of off-label use are robust across a number of specifications. We further deal with an important endogeneity issue by demonstrating that individuals with higher disease severity and multiple conditions do not experience higher off-label use once we account for unobserved individual effects.

The estimates of medical expenditure and work loss cost of off-label use per capita can be used as value to biomedical innovation to demonstrate the use of a drug through regulatory approval. This value is substantial considering that it contributes to on average 12.7% of medical expenditure and work loss cost. While the identified monetary value provides a reasonable approximation, it is possible that this value varies by conditions or patient group and is potentially higher, especially as the immediate prescription drug expenses of off-label drugs higher by $680, or 22% on average. It may be lower when scientific

support of benefits of off-label uses exist which often are not present. The heterogeneity analysis by broad population groups suggest that this value is also higher in children and adolescents and in women. Evidence on the welfare effects of substituting off-label drugs suggests that off-label use could improve welfare due to lower prices of off-label treatments, but has not considered within-individual and -condition differences in expenditures and the quantity of prescriptions used between on- and off-label uses. For depression treatment, Tunçel (2020) demonstrates that banning off-label use may increase prices as market size of a drug is reduced if off-label uses are limited. Bradford et al. (2018) show that insurance expansion increases use of less costly off-label drugs. If there is scientific support for off-label drugs, insurance expansion would increase welfare as uninsured patients would not purchase inferior off-label options. Our results suggest the opposite, as we find that off-label use substantially increases medical expenditure and work loss cost.

Primarily, our study speaks to policy debates regarding the efficiency of U.S. health care provision, particularly those concerning the wrong combinations of inputs or the wrong technology (Glied and Sacarny 2018). We also provide implications about the benefits of regulating innovations through drug approvals and market size of innovations to society (Acemoglu and Linn 2004; Blume-Kohout and Sood 2013). One promising avenue for drug development to expand on-label use of drugs is repositioning of existing drugs to new disease indications (Parisi et al. 2020). This mode of drug development could be more efficient and speed up development compared to the conventional search for molecules. A critical question is how new uses of drugs undergo scientific investigation that consume resources to perform clinical trials and scientific investigation. As we do not distinguish between scientifically proven compared to unproven off-label uses and we acknowledge that some off-label uses may be beneficial to patients, our estimates highlight the total substantial cost burden of off-label use of prescription drugs.

VIII. References

Acemoglu, Daron, and Joshua Linn. 2004. “Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry.” *The Quarterly Journal of Economics* 119 (3): 1049–90.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics*. *Mostly Harmless Econometrics*. Princeton University Press. https://www.degruyter.com/document/doi/10.1515/9781400829828/html.

Awad, Abdelmoneim, and Olivia Hanna. 2019. “Potentially Inappropriate Medication Use among Geriatric Patients in Primary Care Setting: A Cross-Sectional Study Using the Beers, STOPP, FORTA and MAI Criteria.” Edited by Liset Van Dijk. *PLOS ONE* 14 (6): e0218174. https://doi.org/10.1371/journal.pone.0218174.

Beers, Mark H., Joseph G. Ouslander, Irving Rollingher, David B. Reuben, Jacqueline Brooks, and John C. Beck. 1991. “Explicit Criteria for Determining Inappropriate Medication Use in Nursing Home Residents.” *Archives of Internal Medicine* 151 (9): 1825–32.

Blume-Kohout, Margaret E., and Neeraj Sood. 2013. “Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development.” *Journal of Public Economics* 97 (January): 327–36. https://doi.org/10.1016/j.jpubeco.2012.10.003.

Bodenreider, Olivier, and Laritza M. Rodriguez. 2014. “Analyzing U.S. Prescription Lists with RxNorm and the ATC/DDD Index.” *AMIA Annual Symposium Proceedings* 2014 (November): 297–306.

Böhm, Sebastian, Volker Grossmann, and Holger Strulik. 2021. “R&D-Driven Medical Progress, Health Care Costs, and the Future of Human Longevity.” *The Journal of the Economics of Ageing* 18 (February): 100286. https://doi.org/10.1016/j.jeoa.2020.100286.

Bradford, W. David, John L. Turner, and Jonathan W. Williams. 2018. “Off-Label Use Of Pharmaceuticals: A Detection Controlled Estimation Approach.” *The Journal of Industrial Economics* 66 (4): 866–903. https://doi.org/10.1111/joie.12189.

Bryan, Kevin A., and Heidi L. Williams. 2021. “Innovation: Market Failures and Public Policies.” Working Paper 29173. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w29173.

Budish, Eric, Benjamin N. Roin, and Heidi Williams. 2015. “Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials.” *American Economic Review* 105 (7): 2044–85. https://doi.org/10.1257/aer.20131176.

Buxbaum, Jason D., Michael E. Chernew, A. Mark Fendrick, and David M. Cutler. 2020. “Contributions Of Public Health, Pharmaceuticals, And

Other Medical Care To US Life Expectancy Changes, 1990-2015." *Health Affairs* 39 (9): 1546–56. https://doi.org/10.1377/hlthaff.2020.00284.

Chandra, Amitabh, and Jonathan Skinner. 2012. "Technology Growth and Expenditure Growth in Health Care." *Journal of Economic Literature* 50 (3): 645–80.

Cohen, Joel W., Steven B. Cohen, and Jessica S. Banthin. 2009. "The Medical Expenditure Panel Survey: A National Information Resource to Support Healthcare Cost Research and Inform Policy and Practice." *Medical Care* 47 (7_Supplement_1): S44–50.

Cohen, Steven B., Trena Ezzati-Rice, and William Yu. 2006. "The Utility of Extended Longitudinal Profiles in Predicting Future Health Care Expenditures." *Medical Care* 44 (5): I–45.

Dalgaard, Carl-Johan, and Holger Strulik. 2014. "OPTIMAL AGING AND DEATH: UNDERSTANDING THE PRESTON CURVE: Optimal Aging and Death." *Journal of the European Economic Association* 12 (3): 672–701. https://doi.org/10.1111/jeea.12071.

Dresser, Rebecca, and Joel Frader. 2009. "Off-Label Prescribing: A Call for Heightened Professional and Government Oversight." SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=1763292.

Duflos, Gautier, and Frank R. Lichtenberg. 2012. "Does Competition Stimulate Drug Utilization? The Impact of Changes in Market Structure on US Drug Prices, Marketing and Utilization." *International Review of Law and Economics* 32 (1): 95–109. https://doi.org/10.1016/j.irle.2011.08.003.

Eguale, Tewodros, David L. Buckeridge, Aman Verma, Nancy E. Winslade, Andrea Benedetti, James A. Hanley, and Robyn Tamblyn. 2016. "Association of Off-Label Drug Use and Adverse Drug Events in an Adult Population." *JAMA Internal Medicine* 176 (1): 55. https://doi.org/10.1001/jamainternmed.2015.6058.

Fleming-Dutra, Katherine E., Adam L. Hersh, Daniel J. Shapiro, Monina Bartoces, Eva A. Enns, Thomas M. File, Jonathan A. Finkelstein, et al. 2016. "Prevalence of Inappropriate Antibiotic Prescriptions Among US Ambulatory Care Visits, 2010-2011." *JAMA* 315 (17): 1864–73. https://doi.org/10.1001/jama.2016.4151.

Frank, Richard G., and Richard J. Zeckhauser. 2007. "Custom-Made versus Ready-to-Wear Treatments: Behavioral Propensities in Physicians' Choices." *Journal of Health Economics* 26 (6): 1101–27. https://doi.org/10.1016/j.jhealeco.2007.08.002.

Glied, Sherry, and Adam Sacarny. 2018. "Is the US Health Care System Wasteful and Inefficient? A Review of the Evidence." *Journal of Health Politics, Policy and Law* 43 (5): 739–65. https://doi.org/10.1215/03616878-6951103.

Grissinger, Matthew. 2010. "The Five Rights: A Destination Without a Map." *Journal for Formulary Management* 35 (10): 542–542.

Grossman, Gene M., and Elhanan Helpman. 1991. “Quality Ladders in the Theory of Growth.” *The Review of Economic Studies* 58 (1): 43–61. https://doi.org/10.2307/2298044.

Grossman, Michael. 2000. “The Human Capital Model.” *Handbook of Health Economics* 1: 347–408.

Hanlon, Joseph T., and Kenneth E. Schmader. 2013. “The Medication Appropriateness Index at 20: Where It Started, Where It Has Been and Where It May Be Going.” *Drugs & Aging* 30 (11). https://doi.org/10.1007/s40266-013-0118-4.

Husson, M. -C. 2008. “Thériaque® : base de données indépendante sur le médicament, outil de bon usage pour les professionnels de santé.” *Annales Pharmaceutiques Françaises* 66 (5): 268–77. https://doi.org/10.1016/j.pharma.2008.07.009.

“Indication - Medical Definition and More from Merriam-Webster.” 2011. July 14, 2011. https://web.archive.org/web/20110714012214/http:/mw4.m-w.com/medical/indication.

Janakiraman, Ramkumar, Shantanu Dutta, Catarina Sismeiro, and Philip Stern. 2008. “Physicians’ Persistence and Its Implications for Their Response to Promotion of Prescription Drugs.” *Management Science* 54 (6): 1080–93. https://doi.org/10.1287/mnsc.1070.0799.

Jovanovic, Boyan, and Yuri Yatsenko. 2012. “Investment in Vintage Capital.” *Journal of Economic Theory*, Issue in honor of David Cass, 147 (2): 551–69. https://doi.org/10.1016/j.jet.2010.10.017.

Kashoki, Mwango, Zahra Hanaizi, Stella Yordanova, Richard Veselý, Christelle Bouygues, Jordi Llinares, and Sandra L. Kweder. 2020. “A Comparison of EMA and FDA Decisions for New Drug Marketing Applications 2014–2016: Concordance, Discordance, and Why.” *Clinical Pharmacology & Therapeutics* 107 (1): 195–202. https://doi.org/10.1002/cpt.1565.

Kremer, Mirko, Brent Moritz, and Enno Siemsen. 2011. “Demand Forecasting Behavior: System Neglect and Change Detection.” *Management Science* 57 (10): 1827–43. https://doi.org/10.1287/mnsc.1110.1382.

Lat, Ishaq, Scott Micek, Jeffrey Janzen, Henry Cohen, Keith Olsen, and Curtis Haas. 2011. “Off-Label Medication Use in Adult Critical Care Patients.” *Journal of Critical Care* 26 (1): 89–94. https://doi.org/10.1016/j.jcrc.2010.06.012.

Lichtenberg, Frank R. 2014. “The Impact of Pharmaceutical Innovation on Disability Days and the Use of Medical Services in the United States, 1997–2010.” *Journal of Human Capital* 8 (4): 432–80.

———. 2020. “How Cost-Effective Are New Cancer Drugs in the U.S.?” *Expert Review of Pharmacoeconomics & Outcomes Research* 0 (0): 1–17. https://doi.org/10.1080/14737167.2020.1709965.

———. 2022. “The Effect of Pharmaceutical Innovation on Longevity: Evidence from the U.S. and 26 High-Income Countries.” *Economics &*

Human Biology 46 (August): 101124. https://doi.org/10.1016/j.ehb.2022.101124.

Mauvais-Jarvis, Franck, Heiner K. Berthold, Ilaria Campesi, Juan-Jesus Carrero, Santosh Dhakal, Flavia Franconi, Ioanna Gouni-Berthold, et al. 2021. "Sex- and Gender-Based Pharmacological Response to Drugs." Edited by Eric Barker. *Pharmacological Reviews* 73 (2): 730–62. https://doi.org/10.1124/pharmrev.120.000206.

McKibbin, Rebecca Jay. 2020. "The Effect of RCTs on Demand for Off-Label Cancer Drugs." SSRN Scholarly Paper ID 3574623. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3574623.

Molitor, David Paul. 2012. "Physician Behavior and Technology Diffusion in Health Care." Thesis, Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/77795.

Parisi, Daniele, Melissa F. Adasme, Anastasia Sveshnikova, Sarah Naomi Bolz, Yves Moreau, and Michael Schroeder. 2020. "Drug Repositioning or Target Repositioning: A Structural Perspective of Drug-Target-Indication Relationship for Available Repurposed Drugs." *Computational and Structural Biotechnology Journal* 18: 1043. https://doi.org/10.1016/j.csbj.2020.04.004.

Radley, David C., Stan N. Finkelstein, and Randall S. Stafford. 2006. "Off-Label Prescribing Among Office-Based Physicians." *Archives of Internal Medicine* 166 (9): 1021–26. https://doi.org/10.1001/archinte.166.9.1021.

Richardson, Elizabeth. 2016. "Off-Label Drug Promotion," Health Affairs Brief, , June. https://doi.org/10.1377/hpb20160630.920075.

Solow, R. 1960. "Investment and Technological Progress." In *Mathematical Methods in the Social Sciences, 1959*, edited by Kenneth Arrow, S Karlin, and P. Suppes. Stanford, Calif: Stanford University Press.

Strulik, Holger. 2015. "Frailty, Mortality, and the Demand for Medical Care." *The Journal of the Economics of Ageing* 6 (December): 5–12. https://doi.org/10.1016/j.jeoa.2015.09.001.

Tafuri, G. 2013. "Exploring the Regulatory Decision-Making Process for Medicines = Het Besluitvormingsproces Rondom de Toelating van Geneesmiddelen = Il Processo Decisionale per i Farmaci a Livello Regolatorio." S.l.: s.n.].

Tharp, Louis, and Zoe Rothblatt. 2022. "Do Patients Benefit from Legislation Regulating Step Therapy?" *Health Economics, Policy and Law* 17 (3): 282–97. https://doi.org/10.1017/S1744133121000153.

Tunçel, Tuba. 2020. "Should We Prevent Off-Label Drug Prescriptions? Empirical Evidence from France." SSRN Scholarly Paper ID 3694632. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3694632.

US Food and Drug Administration. 2019. "Understanding Unapproved Use of Approved Drugs 'Off Label.'" FDA. FDA. April 18, 2019.

https://www.fda.gov/patients/learn-about-expanded-access-and-other-treatment-options/understanding-unapproved-use-approved-drugs-label.

Weda, Marjolein, Joëlle Hoebert, Marcia Vervloet, Carolina Moltó Puigmarti, Nikky Damen, Sascha Marchange, Joris Langedijk, John Lisman, and Liset van Dijk. 2017. “Study on Off-Label Use of Medicinal Products in the European Union.” European Union.

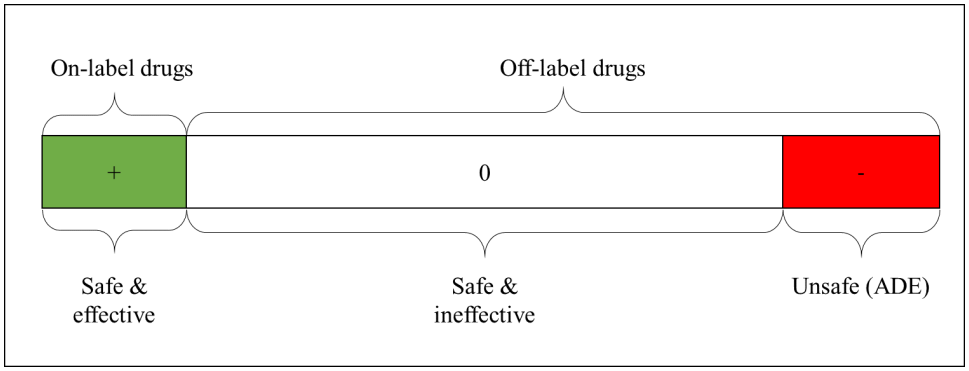
Wishart, David S., Craig Knox, An Chi Guo, Savita Shrivastava, Murtaza Hassanali, Paul Stothard, Zhan Chang, and Jennifer Woolsey. 2006. “DrugBank: A Comprehensive Resource for in Silico Drug Discovery and Exploration.” *Nucleic Acids Research* 34 (Database issue): D668-672. https://doi.org/10.1093/nar/gkj067.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT press. http://books.google.de/books?hl=de&lr=&id=cdBPOJUP4VsC&oi=fnd&pg=PR17&ots=jbccISh8xi&sig=ZOq_tKIBdO3doRg3PWzFp3UA79I.

Yackey, Katelyn, Kristin Stukus, Daniel Cohen, David Kline, Sonia Zhao, and Rachel Stanley. 2019. “Off-Label Medication Prescribing Patterns in Pediatrics: An Update.” *Hospital Pediatrics* 9 (3): 186–93. https://doi.org/10.1542/hpeds.2018-0168.

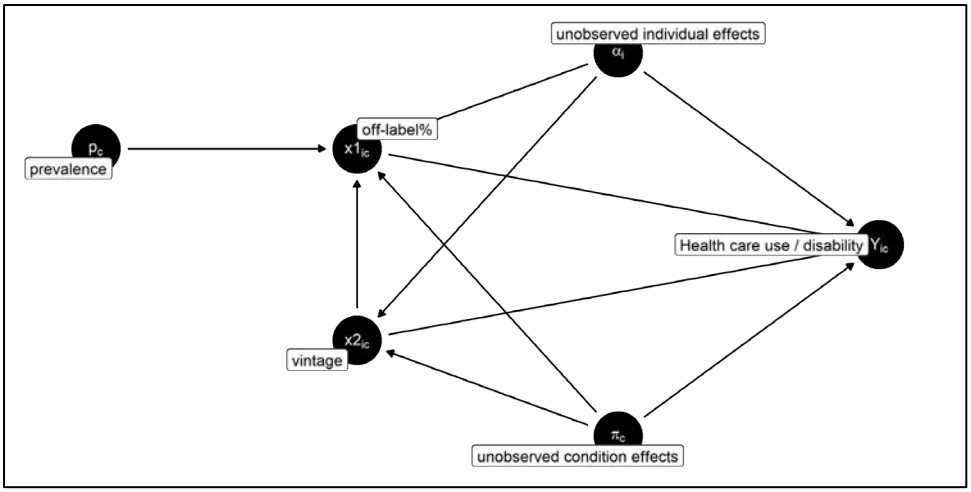
Tables and Figures

FIGURE 1: ON AND OFF-LABEL USE OF DRUGS AND THEIR EFFECTS ON HEALTH



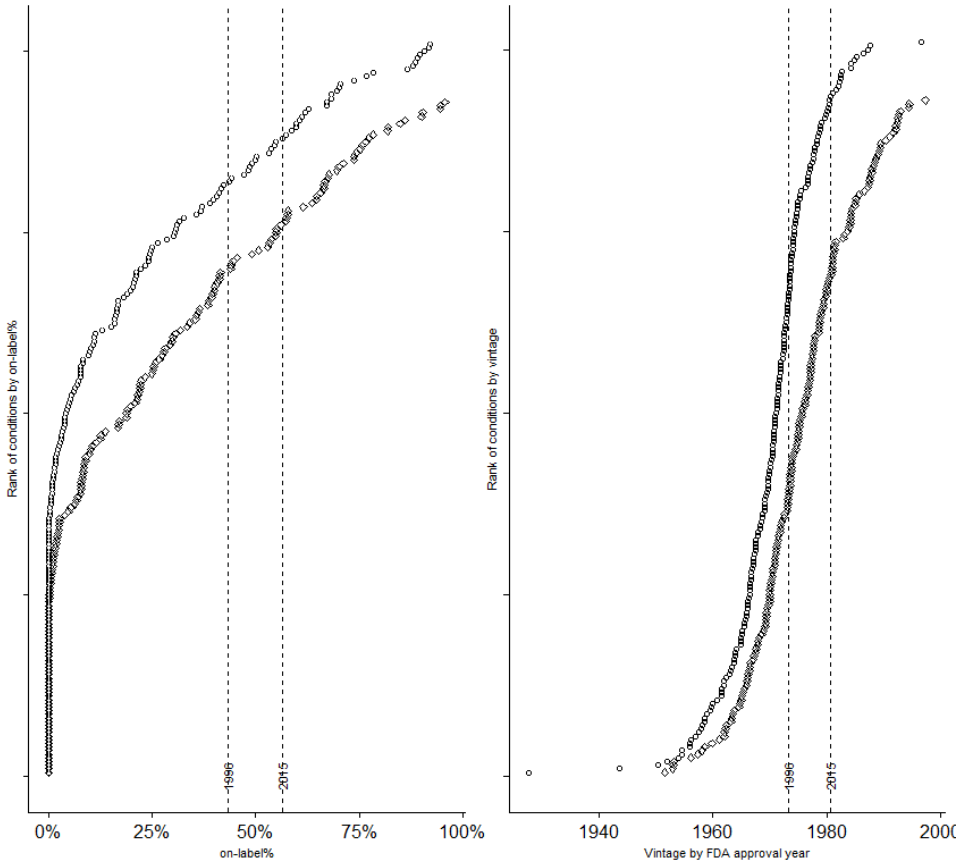
Note: This figure shows the stylized effectiveness of on-label and off-label drugs. Given the data shown and evaluated in regulatory approval, on-label drugs that have received approval are considered safe and effective with some positive patient benefit. Off-label drugs may be safte to use, but ineffective or harm a person causing adverse drug events (ADE). Off-label drugs have not been evaluated through regulatory approval to demonstrate a positive effect.

FIGURE 2: STRUCTURAL RELATIONSHIPS BETWEEN MARKET SIZE (DISEASE PREVALENCE) AND INNOVATION (DRUG VINTAGE) AS CAUSE OF OFF-LABEL USE, AND HEALTH CARE USE / DISABILITY AS CONSEQUENCES OF OFF-LABEL USE



Note: This figure shows the hypothesized relationships between market size (captured by disease prevalence of prescription drug users), the fraction of prescription drugs used off-label, the vintage and approval status of the prescription drug used captured by the FDA approval year of a drug, unobserved individual and condition effects and health care use and disability. Our primary variable of interest is off-label%.

FIGURE 3: FRACTION OF PRESCRIPTION DRUGS USED ON-LABEL AND VINTAGE BY CONDITION, 1996 AND 2015



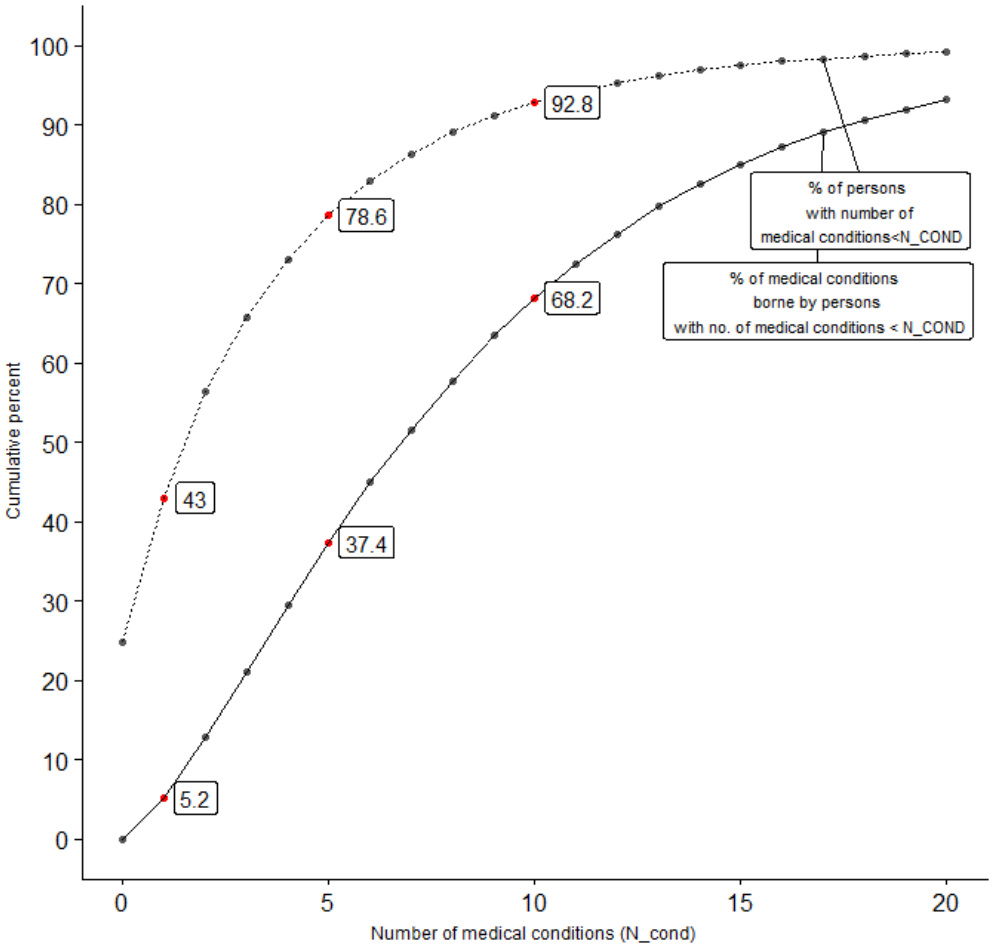
Note: This figure summarizes the variation across individual conditions in the use of drugs on-label and the mean approval year (vintage) used. Off-label corresponds with 100%-on-label%. Conditions are ranked by their average fraction of on-label use in 1996 and 2015. The value of 0% indicates that all drugs used in the condition were used off-label. The dashed lines represent the mean values of the fraction of prescriptions used on-label and vintage in 1996 and 2015. Each circle (square) represents the mean fraction of prescriptions used on-label and vintage by condition in 1996 (2015) according to 247 conditions based on CCS codes from the Clinical Classification Software provided by AHRQ. Data on labelled indications were obtained from Thériaque. Data on drug use were obtained from the Medical Expenditure Panel Survey.

TABLE 1: REGRESSION SUMMARY STATICSTICS OF OUTCOMES BY TYPES OF HEALTH CARE USE AND DISABILITY

Outcome	N	Mean	Std. Dev	Min	Max
Health care use					
# Inpatient events	553,302	0.0373	0.2088	0.0000	13.0000
# Emergency events	553,302	0.0812	0.3179	0.0000	16.0000
# Home health events	553,302	0.0965	0.8292	0.0000	67.0000
# Office-based events	553,302	1.6764	4.1735	0.0000	209.0000
# Out-patient events	553,302	0.1523	1.5191	0.0000	161.5000
# Prescribed medicines	553,302	2.0128	1.7148	0.0000	45.0000
Disability					
Individual had bed days: yes / no	457,898	0.1265	0.3192	0.0000	1.0000
Missed school days: yes / no	457,916	0.0563	0.2241	0.0000	1.0000
Missed work days: yes / no	457,905	0.1034	0.2964	0.0000	1.0000
General					
Number of individuals	201,489				
Number of conditions	247				
Number of conditions, 100% off-label use in 1996	71				
Number of conditions, 100% off-label use in 2015	46				
Number of observations (individual x condition)	553,302				

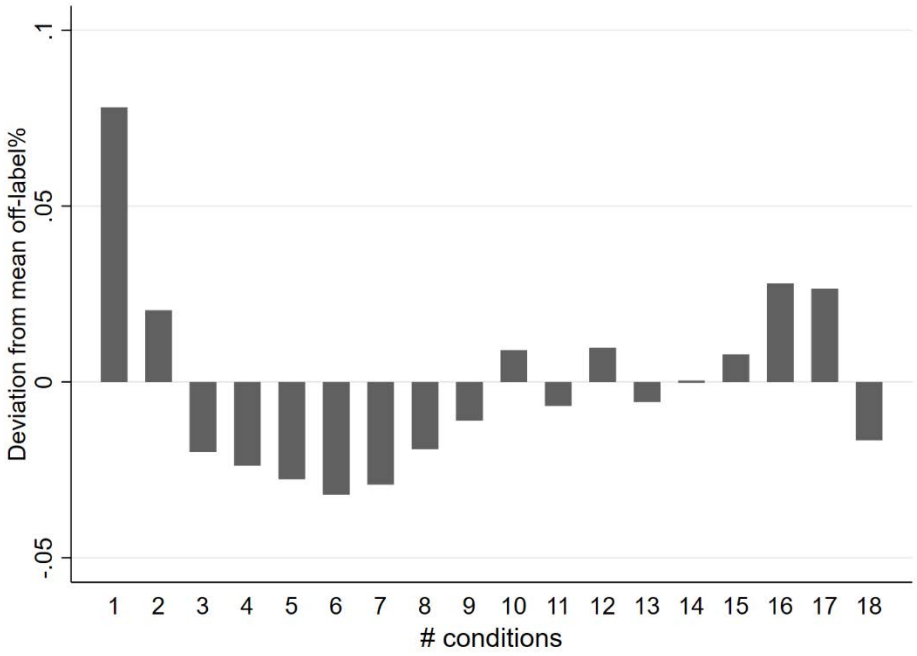
Note: This data summarizes the data on individuals by condition. Statistics correspond to individuals that received a prescription drug between 1996 and 2015 that were linked to the reported outcomes related to health care use and disability for the same condition. Data were obtained from the Medical Expenditure Panel Survey, 1996-2015, N= 553,302 observations individual by condition observations and n=201,489 individuals.

FIGURE 4: DISTRIBUTION OF PERSONS AND MEDICAL CONDITIONS BY NUMBER OF MEDICAL CONDITIONS BORNE BY PERSON, 2015



Note: The figure plots the fraction of persons the number of medical conditions n<N_COND (dotted line) and the fraction of medical conditions borne by persons with the number of medical conditions n<N_COND in the Medical Expenditure Panel Survey (solid line), 2015. Data were obtained from the Medical Expenditure Panel survey.

FIGURE 5: MEAN DEVIATION IN FRACTION OF PRESCRIPTION DRUGS USED OFF-LABEL BY NUMBER OF CONDITIONS PER INDIVIDUAL



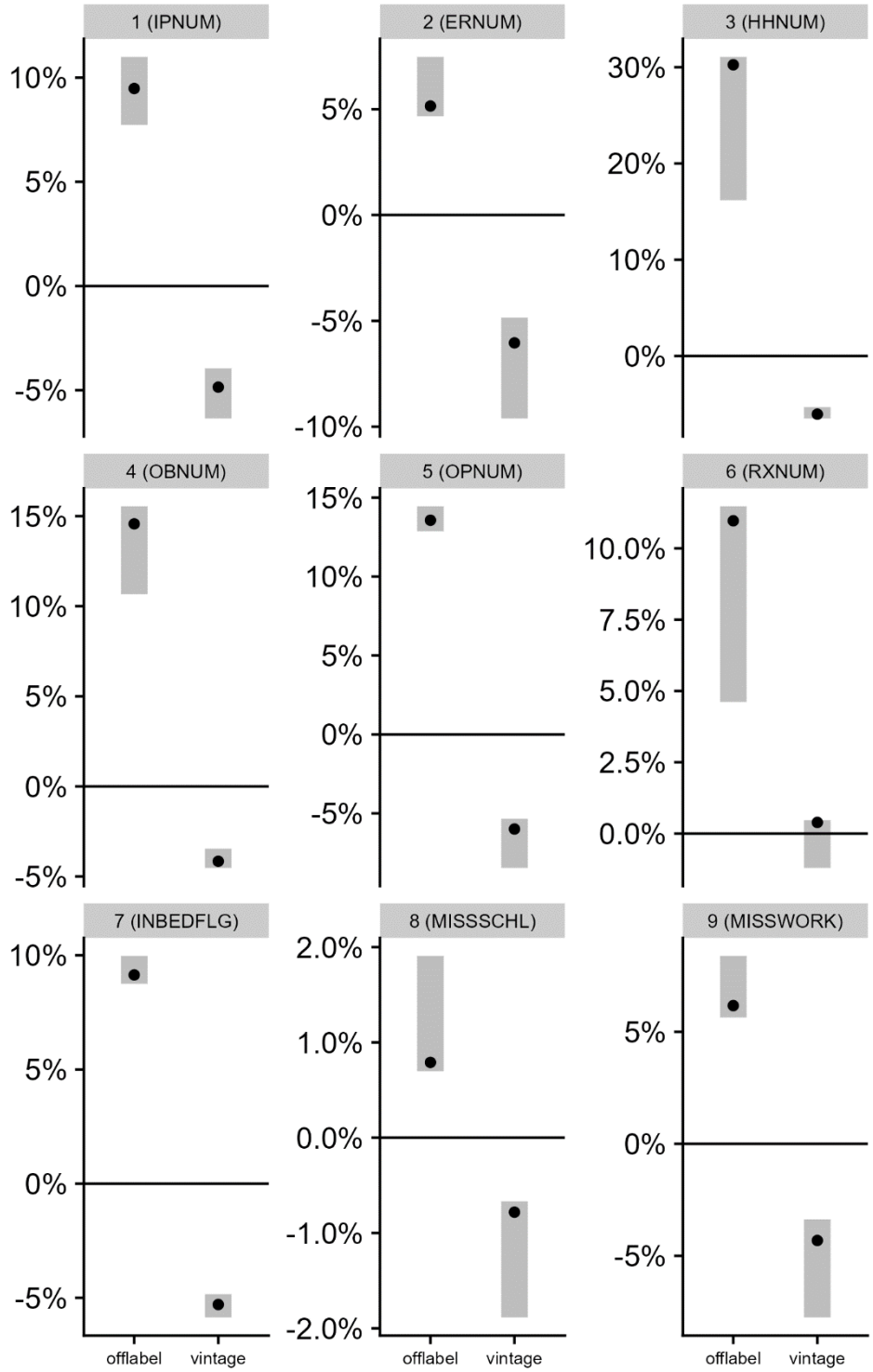
Note: The figure shows the mean deviation of off-label use of prescription drugs by number of conditions reported by the same individual. Positive values indicate that the fraction of drugs used off-label is higher compared to the sample mean. Negative values indicate that the fraction of drugs used off-label is lower than the sample mean. Results of individuals with more than 18 conditions are excluded from the figure. Data were obtained from the Medical Expenditure Panel Survey, 1996-2015, N=553,302 individual by condition combinations.

TABLE 2: HEALTH CARE USE AND DISABILITY OF PRESCRIPTION DRUG USERS, 1996-2015, INPATIENT VISITS

	Dependent: number of inpatient visits ($IPNUM_{ic}$)					
	(1) Baseline	(2) Excl. singletons	(3) Unweighted	(4) Year FE	(5) Excl. vintage	(6) Interaction
off-label%$_{ic}$ ($\hat{\beta}_1$)	0.0029***	0.0029***	0.0039***	0.0029**	0.0032***	
	(0.0009)	(0.0009)	(0.0005)	(0.0009)	(0.0009)	
vintage$_{ic}$ ($\hat{\beta}_2$)	-0.0019*	-0.0019*	-0.0023***	-0.0019*		
	(0.0008)	(0.0008)	(0.0004)	(0.0008)		
year				-0.0011***		
				(0.0003)		
1{(off-label%$_{ic}$)<50%} x vintage$_{ic}$						-0.0001
						(0.0009)
1{(off-label%$_{ic}$)≥50%} x vintage$_{ic}$						-0.0050***
						(0.0012)
Mean	0.039	0.039	0.034	0.039	0.039	0.039
N	13,098,667	13,085,436	860,842	13,098,667	13,098,667	13,098,667
Individuals	201,489	188,258	201,489	201,489	201,489	201,489
R-squared	0.474	0.473	0.343	0.474	0.474	0.474
F-Statistic	8.621	8.621	54.393	10.013	12.654	8.408
RMSE	0.197	0.197	0.201	0.197	0.197	0.197

Note: * p<0.05, ** p<0.01, *** p<0.001; The table shows the effect estimates of one standard deviation increase in off-label% and vintage from two-way fixed-effects regressions for the number of inpatient visits. Column (1) is the baseline estimate including singleton observations, individual and condition fixed effects and observations weighted by the number of prescriptons used by individual i for condition c. Column (2) excludes singleton observations, i.e. individuals reporting one condition only. Column (3) provides unweighted estimates.Column 4 adds fixed effects for the year of the MEPS survey. Column (5) excludes vintages as confounding variable. Column (6) reports interaction effects of off-label use and vintage. Condition and individual fixed effects are included. Standard Errors are clustered at person level. Data on health care use and disability were obtained from the Medical Expenditure Panel Survey, 1997-2015, N=553,051 individual by condition combinations and n= 201,489 individuals. Data on labelled drugs and conditions was obtained from Thériaque. The number of observations (N) in the table reflects the weighted number of observations, except in column (3).

FIGURE 6: PERCENT CHANGE FROM ONE STANDARD DEVIATION INCREASE IN OFF-LABEL% AND VINTAGE BY TYPES OF HEALTH CARE USE AND DISABILITY



Note: This figure shows the baseline estimate and the range of estimates per outcome. The baseline estimates are shown by the black points and are based on two-way fixed effects regressions of off-label% use and vintage, weighted by the number of prescriptions, including condition and individual level fixed effects. The gray bars indicate the range of the estimates based on model variations of the baseline model: excluding singleton observations, regressions unweighted by number of prescriptions, excluding vintage. IPNUM: # inpatient events; ERNUM # emergency events; HHNUM # home health events; OBNUM # office-based events; OPNUM # out-patient events; RXNUM: # precribed medicines; Disability: INBEDFLG individual had bed days: yes / no; MISSSCHL missed school days: yes / no; MISSWORK

TABLE 3: HEALTH CARE USE AND DISABILITY ATTRIBUTED TO OFF-LABEL VS. ON-LABEL USE

	Pooled estimates		
	(1) Estimate off-label%	(2) Estimate as fraction of mean	(3) Cost of off-label use 2015
Outcome	$\hat{\beta}_1$	$\hat{\beta}_1/\bar{y}^k_{2015}$	$(\hat{\beta}_1/\bar{y}^k_{2015})^\star\ \bar{c}^k_{2015}$
Health care use			
# Inpatient events	0.0082	21.91%	$290.56
# Emergency events	0.0100	12.35%	$48.69
# Home health events	0.0676	70.18%	$171.42
# Office-based events	0.5607	33.62%	$268.70
# Out-patient events	0.0485	32.07%	$18.78
# Prescribed medicines	0.5090	25.23%	$292.80
Disability			
Individual had bed days: yes / no	0.0276	21.95%	
Missed school days: yes / no	0.0011	1.92%	
Missed work days: yes / no	0.0153	14.89%	$78.78
Total expenditures			
44% of drugs used off-label replaced by on-label use (% of total cost)			$514.68 (12.92%)

Note: The table shows effect estimates of general linear regressions of the fraction of treatments used off-label including condition and person-level fixed effects by age group category. The estimation sample includes 553,051 condition-drug combinations and 201,489 individuals based on the Medical Expenditure Panel SurveyMean outcomes are for 1996-2015, mean expenditures are based on 2015. The cost of a day of missed work / school were obtained the Bureau of Labour Statistics (https://www.bls.gov/cps/aa2015/cpsaat46.htm) considering the illness or injury lost worktime rate 2012-2015.

TABLE 4: ESTIMATES OF MARKET SIZE AND VINTAGE OF ASSOCIATIONS WITH THE FRACTION OF PRESCRIPTION DRUGS USED OFF-LABEL

Dependent Variable:	off-label%$_{ic}$				
	(1)	(2)	(3)	(4)	(5)
Constant	0.9317***	14.1265***	0.6885***	12.2014***	11.8128***
	(0.0061)	(0.0720)	(0.0011)	(0.0704)	(0.0702)
Market size					
1{(1<prevalence$_c$≤106)>0}	reference			reference	
1{(111<prevalence$_c$≤348)>0}	-0.0156*			-0.0175*	
	(0.0068)			(0.0075)	
1{(357<prevalence$_c$≤1261)>0}	-0.0771***			-0.0737***	
	(0.0064)			(0.0070)	
1{(1305<prevalence$_c$≤4572)>0}	-0.1881***			-0.1737***	
	(0.0063)			(0.0068)	
1{(4614<prevalence$_c$≤87838)>0}	-0.4978***			-0.4570***	
	(0.0061)			(0.0067)	
Vintage		-0.0069***		-0.0057***	
vintage$_{ic}$		(0.0000)		(0.0000)	
		(0.0000)		(0.0000)	
1{(1898<vintage$_{ic}$≤1965)>0}			reference		
1{(1965<vintage$_{ic}$≤1974)>0}			-0.1246***		
			(0.0017)		
1{(1974<vintage$_{ic}$≤1986)>0}			-0.2515***		
			(0.0017)		
1{(1986<vintage$_{ic}$≤1993)>0}			-0.2102***		
			(0.0017)		
1{(1993<vintage$_{ic}$≤2012)>0}			-0.3346***		
			(0.0017)		
Market size x vintage$_{ic}$					
1{(1<prevalence$_c$≤106)>0} x vintage$_{ic}$					-0.0055***
					(0.0000)
1{(111<prevalence$_c$≤348)>0} x vintage$_{ic}$					-0.0055***
					(0.0000)
1{(357<prevalence$_c$≤1261)>0} x vintage$_{ic}$					-0.0056***
					(0.0000)
1{(1305<prevalence$_c$≤4572)>0} x vintage$_{ic}$					-0.0056***
					(0.0000)
1{(4614<prevalence$_c$≤87838)>0} x vintage$_{ic}$					-0.0057***
					(0.0000)
Mean	0.504	0.504	0.504	0.504	0.504
Observations	550,146	550,147	550,147	550,146	550,146
R-squared	0.62	0.58	0.39	0.62	0.62
F-Statistic	260,951	745,501	132,040	208,390	211,964
RMSE	0.41	0.43	0.52	0.41	0.41

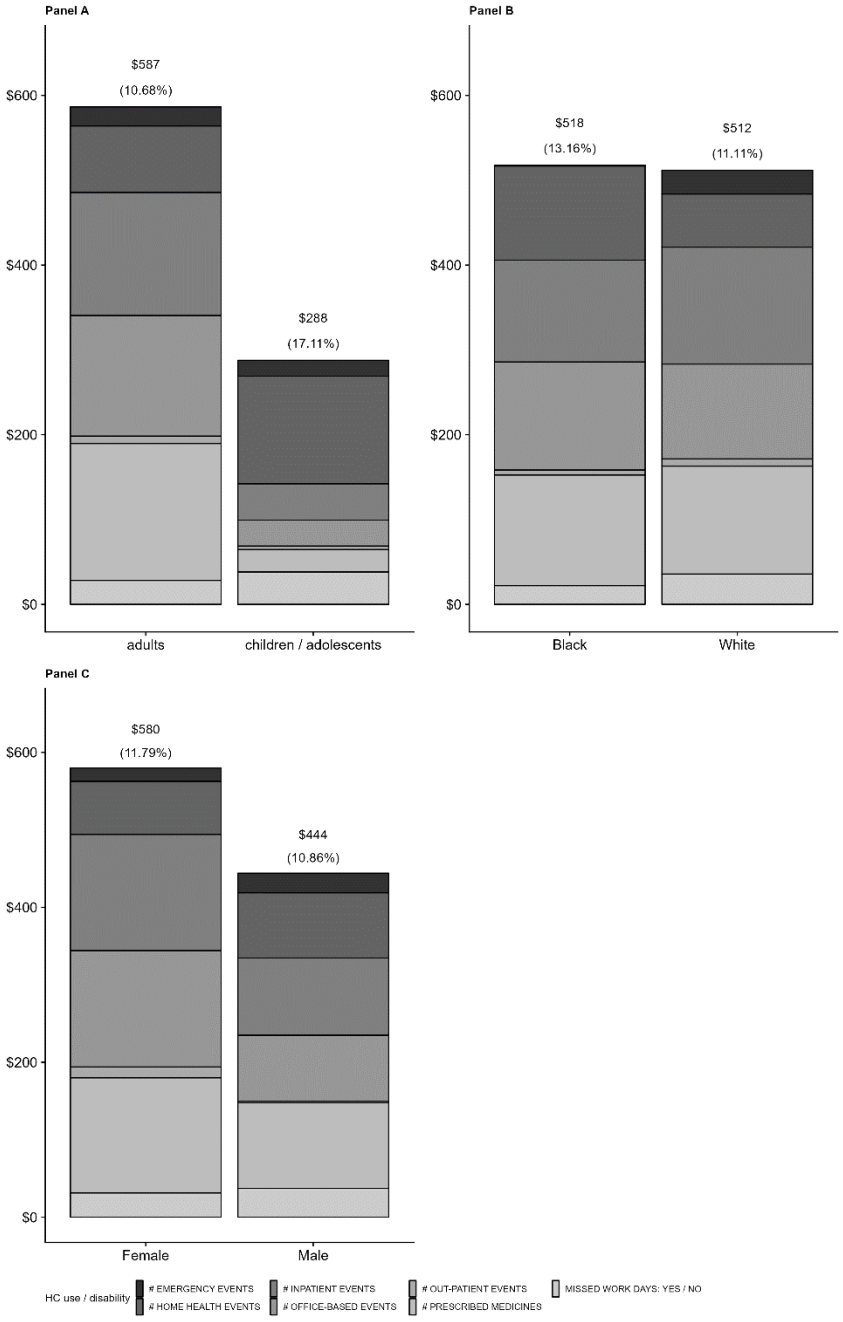
Note: * p<0.05, ** p<0.01, *** p<0.001. This table shows marginal effects of linear regressions of the fraction of prescription drugs uses off-label and market size and the vintage of the drugs used by individual i in condition c. Market size was approximated by the number of patients using a prescription drug for condition c. The estimation sample includes 550,147 condition-drug combinations and 201,489 individuals based on the Medical Expenditure Panel Survey. Estimates are similar when we weight the cross-sectional regressions by the number of prescription drugs used.

TABLE 5: PRESCRIPTION DRUG OFF-LABEL USE AND DISEASE SEVERITY AT INDIVIDUAL LEVEL, 1996-2015

	Dependent: off-label%$_i$					
	(1)	(2)	(3)	(4)	(5)	(6)
# conditions$_i$	0.1067***	0.1044***	0.0204***	0.0177***	-0.0007*	
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)	
Condition severity ($\hat{\pi}_i$)	0.5455***		0.5354***	0.5003***	0.5001***	
		(0.0099)		(0.0067)	(0.0052)	(0.0051)
In-patient events residual ($\hat{e}_i$)		0.0134***		0.0018	-0.0003	-0.0004
		(0.0037)		(0.0023)	(0.0016)	(0.0016)
1{(male$_i$=1)>0}			0.1363***	0.1361***	0.0008	0.0004
			(0.0017)	(0.0017)	(0.0016)	(0.0016)
1{(black$_i$=1)>0}			0.0528***	0.0526***	-0.0259***	-0.0258***
			(0.0025)	(0.0024)	(0.0021)	(0.0021)
education (years)			0.0277***	0.0272***	0.0016***	0.0016***
			(0.0002)	(0.0002)	(0.0002)	(0.0002)
age$_i$			0.0005***	0.0006***	-0.0021***	-0.0020***
			(0.0000)	(0.0000)	(0.0000)	(0.0000)
vintage$_i$					0.0003***	0.0003***
					(0.0000)	(0.0000)
1/2 conditions						reference
3-9 conditions						-0.0096***
						(0.0016)
10 or more conditions						-0.0030
						(0.0031)
N	180,496	180,496	180,496	180,496	180,496	180,496
R-squared	0.405	0.416	0.608	0.619	0.720	0.720
F-Statistic	109.00	35.15	80.80	57.56	80.66	67.63
RMSE	0.473	0.469	0.384	0.379	0.325	0.324

Note: * p<0.05, ** p<0.01, *** p<0.001. Estimates of associations of the number of conditions reported by individual and the mean deviation of off-label%, disease severity based on sum of condition level fixed effects for the number of inpatient events. The dependent variable is the deviation from the mean off-label%. Data were obtained from the Medical Expenditure Panel Survey, 1997-2015. Data are aggregated to person level. The table shows the estimates of general linear models of off-label% at person level. Robust standard errors are reported.

FIGURE 7: HETEROGENEITY IN HEALTH CARE USE AND DISABILITY, BY POPULATION GROUPS



Note: This figure shows the medical expenditure and work loss cost attributed to off-label use by three types of population subgroups. Panel A shows estimates by adults (n=458,043) and children / adolescents up to the age of 18 (n=87,948). Panel B shows separate estimates by whether the individual is black (n=28,968) or white (n=172,497). Panel C shows separate estimates by males (n=89,327) and females (n=112,138). Estimates were obtained from two-way fixed effects regressions where the dependent variable is one of the seven outcomes of

health care use / disability for person i and condition c. The estimate of interest is the share of prescription drugs used off-label by person in condition c. All regressions include vintage, person-level and condition-level fixed effects. Estimates of off-label use were multiplied by the average cost per health care or disability related measure in 2015. Mean outcomes are for 1996-2015, mean expenditures are based on 2015. The cost of a day of missed work / school were obtained the Bureau of Labour Statistics (https://www.bls.gov/cps/aa2015/cpsaat46.htm) considering the illness or injury lost worktime rate 2012-2015. Percentages in brackets display the aggregate share of medical expenditure and work loss cost that is attributed to eliminating off-label use at the average 2015 population average, i.e. 44%.

Appendix

Off-label use, approval status and health care use- theoretical background

To describe how the use of technology like prescription drugs is leading to changes in health care use and disability, we draw from the considerations how individuals accumulate health deficits over time and the literature (Böhm, Grossmann, and Strulik 2021; Dalgaard and Strulik 2014; G. M. Grossman and Helpman 1991) on technological change. The purpose is to specify parameters for the empirical investigation and their structural relationships.

We assume that there is a continuum of medical conditions $c = [1, \dots, C]$ and a continuum of technologies $d = [1, \dots, D]$ with a certain quality q. For technology d to be appropriate to treat medical condition c designated for on-label use, the technology needs to prove its effectiveness such that $q_{cd} = \alpha > 0$. In biopharmaceutical research and development, this form of effectiveness is typically investigated by clinical trials of phase III. This evidence is used in applications to marketing authorization bodies like the US FDA, which investigate whether the specific technology-condition combination $\{c, d\}$ proves this type of effectiveness. In a large number of technologies, quality is assessed in comparison with a placebo technology such that we can assume that the effect just needs to be larger than zero. Accordingly, for each condition, there is a sub-set of technologies $d_c = [1, \dots D_c] \subset D$ that proves to be effective and is authorized to treat condition c. Technologies that do not show any effect (that means they do not obtain marketing authorization), or drug-condition combinations $\{c^-, d\}$ for which marketing authorization has not been requested, are considered inappropriate and not effective, that means $q_{c^- d} = 0$.

To consider variation in quality by condition and the role of innovation, we rely on endogenous growth theory which suggests that biopharmaceutical research and development typically aims to advance treatment quality. The general assumption is that newer vintage goods are, on average, of higher quality than older vintage goods. Grossman and Helpman (1991) have demonstrated that every industry (medical conditions that are targeted by biopharmaceutical therapies that stem out from R&D in our case) has its unique quality ladder. Most importantly, "technology-based growth is a process of generating an ever-expanding variety of horizontally differentiated products." (G. M. Grossman and Helpman 1991, 44). That means that over time, successful innovators will improve the quality level of technologies appropriate to treat condition c. Accordingly, we assume that technologies within a condition-specific subset $d_c = [1, \dots D_c] \subset D$ improve in quality over time. That means that a technology d that is commercialized in $t+1$ has the quality $q_{t+1}\{cd\} = \alpha + \gamma > q_t\{cd\} > 0$. The expected quality of a technology to effectively treat condition c will then be determined whether it is appropriate to use for condition c and its relative quality (captured by its vintage) within the continuum of qualities in the sub-set of technologies d_c.[17] An important implication is that across conditions and time, the average quality of the latest vintages ("stock of medical knowledge") $Q_t \equiv \int_0^1 q_t(c)dc$ will depend whether there is at least one appropriate active pharmaceutical ingredient and the vintages of active pharmaceutical ingredients available in the subset of technologies that are appropriate to treat condition c.

To describe the relationship between health inputs, the quality of the technology and subsequent outputs in terms of health care use and disability, assume that individual i suffers from a condition c. We further assume that the

[17] Böhm et al. (2021) consider a dynamic case where individuals accumulate health deficits based on a set of illnesses (medical conditions) that an individual suffers from. For the moment, we refrain from cumulating health deficits acquired through different illnesses and analyze the effects at the condition-by-patient level.

individual generally has access to all available technology authorized by the FDA, that means drugs that have obtained marketing authorization independent of whether the drug is used on or off-label.[18] Endogenous growth theory suggests that newer vintage drugs are of higher quality compared to older vintage drugs which leads us to assume that, on average, newer drugs with a designated label are better than older drugs (G. M. Grossman and Helpman 1991; Solow 1960). The expected input for the treatment of medical condition c is then defined by the quality of the technology used to treat condition c is a function of whether it is indicated to treat condition c based on regulatory approval and its vintage v, that means

$$q_c(v) = \begin{cases} > 0, \text{ if used on-label in condition } c \\ 0, \text{ if used off-label in condition } c \end{cases}, \text{ and } \frac{\partial q_c}{\partial v} > 0 \quad \text{(A1)}$$

As there may be off-label uses that prove effective outside regulatory approval that typically requires more comprehensive data compared to clinical trials, we could relax this assumption and assume that $q_{c,on-label}(v) - q_{c,off-label}(v) > 0$, that means all on-label use is superior to any off-label use for the same condition.

The expected effect E_{ic}^{k} on outcomes k in the dimensions of health care use and disability from using treatments for individual i is then dependent on whether these are used on-label and therefore whether they have been approved for their designated use and their vintage that means when they have been approved for their designated use:

$$E_{ic}^{k} = \int_d \; q_c(v) + \varepsilon_i \qquad \text{(A2)}$$

In this model, the outcomes denoted by E_{ic}^{k} reflect that using a drug has direct consequences on health care related events such as hospital or ambulatory care visits that are indicative of the effectiveness of the prescription drug treatment. In addition, given that the use of an on-label drug shows a positive effect, it will reduce

[18] This may be a strong assumption as health insurance and pharmacy benefit plans may restrict access to certain technology.

an individual's disability that we will operationalize in the form of having to spend a day in bed, missing school or work in the empirical application. ε_i accounts for the fact that the effectiveness of the treatment depends on whether physicians and individuals themselves make the right choice, behavioral factors and pre-existing health deficits of the individual. The term ε_i allows for the possibility that consumers make irrational decisions when choosing technologies and that the same treatment may be more effective in one individual compared to another. That means that individuals use products that are not necessarily good for them. Variation in decision-making across individuals may be partly explained by factors such as socio-economic characteristics and severity of the disease besides differences in genetics, metabolism and hormonal status. In addition, choosing the right pharmaceutical not only is the decision of the individual that suffers from a condition, but always involves physician decision-making. Physicians are typically the ones held responsible for making the wrong choices about a prescription drug. It has been demonstrated that physicians often deviate from the best available treatment option (Frank and Zeckhauser 2007; Janakiraman et al. 2008).