Happy pills? The health consequences of the dramatic increase in antidepressant use*

Giuliano Masiero† Fabrizio Mazzonna‡ Sandro Steinbach§

July 12, 2021

Abstract

This paper investigates the impact of the dramatic increase in antidepressant consumption on population health and labor market outcomes in Switzerland between 2002 and 2014. Our research design is based on a modified version of the popular shift-share instrument. This instrument exploits the large and heterogenous growth in antidepressant sales due to product innovation (the shift), and the local market power of pharmaceutical companies (the share). Our estimates show that an increase in antidepressant sales causes a sharp increase in hospital admissions related to depression symptoms but we do not find any economically relevant effects on labor market outcomes.

Keywords: Depression; antidepressant treatment; suicides; mental health
JEL codes: I12, I18, J21

*We thank Olaf Verbeek for his initial research support on this project and Panka Bencsik, Christopher Cronin, Osea Giuntella, and Nicolas Ziebarth for insightful comments on early drafts of this paper. We also thank seminar participants at the University of Pittsburgh 2020 and session participants at CINCH+ Mental Health Workshop 2020 and ASSA meeting 2021.
†Department of Economics (DSE), University of Bergamo, Italy, and Institute of Economics (IdEP), Università della Svizzera Italiana (USI), Switzerland.
‡Corresponding author: Institute of Economics (IdEP), Università della Svizzera Italiana (USI), Switzerland (fabrizio.mazzonna@usi.ch), Institute for Labor Economics (IZA) and Munich Center for the Economics of Aging (MEA).
§Department of Agricultural and Resource Economics, University of Connecticut, United States.

1. Introduction

Antidepressants are among the most prescribed drugs in the world, and their use in OECD countries has more than doubled in the last 15 years (OECD, 2019).[1] In the US, antidepressants are the third most prescribed class of drugs, with 13% of people aged 12 and over who reported taking antidepressants within the last month (Pratt et al., 2017). This considerable growth can be seen as a positive outcome for society, given the well-known economic burden of depression associated with increasing disability and absenteeism (e.g., James et al., 2018) and lower productivity (Stewart et al., 2003). However, in many countries, the boost in antidepressant use has been driven by non-psychiatric prescriptions in primary care, most of them without depression diagnoses, as a result of off-label use (e.g., Mojtabai and Olfson, 2011; Wong et al., 2017). There is also rising concern over the efficacy of antidepressants and the possible occurrence of severe side-effects, predominantly related to overprescription and their use in the long-term.[2] Although initial evidence about antidepressant efficacy (and drugs approval) is mostly grounded on randomized controlled trials (RCT), several recent meta-analyses find evidence of clinical efficacy only for the most severely depressed (e.g., Fournier et al., 2010), and in most cases, the placebo effect accounts for about 80% of the total measured effect (Currie and MacLeod, 2020).

This study investigates the health consequences of the recent upsurge in antidepressant use in Switzerland, showing causal evidence of adverse health outcomes. We exploit the large geographical and temporal variation in antidepressant sales at the product level and individual hospital admissions and suicide events for 13 years, between 2002 and 2014, for the whole Swiss country (see Figure 1). The Swiss case is particularly relevant since antidepressant consumption is very close to the OECD average and has increased in the last 15 years by over 60%. New antidepressant drugs have enlarged the market by reaching new consumers rather than reducing the market for previously branded drugs. In 2016, the prevalence rate of antidepressant prescriptions was 8.7%, far above the estimated number of people with major depressive symptoms, 5.2% (Haller et al., 2019). Moreover, the country is home to many large pharmaceutical companies, and the pharmaceutical industry represents almost 5% of Swiss GDP.

[1] The worldwide increase in antidepressant use can be traced back to the beginning of the '90s, with the introduction of the Selective Serotonin Reuptake Inhibitors (SSRI), which are known to have better tolerability than the old classes of antidepressants, the Tricyclic antidepressants (TCA).

[2] Awareness for adverse events was spurred in October 2004 when the US Food and Drug Administration (FDA) issued a black box warning for all antidepressants. It is worth mentioning that Busch et al. (2014) find that the FDA warning had some unintended consequences on human capital development in adolescents, affecting their educational performance and delinquency outcomes.

We address the endogeneity concern that arises in ecological studies using two complementary instrumental variable (IV) approaches. Our main instrument is inspired by the popular shift-share approach (Bartik, 1991), but substantially diverges regarding its implementation. Similar to the standard shift-share instrument, the predicted antidepressant sales in a region[3] is a weighted average of the national pharmaceutical company growth rates (the shifts), but the weights depend on the pharmaceutical company's regional market shares for non-antidepressant drugs in the base year (the shares). The variation in the shares comes from historically grown differences in market power between pharmaceutical companies in different Swiss regions, while the variation in the growth rates (net of year and region fixed effects) mainly comes from the introduction of new products in the market (as reported in Figure 2), especially generic drugs. The nature of the instrument allows us to alleviate concerns regarding the correlation between the initial shares and health conditions of the region because we exploit the plausibly exogenous variation in the market power of pharmaceutical companies in the non-antidepressant market. So the identifying variation comes mainly from multi-product companies (generics drug makers and large companies) that are more likely to sell new antidepressants in regions with larger non-antidepressant market shares as compared to their competitors.

Using the before described research design, we find that an increase in sales of one defined daily dose per 1,000 inhabitants (roughly 3% of 2003 sales) increases emergency hospital admissions for mental disorders by 1.8%, driven by an even larger increase in hospitalizations for depression symptoms (5.6%). We also estimate a log-log model that leads to comparable estimates in elasticity terms (.8 for mental disorders and 2.5 for depression). The evidence on suicides is generally positive, although point estimates are noisier. Furthermore, since our results might only reflect the side effect of an otherwise successful increase in depression treatment, we use data from the Swiss labor force survey to investigate the presence of spillovers into the labor market. Conversely, from the recent evidence by Bütikofer et al. (2020) and Shapiro (2020), supporting the presence of positive labor market effects from antidepressant use, we can reasonably reject any economically relevant effect on unemployment, labor force participation, income, and working hours.

Following the recent literature that investigates the formal conditions underlying the validity of the Bartik instrument (e.g., Borusyak et al., 2018; Goldsmith-Pinkham et al., 2020), we provide

[3] Our regions are very small local labor markets. Their average population is less than half the average population of US counties.

more formal tests about the validity of our identification strategy. As shown by Goldsmith-Pinkham et al. (2020), the Bartik instrument is equivalent to using the initial shares of each pharmaceutical company (interacted with time fixed effects) as instruments in a weighted GMM estimation. Therefore, we calculate the so-called "Rotemberg weights" on these instruments, which allows us to pick out pharmaceutical companies that account for the largest share of the identifying variation. We observe that two-thirds of the variation comes from one company, which experiences a dramatic increase in antidepressant sales in the Swiss market over the observation window. The reason rests on the introduction of new generics drugs. As shown in Figure 3, we find no evidence of differential pre-trends in the main health outcome between regions with high and low market shares for this company. This result also holds for the other major weight companies. Even when the top weight company is excluded from our regression, the estimated effects are essentially unchanged.

In addition, we run some analyses to test indirectly the exogeneity of our instrument using placebo outcomes, i.e., hospital admissions for diseases that should not be affected by antidepressant use. Here, we remark the lack of relevant correlation between the increase in antidepressant sales induced by our instrument and placebo hospital admissions. As a final robustness check, we use a second instrument that relies on the hypothesis that antidepressant prescribing practices in relatively big neighboring countries generate spillover effects in small Swiss regions and influence doctors' prescribing behavior. This is especially true for Switzerland, where more than 25% of doctors study in one of the three main neighboring countries (France, Germany, and Italy). Therefore, our second instrument builds on the assumption that the magnitude of medical practice spillovers is inversely related to geographical distance. We calculate spatially weighted averages of antidepressant sales in neighboring countries and assign them to each small region. It is reassuring to see that the point estimates obtained with this alternative estimation strategy and the previous ones are very similar.

This paper relates to the extensive literature on the spatial correlation between antidepressant consumption and suicides to evaluate the population health effects in the "real world". Our paper contributes to this literature by providing new causal evidence about the health consequences of the recent surge in antidepressant use. As far as we know, Ludwig et al. (2009) is the only study providing plausible causal evidence on the relationship between antidepressant consumption and suicides at the population level. The authors exploit the differential introduction of SSRI

drugs across countries between 1980 and 2000 and find that an increase of SSRI sales by one pill per capita reduces suicides by 5%. Although insightful, this result is no longer applicable to the current level of antidepressant consumption. Nowadays, a large part of antidepressants consumed is SSRI. Likely, the marginal patient treated with antidepressants suffers from a milder form of depression as compared to the marginal patient at the beginning of the '80s, when SSRI was introduced, even without considering off-label use. Also, antidepressant consumption in the US has increased by 400% between the early nineties and the beginning of this century, and similar trends are observed in most other developed countries.

Another important contribution of our paper is the focus on the effect of antidepressants on hospitalization rather than relying only on suicides (e.g., Ludwig et al., 2009). While suicide is a very low probability event and its statistical power tends to be low (Currie and MacLeod, 2020), hospitalizations are by large more frequent and have a remarkable impact on health care costs. Our results complement recent evidence by Cuddy and Currie (2020a,b) who show that inappropriate mental health treatment with antidepressants in adolescent children has an impact on the future total cost of care and poses health risks. In particular, children who receive inappropriate drug treatment ("red-flag drugs") have a 131% higher probability of using an emergency room or being hospitalized. Finally, our findings are also consistent with Currie and MacLeod (2020), who show that doctor violation of practice guidelines for adult patients is associated with increased emergency room visits and hospitalizations.

Our paper is further related to the large literature in economics on the causal link between pharmaceutical marketing and prescription drug utilization, including antidepressants. However, most of this literature focuses on direct-to-consumer advertising (e.g., Shapiro, 2018; Sinkinson and Starc, 2019).[4] We contribute to this literature by showing evidence of a different marketing channel. The robust predicting power of our instrument implies that at least part of the increase in antidepressant use can be attributed to the market power of pharmaceutical companies and their ability to promote products among physicians. This evidence is supported by a complementary investigation on the relationship between the market share of pharmaceutical companies and payments to health care professionals (provided through voluntary disclosure).

The remainder of the paper is organized as follows. In the next section, we introduce a simple conceptual framework to explain the relationship between depression severity and the benefits

[4] As explained in Section 2.2, direct-to-consumer advertising is not allowed in Switzerland, and prices are set at the federal level, implying that pharmaceutical companies compete over physician detailing.

of antidepressant treatment. We also present the Swiss institutional setting where we conduct our analysis. In Section 3, we describe our data and their aggregation process. In Section 4, we discuss our identification strategy, while in Section 5 we report our results and complement the discussion with some robustness checks. Finally, Section 6 concludes.

2. Background

2.1 Conceptual framework

As already mentioned, there are growing concerns over the efficacy of antidepressants. Critics point to methodological issues (e.g., RCT duration and sample size), under-reporting of adverse events during the trials, and questionable clinical significance (e.g., Jakobsen et al., 2017). Moreover, their use is associated with an increase in suicide risk among pediatric patients and adolescents (Cipriani et al., 2016; Stone et al., 2009; Bridge et al., 2007), and with several adverse health outcomes among the elderly (Coupland et al., 2011). Increasing suicidal behavior seems to be the result of an early "activating effect" that gives patients the energy to follow their suicidal impulses (Stone et al., 2009). Also, the inappropriate use of antidepressants, especially SSRI, is frequently associated to withdrawal symptoms, which are particularly severe in some cases (Davies and Read, 2019).[5] Not less important are the unknown consequences of off-label prescriptions, which may "*expose patients to unknown health risks if their clinical characteristics differ from the patient population studied in clinical trials*" (Wong et al., 2017).

Taking this evidence into account, we present a simple conceptual framework that describes the potential relationship between depression severity and the benefits of antidepressant treatment (Figure 4). In particular, we assume that the more severe the depression, the greater the expected benefits of drug treatment. However, in the case of mild to moderate depression the expected benefits are small or even negative.

For simplicity, we assume a positive non-linear relationship between depression severity and the expected effects of antidepressant treatment for both classes of antidepressants, namely SSRI and TCA. Since SSRIs are expected to have milder or fewer side effects than TCA, i.e., the older class of antidepressants, we draw the SSRI benefit line above the TCA line. As a result, prescription thresholds at the end of the '80s —based on TCA guidelines and their expected benefits— were too stringent for SSRI, the newer class of antidepressants. This might

[5] Typical antidepressant withdrawal reactions include increased anxiety, flu-like symptoms, insomnia, nausea, imbalance, sensory disturbances, and hyperarousal.

explain why Ludwig et al. (2009) find evidence of positive mental health effects after the SSRI introduction. Since then, prescription thresholds shifted towards milder forms of depression (Moore et al., 2009), as supported by the extensive evidence of over-treatment for mild cases of depression. Graphically, this is represented by the current prescription threshold laying on the left-hand side of the prescription threshold suggested by the current guidelines. As a result, we might expect negative effects on the marginal patient treated today with antidepressant drugs, while greater adherence to guidelines would provide non-negative benefits to all people treated with antidepressants. It is worth noting that, even if treatment with antidepressants provides positive benefits to several patients, an empirical investigation that focuses on the marginal patient might find evidence of negative effects.

Finally, Figure 4 assumes some form of monotonicity in drug treatment, but we are aware that the current situation is more complicated. In other words, there is both evidence of over-treatment and under-treatment, meaning that we might have people with severe depression who are not currently treated, while other people with very mild forms of depression resort to antidepressants (e.g., Editors et al., 2013).

2.2 Institutional setting

Switzerland is a confederation of 26 cantons with considerable autonomy in the organization and the provision of health care services. The supply of mental health care is a cantonal responsibility, though the federal state organizes some of the fundamental financial aspects (Biller-Andorno and Zeltner, 2015). Private health insurance is mandatory and regulated by federal laws. The insurance plan covers an extensive list of prescription drugs and, therefore, Swiss consumers face almost no costs when using antidepressants. A consumer can opt for a Health Maintenance Organization (HMO) type of health insurance or a general practitioner (GP) scheme, which allows the consumer to reduce the insurance premium. Cantonal authorities provide subsidies for those consumers facing financial hardship. The minimum annual deductible amounts to 300 CHF (1 CHF $\approx$ 1 US \$), but the consumer can choose a higher deductible, up to 2,500 CHF, against a decrease in the insurance premium. After the deductible is exhausted, the consumer contributes by 10% to all health care expenses, up to a stop-loss amount of 700 CHF. Moreover, the federal government introduced a 20% co-insurance rate for off-patent brand name medications in 2006.

Individuals who suffer from mental disorders generally opt for the minimum deductible. More-

over, the deductible is quickly exhausted by physician visits and psychotherapy consultations.[6] To provide an idea of the potential costs of antidepressant treatment for a patient, the price per defined daily dose for the most prescribed (brand name) drug in Switzerland (Cipralex 10mg) is about 1.32 CHF, or 480 CHF a year. According to the drug list, this drug has a 20% co-insurance rate. Therefore, a patient treated with this drug pays at most 336 CHF a year out of pocket. No antidepressant is available "over the counter" since all antidepressants without exception are prescription drugs. Lastly, Masiero et al. (2018) find that antidepressant consumption in Switzerland is associated with physician density, suggesting that supplies may induce demand at least to some extent. All in all, consumers determine the demand for antidepressant drugs only to a small degree.

The Federal Office of Public Health (FOPH) sets prices for prescription drugs in Switzerland. After a drug has been granted access to the Swiss market by the federal authority (Swissmedic), the FOPH decides whether to include the drug in the list for reimbursement (Spezialitätenlist - SL), upon evaluation of its efficacy. Antidepressants are relatively expensive in Switzerland, and the price difference between brand names and generic drugs is not very large. Generic drugs are at least 50% more costly than in other European countries.[7] These market characteristics suggest that pharmaceutical companies are likely to compete on quantity rather than in prices. Since only physicians can prescribe antidepressants and federal laws prohibit direct-to-consumer advertising for prescription drugs, pharmaceutical companies can only influence their sales through physician detailing.

During the observation window, pharmaceutical companies were not legally forced to disclose any information regarding their physician detailing activity. Only at the end of 2013, the majority of pharmaceutical companies agreed to adopt a code of conduct that includes the voluntary disclosure on their website of pecuniary benefits provided to health care professionals.[8] We will use this information to bring evidence on the strong correlation between physical detailing and local market shares of pharmaceutical companies.

[6] According to the Swiss tariff system for out-patient medical services (TARMED), the cost for a psychiatric consultation amounts to 11.20 CHF per five minutes. A psychotherapeutic consultation or a GP consultation amount to 10.42 CHF per five minutes. For instance, with only two hours of treatment per month, the deductible is already exhausted in a couple of months. For the remaining part of the year, the patient only pays the co-insurance rate.

[7] See the recent press release by the Swiss health insurance association (Santésuisse, 2017) on the international comparison between generic drug prices in Switzerland and prices in Belgium, Denmark, Germany, Finland, France, Great Britain, the Netherlands, Austria, and Sweden.

[8] Here the link to the document: `https://www.scienceindustries.ch/_file/24299/pharma-kooperations-kodex-april-2019-e.pdf`

The prevalence rate of mental health problems in Switzerland is similar to other developed countries (Schuler and Burla, 2012). Severe cases can be treated both in private and public hospitals. Hospitals charge a daily fee which decreases with the length of stay. Cantons and health insurance providers share the costs of psychiatric hospital stays. Although hospital admissions for mental health disorders have increased over time, the number of psychiatric hospital beds per capita has declined, and a growing number of patients is treated in outpatient settings. Similar trends are observable for other European countries (Priebe et al., 2008). The fees for the services provided by outpatient departments/clinics are standardized in the TARMED tariff system to avoid differential treatment of patients with different insurance plans.

3. Data

We exploit two primary datasets on antidepressant sales and hospital discharge data for Switzerland, covering 2002 to 2014. The data are aggregated at small region level (SMR - spatial mobility region). This level divides Switzerland into 106 SMRs, each of which accounts for approximately 45,000 individuals. An SMR is a statistical subdivision of the country based on economic activity around an agglomeration hub. As such, each region represents a local labor market or commuting zone.

The level of disaggregation allows us to account for population characteristics and neglect possible consumption spillovers across regions. Indeed, people living in one region are highly unlikely to work in a neighboring region and, therefore, to shop for antidepressants outside the SMR of residence. Thus, measuring antidepressant consumption at the level of commuting zones represents an effective way to deal with the potential for measurement error. Nonetheless, an additional source of measurement error may arise from the use of wholesale data (from the pharmaceutical company to the pharmacy/drugstore) since we measure what the pharmacist stocks rather than the sales to the final consumer. Although our observations of final consumption are on average correct, we could overestimate antidepressant use in some cases.

We supplement our primary datasets with three additional sources of data: *i*) labor market outcomes from the Swiss Labor Force Survey (SLFS); *ii*) mortality data from the official Swiss mortality statistics; *iii*) essential covariates for each region and year from various sources. In the following, we provide a detailed description of each data source.

Antidepressant sales

We obtained data on antidepressant sales from IMS Health Switzerland (now IQVIA). Our dataset contains annual antidepressant sales at the product level by pharmaceutical sales region (237 regions) from 2002 to 2014. This level of aggregation includes at least five pharmacies to avoid the identification of specific retailers. The level of detail allows us to calculate the consumption of each antidepressant product in defined daily doses (DDD) per 1,000 inhabitants per year based on information from the WHO dataset on daily doses by active ingredient. In particular, we consider sales data for the following Anatomical Therapeutic Chemical classes (ATC): N06A4 (Selective Serotonin Reuptake Inhibitors - SSRI), N06A5 (Serotonin and Norepinephrine Reuptake Inhibitors - SNRI), and N06A9 (Other antidepressants, including Tricyclic antidepressants - TCA) In Table A.1 we report the active ingredients included in each class. Although herbal medicines (class N06A2) enjoy a high level of acceptance in the Swiss population, we exclude them from our analysis since we cannot define the daily dose for this class of antidepressants.

OECD data on antidepressant consumption in DDD for France, Germany, and Italy from 2003 to 2014 (see Table A.2 for descriptive statistics) are also used to construct our second instrument based on spatial spillovers in prescription practices.[9]

Hospital discharge data and suicides

We obtained individual-level data on hospital discharge from the Federal Statistical Office (FSO). The most detailed geographical aggregation at which hospital admission data are available for Switzerland is the MedStat region level. The MedStat region is a geographical concept used by the FSO to anonymize individual-level hospital admission data.[10] We use a population-weighted matching procedure to reassign data aggregated at the MedStat level to the SMR level, and from the pharmaceutical sales region to the SMR level. The matching method allows

[9] Since national data for France after 2012 are not released, we approximate them using the natural cubic spline interpolation, which code was originally given by Herriot and Reinsch (1973). We do not include Austria, a minor neighboring country sharing a relatively small border with Switzerland in a mostly mountainous area. Not surprisingly, doctors from Austria represent a negligible minority in Switzerland.

[10] An advantage of these data is that the 604 MedStat regions are homogenous regarding the population size, with each of them containing about 12,000 people. It is important to note that the spatial definition was updated in 2008 to account for population growth. Based on postal codes for 2007, the old MedStat regions were split up or combined to form new MedStat regions. Therefore, it is impossible to study hospital admissions over the structural break without reassigning the data from the new to the old definition of MedStat region. We accomplish this task by matching postal codes underlying the MedStat regions over the structural break. To do this, we follow the approach developed by Filippini et al. (2019). In particular, we use detailed information on the general population at the postal code level for 2010 from the FSO to create weights and to recode the location information. As a result, we obtain a match between the new and the old definition of MedStat regions. We then reassign the morbidity data over the structural break using population weights. A further discussion of the spatial concepts is provided in Appendix A.

us to build a final dataset with comparable spatial data on both antidepressant consumption and health outcomes. For descriptive purposes, we express hospital admissions in terms of annual prevalence per 10,000 inhabitants by SMR throughout our analysis.

In our main investigation, we consider total hospital admissions and hospital admissions related to mental health problems. First, we account for hospital admissions for all mental health conditions (Chapter V).[11] Second, we focus specifically on hospital admissions related to depression (depressive episode - F32, and Recurrent Depressive Disorder - F33). Third, to capture the impact of antidepressant drugs on suicides, we combine hospital admissions for suicide attempts (Intentional Self-harm - X60-X84) and completed suicides from official mortality data from FSO. As robustness checks, we split hospital admissions between emergency and elective, and use hospital admissions for infectious and parasitic diseases (ICD code A-B), bone fractures (ICD codes S), and pregnancy, childbirth and the puerperium (ICD codes O) as placebo outcomes.

Swiss Labour Force Survey (SLFS)

The Swiss Labor Force Survey (SLFS) is the main source of information in Switzerland about the structure of the labour force and employment behavior patterns. To account for eligibility for statutory retirement, we focus on the working age population, namely 16–64 for men and 16–63 for women, and drop individuals in military service or in education. In our observation window (2003–2014), we have almost 600,000 observations that we collapse at SMR level using survey weights. The average number of respondents per cell is over 300. Cells with a relatively small number of respondents (less than 30) are excluded. We focus on the following labor market outcomes: unemployment, labor force participation, individual income, mean and total working hours in a week.

Additional datasets

To construct our main instrument, we obtain total drug sales, excluding antidepressants, by pharmaceutical company and region in 2002 (source: IMS Health Switzerland).[12] Since we use such data to predict antidepressant sales, this additional dataset includes only the top 16 companies in the antidepressant drugs market, representing almost 98% of the market over the observation window. Non-antidepressant sales are based on ex-factory prices because we do not have information on retail prices for single products.

[11] All disease codes refer to the International Statistical Classification of Diseases and Related Health Problems, 10th Revision, of the WHO.

[12] A negligible number of drugs with a retail price greater than $5,000$ CHF is also excluded.

We collect the following covariates for each region and year from the FSO: the distribution of the population across gender and age, the share of German-speaking people, the share of foreigners, and the average municipal official unemployment rate (based on unemployment benefit claims).

To investigate the relationship between physician detailing and local pharmaceutical market shares, we collect voluntary disclosure information on payments to health care professionals from 2015 to 2017 from the website of eleven pharmaceutical companies. Such information is then matched at SMR level, year and pharmaceutical company, with pharmaceutical drug sales obtained from the association of Swiss pharmacies (OFAC).[13]

3.1 Descriptive evidence

We summarize the main variables used in our analysis in Table 1. The average antidepressant consumption across SMRs, and for the whole period under study, is more than 40 DDD per 1'000 inhabitants, with an increase of almost twenty DDD over the last decade. The spatial and temporal variation in antidepressant consumption are illustrated in Figure 1. We observe a sharp increase over time and substantial variation across regions. Interestingly, the consumption is mostly concentrated in South-Western regions in 2003, while there is no clear concentration in any region in 2014. Hence, the increase in consumption over time characterizes a catching-up process with the strongest increase in the North-Eastern regions.

A similar pattern is observable for mental health outcomes in Figure 5. In particular, hospital admissions for mental disorders and depression show a significant increase over time. Subfigures (a) and (c) highlight the concentration of mental health disorders and depression in South-Western regions in 2003. Then, according to subfigures (b) and (d), the prevalence of both mental disorders and depression appear to increase in North-Eastern regions over time. The variation in suicide rates provides a far more complex picture, probably due to the rare-event nature of suicides. The comparison of subfigures (e) and (f) does not seem to suggest a clear spatial or temporal pattern, although we see evidence for an increasing number of cases in some Eastern regions.

Figure 2 provides a detailed description of the Swiss antidepressant market over the observation window. Subfigure (a) shows that the market is characterized by substantial product innovation

[13] Since our sales data from IMS Health Switzerland end in 2014, we use the data from the association of pharmacies to cover the period 2015-2017.

with the introduction of several new antidepressants, especially generics.[14] In subfigure (b), we illustrate how this innovation affected overall sales, distinguishing between generic and branded drugs. The figure shows that, at the beginning of the period, the market was dominated by branded drugs. The introduction of various generic products between 2002 and 2005 (due to a number of expired patents) resulted in an exponential increase of generic drug sales. However, the penetration of generic drugs led to an overall expansion of antidepressants, which increased by over 60%, rather than a mere substitution of branded drugs (see Masiero et al., 2018 for details). We discuss the use of such variation to construct our instrument for antidepressant consumption in the next section.

4. Empirical strategy

In this section we outline our empirical strategy to identify the effect of antidepressant sales on hospitalizations and suicides. Following previous literature on the effect of antidepressant use on suicides (e.g. Ludwig et al., 2009), we apply the natural log to hospital admissions and suicides to approximate a normal distribution for our data.[15] We report the results of an alternative log-log specification (taking also the log of antidepressant sales) and of a Poisson count data model in the Appendix. More formally, we rely on the following empirical model:

$$\ln(y_{rt}) = \beta_0 + \beta_1 AD_{rt} + \beta_2 X_{rt} + \vartheta_r + \lambda_t + \epsilon_{rt}, \tag{1}$$

where y_{rt} is our outcome variable (number of hospital admissions or number of suicides)[16] in region r at time t; AD_{rt} represents antidepressant sales expressed in DDD per 1,000 inhabitants in year t; X_{rt} is a vector of controls, including demographics (the age distribution of the population, the share of females, the share of German speakers, and the share of foreigners); ϑ_r are region fixed effects, λ_t are time fixed effects, and ϵ_{rt} is an idiosyncratic error term. This empirical model is also used to estimate spillover effects of antidepressant sales on labor force outcomes.

We use the fixed effects (FE) estimator as our benchmark model to estimate Equation (1). Given the substantial scale differences between small regions (we move from almost half a million

[14] Additional details (pharmaceutical company, active ingredient and year of introduction) for new brand name products and generic drugs are provided respectively in Table A.3 and Table A.4.

[15] We take care of zeros for suicides by adding a small constant.

[16] Ludwig et al. (2009) use the log of suicides rate per 100,000, but this implies having the population on both the left- and right-hand side of our regression. However, the results are very similar when we use the log rate instead of the log count.

inhabitants in Zurich to less than 10,000 in Appenzell-Innerhoden), we weight our estimates for the population. This weighting approach allows us to correct for heteroskedasticity in the error term (Solon et al., 2015), and alleviate the measurement error problem. Indeed, more populated regions show a higher signal-to-noise ratio, which is an issue, especially when dealing with a low-frequency outcome such as suicides.[17]

Antidepressant consumption is endogenous to the conditions that influence mental health outcomes. The inclusion of region fixed effects allows us to remove all time-invariant unobservables, but this does not allow us to get rid of all the endogeneity concerns. Several omitted time-varying factors may still bias our estimates. On the one hand, the latent (mental) health status of the population may affect both the use of antidepressants and the prevalence of mental health disorders in a region, causing an upward bias of our FE estimates. On the other hand, there have been attempts to build awareness of depression and decrease stigma in the general population and among health care practitioners.[18] We would expect such policy interventions to have a positive effect on antidepressant sales since they encourage the uptake of antidepressant treatment, and possibly a negative impact on health outcomes, introducing a downward bias in our estimates. To account for these issues, we rely on an instrumental variable strategy based on two different instruments, as following described.

4.1 IV 1: Pharmaceutical industry market power

The primary instrument for our identification strategy is an adaptation of the traditional shift-share instrument (Bartik, 1991), where national levels of antidepressant sales for each pharmaceutical company are assigned to small regions using the supply-driven increase in antidepressant sales. More specifically, we allocate the annual sales of antidepressant drugs of each pharmaceutical company using its regional market share for non-antidepressant drugs.

Our instrument diverges from the standard shift-share approach in two important ways. First, we use regional market shares, rather than regional sales relative to national sales, to calculate the regional shares. Indeed, the use of regional sales to calculate the regional shares would be endogenous since most pharmaceutical companies are likely to sell more in regions where the mental health conditions of the population are poor. The use of market shares allows us

[17] Comparing Zurich (a densely populated city) and Appenzell Innerhoden (a scarcely populated rural region), we observe a large variability in mental health outcomes (see Figure A.2).

[18] For instance, the "Alliance against Depression" is active in several (German-speaking) Swiss cantons and improves awareness for depression in the general population, and among physicians, teachers, etc. However, the scope, the length and the stakeholders of the program are up to the discretion of the cantons.

to overcome this problem since we consider the sales of each pharmaceutical company relative to its competitors. Second, the market shares are computed using non-antidepressant drugs sales. Therefore, we use a different market to exploit the market power of each pharmaceutical company in a region, and to overcome the residual concerns regarding the potential endogeneity of our shares. In practice, we exploit the fact that pharmaceutical companies tend to sell more in regions where they have higher market power (relative to their competitors).

The variation exploited by our instrument also comes from the differential (national) growth rate among pharmaceutical companies (net of the overall yearly change in antidepressant sales captured by the time fixed effects). As reported in Figure 2, the variation arises from the introduction of several new products in the market, some of which are covered by a new patent (brand name introduction) and some others introduced as new generic drugs (first or secondary introduction).

Because of data restrictions, our main estimates are obtained using 2002 as a base year to construct the shares, while data from 2003 to 2014 are used to analyze the relationship between antidepressant sales and mental health outcomes. Although our shift shares are different from those commonly employed in the literature, some concerns might still arise if there is no sufficient time gap between the base year and the years for which we estimate the effect of antidepressant sales on mental health outcomes. For this reason, we substantiate the validity our instrument by re-estimating the model using increasing time gaps between our base year (2002) and the years used for the estimation (using incrementally fewer data in steps of one year).[19]

More formally, our instrument is constructed as follows. Let the annual national stock of antidepressant for each pharmaceutical company be defined by $AD_{mt} = \sum_r AD_{mrt}$, where AD_{mrt} represents the antidepressant sales for pharmaceutical company m in region r at time t. This stock is used to calculate the shifts, i.e. the variation in antidepressant sales over time. The shares instead are calculated as:

$$\widetilde{S}_{mr2002} = \frac{v_{mr2002}}{\sum_m v_{mr2002}}. \tag{2}$$

In other words, $\widetilde{S}_{mr2002}$ represents the market share from wholesales of non-antidepressants for

[19] A time gap of one year implies that we are using data from 2004 to 2014; a time gap of two years implies that we are using data from 2005 to 2014, and so on.

pharmaceutical company m, in region r, in the base year 2002.[20] This allows us to redistribute the national stock AD_{mt} as follows:

$$\widetilde{AD}_{mrt} = \widetilde{S}_{mr2002} \times AD_{mt}, \tag{3}$$

where the regional variation in $\widetilde{AD}_{mrt}$ comes from differences in the pharmaceutical company's non-antidepressant market shares in the base year, and the temporal variation comes from the national growth rates of the company. Finally, we sum $\widetilde{AD}_{mrt}$ over m to obtain our instrument as follows:

$$\widetilde{AD}_{rt} = \sum_{m} \widetilde{AD}_{mrt}. \tag{4}$$

We can now exploit our instrument to estimate the model in (1) using a two-stage least squares fixed-effects estimator (2SLS-FE). The first stage can then be written as

$$AD_{rt} = \alpha_0 + \alpha_1 \widetilde{AD}_{rt} + \alpha_2 X_{rt} + \theta_r + \tau_t + \varepsilon_{rt}. \tag{5}$$

Since we interpret $\widetilde{AD}_{rt}$ as the sales due to market power, we expect the sign of α_1 to be positive if pharmaceutical companies are more successful in pushing their sales in regions where they have a larger market share as compared to other regions.

The estimate obtained from the first stage is then used in the second stage to recover the parameter of interest β_1 from equation (1). One should bear in mind that, in the presence of treatment heterogeneity, we do not estimate the average treatment effect, but a local average treatment effect (LATE) (Angrist et al., 1996). In particular, we measure the effect of antidepressant consumption on the health outcomes of interest for those people who consume antidepressants because of the pharmaceutical company's market power, but would not use antidepressants without market power.

It is worth reminding that antidepressants are exclusively prescribed by physicians, and their demand is unlikely to be driven by the patient who faces virtually no costs. Moreover, prices are set at the federal level, so pharmaceutical companies can only compete on quantity by physician detailing. As such, our instrument captures the potential influence that a company

[20] In a previous version of this paper, we use to normalize the shares by dividing them for the sum of the regional market shares of each company, $\sum_r \widetilde{S}_{mr2002}$. While this normalization allows to have shares ($\widetilde{S}_{mr2002}$) that sum up to one at company level, it leads the first stage to be slightly less powerful without affecting our point estimates.

can exert in a certain region with respect to its competitors. This means that α_1 captures how much this market power actually influences antidepressant sales. In the Appendix, we provide supporting evidence for this mechanism: we show that the local market share significantly increases in regions where pharmaceutical companies spend more in physician detailing (and other promotional activities towards physicians).

Rotemberg weights

We calculate the so-called "Rotemberg weights" following Goldsmith-Pinkham et al. (2020), who show that the Bartik-IV can be decomposed into a weighted combination of just-identified estimates, each of them using—in our specific case—the variation provided by a single pharmaceutical company as instrument.[21] The calculation of these weights make our identifying variation more transparent for two main reasons. First, it highlights which companies (among the 16 in our dataset) drive our overall estimate. The higher the weight of a company, the higher the influence on our estimate. Second, we can investigate whether these weights are driven by the shift component, the differential growth rate across companies over the observation window, or by the variation in pharmaceutical company (market) shares across regions.

The computation of Rotemberg weights also allows us to implement further tests about the validity of our research design. As already mentioned, we test for parallel pre-trends in the outcome variables for regions with high and low exposure (i.e., first versus fourth quartile of non-antidepressant market share in 2002) to the top-five weight companies, as in a standard DiD design. Furthermore, we investigate the level of heterogeneity resulting from these different instruments, one for each company ($\widetilde{AD}_{mrt}$), as reported in equation (3). Since we are investigating the effect of different drugs (often with different active ingredients) within the same class, we expect some level of heterogeneity. Moreover, we use an overidentified estimator based on these multiple Bartik instruments (i.e., 16 instruments, one for each company) to verify whether our results substantially change when we remove the instrument associated to the pharmaceutical company with the largest weight.

IV 2: Prescribing practice spillovers

The second instrument for our identification strategy relies on prescribing practice spillovers from outside Switzerland. This instrument is meant to take into account that around 30%

[21] We use the Stata package "Rotemberg Weight Package" provided by Paul Goldsmith-Pinkam on his personal website (`https://github.com/paulgp/bartik-weight`).

of doctors practicing in Switzerland have foreign qualifications, and almost all of them (25%) studied in one of the three big neighboring countries, namely France, Germany, and Italy. As stated above, we exclude Austria because it shares a relatively small border with Switzerland in a mostly mountainous area, and doctors from this country represent a negligible minority in Switzerland. In Figure A.3, we report the average DDD of antidepressant consumption per 1,000 inhabitants from 2002 to 2014 for the three main bordering countries. It is remarkable that both consumption levels and trends over time are very similar to those reported in Figure 1 for the different Swiss regions that are closer to these countries. Based on this observation, we build our instrument using spatially weighted averages of antidepressant sales in neighboring countries and assign them to small regions as follows:

$$\widetilde{AD}_{rt} = \frac{\sum_c w_{cr} AD_{ct}}{\sum_c w_{cr}}, \tag{6}$$

where $\widetilde{AD}_{rt}$ is a measure of antidepressant sales in region r at time t based on spillovers effects generated by exogenous prescribing practices, AD_{ct} is antidepressant sales (expressed in DDD per 1,000 inhabitants) in country c and year t, and w_{cr} is the squared inverse of geographical distance between country c and the centroid of region r. So, we assume that the magnitude of the spatial spillovers is inversely related to geographical distance (squared). This is reasonable considering the language barrier that might prevent foreign doctors to work further away in a different linguistic region.

Given the very high correlation in antidepressant use of the bordering regions with their neighboring countries, the relevance of this instrument does not represent a concern. Therefore, the validity of this instrument substantially relies on the assumption that antidepressant consumption in neighboring countries affects the mental health of the neighboring Swiss regions only through spatial spillovers in prescription practices (exclusion restriction). The inclusion in our specification of a full set of region fixed effects, that account for other time invariant cultural confounders, makes this assumption more reasonable. Although unlikely, we cannot exclude that other unobserved treatment practices correlated with changes in antidepressant prescriptions might violate the exclusion restriction. For this reason, we mainly rely on the identification based on the pharmaceutical industry market power (first instrument), using this second instrument essentially as a robustness check.

5. Results

In this section, we first show our main results on the effect of antidepressant sales on hospitalizations using the pharmaceutical industry market power instrument and the alternative instrument based on spatial spillovers. Then, we report the results of our investigation of the potential spillovers on labor market outcomes. Robustness checks are reported at the end of this section.

5.1 2SLS estimates using pharmaceutical industry market power

Table 2 shows the baseline results of our analysis using our main instrument, based on the pharmaceutical industry market power, for each health outcome. In particular, the table reports the estimated effect of antidepressant sales on all hospital admissions (column 1), hospital admissions related to mental disorder (column 2), hospital admissions related to depression symptoms (column 3), all hospital admissions excluding mental disorders (column 4), and suicides (column 5). Standard errors appear within brackets and are robust and clustered at region level. This is to say that standard errors are resilient to arbitrary within-panel autocorrelation (region) and to contemporaneous cross-panel correlation (year).

As a benchmark, we report the results obtained from the FE estimation in the first row of the table. The FE estimates indicate that a one-unit increase in defined daily doses of antidepressant per 1,000 inhabitants in a region (1 DDD corresponds to almost 3% of 2003 sales) is associated with approximately .1% more admissions, although the effect is not significant. This effect is mainly driven by hospital admissions for depression, for which the estimated coefficient is highly significant and suggests that hospital admissions may increase by 1.2%. Conversely, there is no evidence of a significant correlation between antidepressant sales and hospitalizations for mental disorders in general. Also, the estimated effect of antidepressant sales on suicides is not statistically significant and point estimates are even negative.

The second row of Table 2 shows the results of our 2SLS estimates, while the third and the fourth rows report the first stage and the reduced form results, respectively. When we exploit the arguably exogenous variation in antidepressant sales, the estimated effects are larger in magnitude than those estimated with the FE approach. The estimated effect for all hospital admissions is +.4%, although still not significant, and is driven by hospitalizations related to mental health disorders (+1.8%), in particular for depression (+5.6%). When we exclude hospital admissions

related to mental health from all admissions reported in column (1), we do not find evidence of large and significant effects on non-mental health related admissions. As for the estimated effect of antidepressant sales on suicides, we find no evidence of statistically significant effects but, differently from the FE results, point estimates are now positive. All these effects are similar both in magnitude and significance to those reported in Table A.6 of the Appendix for the log-log model specification. The only exception is that the effect of all admissions is now statistically significant.

The estimates of the first stage reported in the third row of Table 2 show that our instrument predicts antidepressant sales quite well. The Kleijbergen-Paap F-statistic on the excluded instrument is just above 15, but the Anderson-Rubin test allows us to reasonably reject the null hypothesis of weak instrument (and/or $\beta_1 = 0$) for mental disorders and depression. This is consistent with the results reported in the fourth row of Table 2, which provide evidence for the presence of significant reduced-form effects for these two outcomes. Moreover, our results are confirmed by the log-log specification where the value of the Kleijbergen-Paap F-statistic is almost twice as large as in the log-linear specification (Table A.6).

In the Appendix, we also present the estimation results obtained when we exploit the heterogeneity across gender and age groups (Table A.7). In the case of hospitalizations for mental disorders and depression, we find some evidence of heterogeneity. The estimated effect is generally lower for the elderly, especially for depression. However, larger heterogeneity is found for suicides. In particular, we estimate a slightly significant increase in suicides for women, but we do not find similar evidence of heterogeneity for the other mental health outcomes. Given the high risk of a "false positive" when breaking down estimates for such a rare event like suicides, we avoid speculating on these differences. Consider also that antidepressant sales and the instrument do not vary by gender and age, while our outcomes do.

In what follows, we open the black box of our Bartik-type instrument. We start from Table 3, which shows summary statistics about the Rotemberg weights collapsed at pharmaceutical company level (as in Goldsmith-Pinkham et al., 2020). Although Panel A shows the presence of negative weights on some pharmaceutical companies, these weights are on average very small. If ever, they reduce the estimated effect of antidepressant on hospitalizations for depression (see the values in the "α-weighted β" column). The correlations between pharmaceutical company aggregates reported in Panel B indicate that the weights (α_m) are largely associated with national

growth rates (g_m) and, to a lower extent, with the variation in shares across regions (Var(z_m)) and the power of the first-stage F statistic (F_m). This suggests that the variation exploited by our estimator derives mostly (although not exclusively) from differences in product innovation across pharmaceutical companies in the observation window, rather than the variation in market shares across regions. This is reflected also in Panel C, where the top-three weight pharmaceutical companies account for 85% of positive weights,[22] despite they are characterized by a quite low market share (last column of Panel C). The top weight company, Mepha-Teva, explains alone 67% of positive weights. Indeed, Mepha-Teva experienced an extraordinary growth (the value of g_m in the second column), thanks to the introduction of several new generic drugs since 2004 (see also Table A.4). Considering that our estimates are largely driven by Mepha-Teva, we test the validity of our research design—as in a difference-in-difference (DiD)— by comparing unconditional trends in average hospital admissions for depression in regions with high and low market shares of this company. This is illustrated in Figure 3, where we use the top and the bottom quartile of non-antidepressant regional market shares in 2002 as a sort of measure of treatment exposure. In this specific case, the treatment is represented by the introduction of new generics in the Swiss market since 2004. It is reassuring to observe that the trends in hospital admissions for depression in the two groups of regions between 1999 and 2003 are parallel. Conversely, the two trends clearly diverge starting from 2004, with high-share regions experiencing a larger increase in hospital admissions.[23]

Since the "beginning of treatment" for the other four companies is not as clear as for the top weight company (these pharmaceutical companies started to introduce new drugs at very different points in time), we only show their pre-trends (before 2003) in the main outcome in Figure A.4. Except for one company (Sandoz) —where there is a potential (small) anticipation effect one year before— trends appear parallel between high and low share regions between 1999 and 2002. In the Appendix (Table A.8), we test more formally for pre-trends by regressing the growth in hospitalizations for mental disorders and depression in the period 1999 and 2002 on regional market shares of the top-five weight companies used to construct the instrument.

Although most of the variation is driven by one company, our estimate can be generalized also to antidepressants sold by other companies. In the Appendix (Figure A.5), we show that the

[22] Note that total positive weights are equal to 1.179, as reported in Panel A.

[23] In the figure, regions with a larger exposure to this company are characterized by a lower number of hospital admissions in the pre-treatment period. Part of this initial difference in levels is due to population size. However, differences in trends are similar if we account for the population.

point estimates associated with the other major companies are all positive with some evidence of heterogeneity. This should not be surprising since these companies often sell different products that might have different (side-) effects. For this reason, we also verify that our results are not affected by the exclusion of the top-weight company (Mepha-Teva) when using an overidentified estimator based on these multiple Bartik instruments, one for each company (more details are provided in Section 5.4).

We also report evidence regarding the influence of pharmaceutical companies over doctors' prescribing practices through physician detailing (Table A.5). This helps to explain the local market power that is exploited in our identification strategy. More precisely, we show that the regional expenditure in detailing activities by pharmaceutical companies is strongly correlated with revenues and market shares. This correlation is robust even if we account for region fixed effects or company and canton fixed effects.

5.2 2SLS estimates using spatial spillovers

In Table 4 we replicate our main analysis using the alternative instrument based on spatial spillovers in prescription practices. It is remarkable that our estimates based on this alternative instrument lead to very similar point estimates. For instance, the estimated effects are slightly larger (and statistically significant) for all hospital admissions (column 1), and almost identical for mental health (column 2) and depression (column 3), but standard errors largely overlap with the estimates reported in Table 2. It is also worth noting that this instrument is particularly powerful and should reject any residual weak instrument concerns. Since point estimates are almost identical, it is pointless to show the results of the overidentification test using the two instruments jointly.

5.3 Effects on labor market outcomes

To add more evidence, we investigate the potential spillovers of antidepressant sales on labor market outcomes using data from the Swiss Labour Force Survey (Table 5). Fixed effects estimates suggest the presence of small and positive spillovers on the labor market. An increase of one DDD in antidepressant sales is associated with an increase of 0.1 percentage points in labor market participation. Conversely, the effect on income and working hours, conditional on being in the labor force, is negative. Such estimates might imply a small increase in the share of people only marginally included in the labor force (which reduces the average income and working hours conditional on participation). However, the result is not confirmed by our IV

estimates. Except for income, IV estimates point in the same direction, but the coefficients are not statistically different from zero. All in all, the standard errors are small enough to discard the presence of economically relevant effects.

Finally, it is worth noting that even our estimates based on the FE model do not clearly show the presence of a negative correlation between antidepressant sales and labor market outcomes. This result should reassure on the presence of any residual time-varying unobservable factor that could bias our IV estimates, and can possibly explain the estimated effect on hospitalizations for mental health problems.

5.4 Robustness checks

In our baseline regression, we estimate the treatment effect for the period 2003-2014 using 2002 as the base year. As a robustness check, we keep the base year at 2002 and, in succeeding steps of one year, we use incrementally fewer data to estimate the treatment effect. In practice, we increase the time gap between the base year used to construct the shares and our observation window. In such a way, we effectively decrease the sample size and, therefore, the variation that we can exploit. Nonetheless, we find consistent point estimates and standard errors for all the three outcomes (Figure A.1). If ever, the Figure shows larger point estimates, especially for mental health disorders, when we use increasingly more recent data.

In Table A.9 we report 2SLS estimates of the effect of increasing antidepressant sales on elective hospital admissions and a set of other placebo outcomes, namely infectious diseases, bone fractures and pregnancy. These placebo outcomes are selected to match the mean and standard deviation of our primary outcomes. Reassuringly, none of the placebo outcomes show a significant (and relevant) correlation with the instrumented antidepressant sales.

Our results are also robust when we consider the count nature of the health outcomes (before the log transformation). In this case, we take the endogeneity of antidepressant sales into account in a non-linear model using the so-called "two-stage residual inclusion" (Terza et al., 2008), which is a two-step procedure similar to the control function approach where the residuals from the first stage are included in the main regression equation. The point estimates reported in Table A.10 are very similar to those of Table 2 and 4, although larger for mental disorders using the market share instrument (IV 1).

As already mentioned, we formally test for pre-trends (before 2003) in the main outcomes

of interest (hospital admissions for mental disorders, depression and suicide) for the top-five Rotemberg weight companies (Table A.8). For each mental health outcome, columns (1), (3) and (5) show that there is no evidence of significant correlation between changes in hospital admissions before 2003 and the non-antidepressant market share of these five companies in 2002, $\widetilde{S}_{mr2002}$ in equation (4.1). In addition, in column (2), (4) and (6), we show that pre-trends in the three outcomes of interest are not significantly correlated with the change of our instrument over the observation window (2003-2014).

Finally, in Table A.11, we exploit the overidentified estimator based on the multiple Bartik instruments, one for each company, to investigate the heterogeneity in the effect of interest and the exclusion of the top-weight company. Given the bias of the overidentified 2SLS in finite samples, we use the Limited Information Maximum Likelihood (LIML) that has better sample properties than the 2SLS (Chao and Swanson, 2005). In Panel A of Table A.11 we show the results using all the 16 company-instruments. Then, in Panel B we use only the Mepha-Teva instrument, while in Panel C we exclude only Mepha-Teva from the instruments set. The point estimates for the LIML estimator reported in Panel A are smaller than our Bartik instrument estimates reported in the main text, but still positive and statistically significant for depression. While Mepha-Teva estimates are larger in size (Panel B), the exclusion of this company from the instrument list (Panel C) only slightly reduces the estimated effect, which remains statistically significant for depression.

6. Conclusion

This research sheds light on the health effects of the dramatic increase in antidepressant consumption observed in Switzerland in the last two decades – a phenomenon that is also present in many other developed countries. Using plausibly exogenous variations in local market shares of pharmaceutical companies and product innovation, we find that antidepressant sales increase hospital admissions for mental disorders, especially for depression symptoms. The estimated effects on suicides are not statistically significant but points in the same direction. We show that our identifying variation is largely (although not exclusively) driven by one pharmaceutical company that introduced several new generic products over the observation window. Yet, the exclusion of this company does not undermine our results. We find almost identical results when we use a second instrument, which exploits spatial spillovers in antidepressant prescribing practices from the main neighboring countries of Switzerland.

Our results are in contrast with earlier evidence by Ludwig et al. (2009), who find that the increase in the use of SSRI antidepressants decreases suicide mortality by 5%. These authors compare the use of SSRI across countries and over time in the 80s and the 90s. At that time, SSRI were promoted as being more efficient than TCA antidepressants, particularly in terms of reduced side effects. Since this analysis includes the introduction period of SSRI, the study probably captures the initial impact of their uptake. The current market, however, is dominated by antidepressant drugs in the SSRI class.

As it is often the case in IV settings, our estimates allow to recover only the effect for the subpopulation of compliers (LATE), which may not coincide with the average effect for the whole population (ATE). However, the fact that we find very similar results using two instruments based on very different research designs might help to generalize our findings, at least in the light of current levels of treatment with antidepressants. Evidence suggests that, over time, the prescription threshold for antidepressants has shifted towards the lower end of the severity distribution of depression, despite prescription guidelines dictate psychotherapy for mild depression and, at most, a combination of psychotherapy and pharmacotherapy for moderate cases. Our results imply that the marginal patient treated with antidepressants nowadays may no longer benefit from antidepressant treatment. Therefore, a policy recommendation would be to advance measures ensuring adherence to prescription guidelines, and to emphasize the importance of alternatives to pharmacotherapy, especially for mild and moderate depression.

In accordance with previous evidence on extensive off-label prescription practices, our research sheds light on one of the causes of over-treatment with pharmacotherapy, namely the influence of pharmaceutical companies over doctors' prescribing practices. However, there are other potential causes of over-treatment that are not addressed in this paper. First, over-treatment with antidepressants might be the result of undercapacity of psychotherapists. Second, psychotherapy is a form of treatment more time-consuming than prescribing antidepressants. Hence, decreasing stigma and increasing awareness may have led the number of patients to grow to such an extent that physicians have resorted to pharmacotherapy, even if this therapy is not the best treatment option. Moreover, patient bias in beliefs about treatment effects (Cronin et al., 2020) might also explain why, despite clear guidelines and increasing evidence on the benefits of psychotherapy, people may still prefer the drug treatment. Further research in needed to pin down the mechanisms behind the extensive use of antidepressants with little efficacy, and potentially inducing

adverse health effects and increasing health care costs.

References

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Bartik, T. J. (1991). Who benefits from state and local economic development policies? Technical report, Upjohn Institute for Employment Research.

Biller-Andorno, N. and Zeltner, T. (2015). Individual responsibility and community solidarity – The Swiss health care system. *New England Journal of Medicine*, 373(23):2193–2197.

Borusyak, K., Hull, P., and Jaravel, X. (2018). Quasi-experimental shift-share research designs. Technical report, National Bureau of Economic Research.

Bridge, J. A., Iyengar, S., Salary, C. B., Barbe, R. P., Birmaher, B., Pincus, H. A., Ren, L., and Brent, D. A. (2007). Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. *Jama*, 297(15):1683–1696.

Busch, S. H., Golberstein, E., and Meara, E. (2014). The FDA and ABCs Unintended Consequences of Antidepressant Warnings on Human Capital. *Journal of Human Resources*, 49(3):540–571.

Bütikofer, A., Cronin, C. J., and Skira, M. M. (2020). Employment effects of healthcare policy: Evidence from the 2007 fda black box warning on antidepressants. *Journal of Health Economics*, 73:102348.

Chao, J. C. and Swanson, N. R. (2005). Consistent estimation with a large number of weak instruments. *Econometrica*, 73(5):1673–1692.

Cipriani, A., Zhou, X., Del Giovane, C., Hetrick, S. E., Qin, B., Whittington, C., Coghill, D., Zhang, Y., Hazell, P., Leucht, S., et al. (2016). Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: A network meta-analysis. *Lancet*, 388(10047):881–890.

Coupland, C., Dhiman, P., Morriss, R., Arthur, A., Barton, G., and Hippisley-Cox, J. (2011). Antidepressant use and risk of adverse outcomes in older people: Population based cohort study. *British Medical Journal*, 343:d4551.

Cronin, C. J., Forsstrom, M. P., and Papageorge, N. W. (2020). What good are treatment effects without treatment? mental health and the reluctance to use talk therapy. Technical report, National Bureau of Economic Research.

Cuddy, E. and Currie, J. (2020a). Rules vs. discretion: Treatment of mental illness in us adolescents. Technical report, National Bureau of Economic Research.

Cuddy, E. and Currie, J. (2020b). Treatment of mental illness in american adolescents varies widely within and across areas. *Proceedings of the National Academy of Sciences*, 117(39):24039–24046.

Currie, J. M. and MacLeod, W. B. (2020). Understanding physician decision making: The case of depression. *Econometrica*, 88:847–878.

Davies, J. and Read, J. (2019). A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: Are guidelines evidence-based? *Addictive behaviors*, 97:111–121.

Editors, P. M. et al. (2013). The paradox of mental health: over-treatment and under-recognition. *PLoS Med*, 10(5):e1001456.

Filippini, M., Masiero, G., and Steinbach, S. (2019). The impact of ambient air pollution on hospital admissions. *The European Journal of Health Economics*, 20(6):919–931.

Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., and Fawcett, J. (2010). Antidepressant drug effects and depression severity: A patient-level meta-analysis. *Journal of the American Medical Association (JAMA)*, 303(1):47–53.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110:2586–2624.

Haller, E., Watzke, B., Blozik, E., Rosemann, T., Reich, O., Huber, C. A., and Wolf, M. (2019). Antidepressant prescription practice and related factors in switzerland: a cross-sectional analysis of health claims data. *BMC psychiatry*, 19(1):196.

Herriot, J. G. and Reinsch, C. H. (1973). Algorithm 472: procedures for natural spline interpolation. *Communications of the Association for Computing Machinery*, 16:763–768.

Jakobsen, J. C., Katakam, K. K., Schou, A., Hellmuth, S. G., Stallknecht, S. E., Leth-Møller, K., Iversen, M., Banke, M. B., Petersen, I. J., Klingenberg, S. L., et al. (2017). Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and trial sequential analysis. *BMC Psychiatry*, 17(1):58.

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., et al. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1789–1858.

Ludwig, J., Marcotte, D. E., and Norberg, K. (2009). Anti-depressants and suicide. *Journal of Health Economics*, 28(3):659–676.

Masiero, G., Mazzonna, F., and Verbeek, O. (2018). What drives the rise of antidepressant consumption? Evidence from Switzerland. Technical Report 2018/1, IdEP Economic Papers, Università della Svizzera italiana.

Mojtabai, R. and Olfson, M. (2011). Proportion of antidepressants prescribed without a psychiatric diagnosis is growing. *Health Affairs*, 30(8):1434–1442.

Moore, M., Yuen, H. M., Dunn, N., Mullee, M. A., Maskell, J., and Kendrick, T. (2009). Explaining the rise in antidepressant prescribing: A descriptive study using the general practice research database. *British Medical Journal*, 339:b3999.

OECD (2019). Health at a glance 2019. page 220.

Pratt, L. A., Brody, D. J., and Gu, Q. (2017). Antidepressant use among persons aged 12 and over: United states, 2011-2014. *NCHS data brief*, (283):1.

Priebe, S., Frottier, P., Gaddini, A., Kilian, R., Lauber, C., Martínez-Leal, R., Munk-Jørgensen, P., Walsh, D., Wiersma, D., and Wright, D. (2008). Mental health care institutions in nine European countries, 2002 to 2006. *Psychiatric Services*, 59(5):570–573.

Santésuisse (2017). Press release, 4 June 2017. *https://www.santesuisse.ch/fileadmin/-sas_content/Comm_2017-04-06-d.pdf*.

Schuler, D. and Burla, L. (2012). *Psychische Gesundheit in der Schweiz: Monitoring 2012*. Schweizerisches Gesundheitsobservatorium (Obsan).

Shapiro, B. (2020). Promoting wellness or waste? evidence from antidepressant advertising. *American Economic Journal: microeconomics*, (Forthcoming).

Shapiro, B. T. (2018). Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy*, 126(1):381–437.

Sinkinson, M. and Starc, A. (2019). Ask your doctor? direct-to-consumer advertising of pharmaceuticals. *The Review of Economic Studies*, 86(2):836–881.

Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.

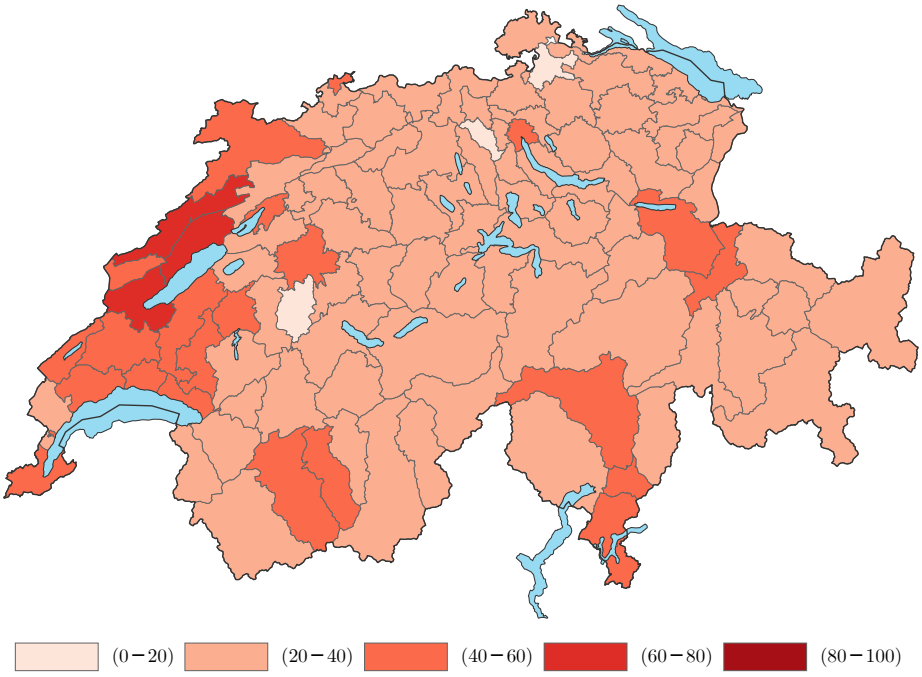
Stewart, W. F., Ricci, J. A., Chee, E., Hahn, S. R., and Morganstein, D. (2003). Cost of lost productive work time among us workers with depression. *Journal of the American Medical Association (JAMA)*, 289(23):3135–3144.

Stone, M., Laughren, T., Jones, M. L., Levenson, M., Holland, P. C., Hughes, A., Hammad, T. A., Temple, R., and Rochester, G. (2009). Risk of suicidality in clinical trials of antidepressants in adults: Analysis of proprietary data submitted to US Food and Drug Administration. *British Medical Journal*, 339:b2880.

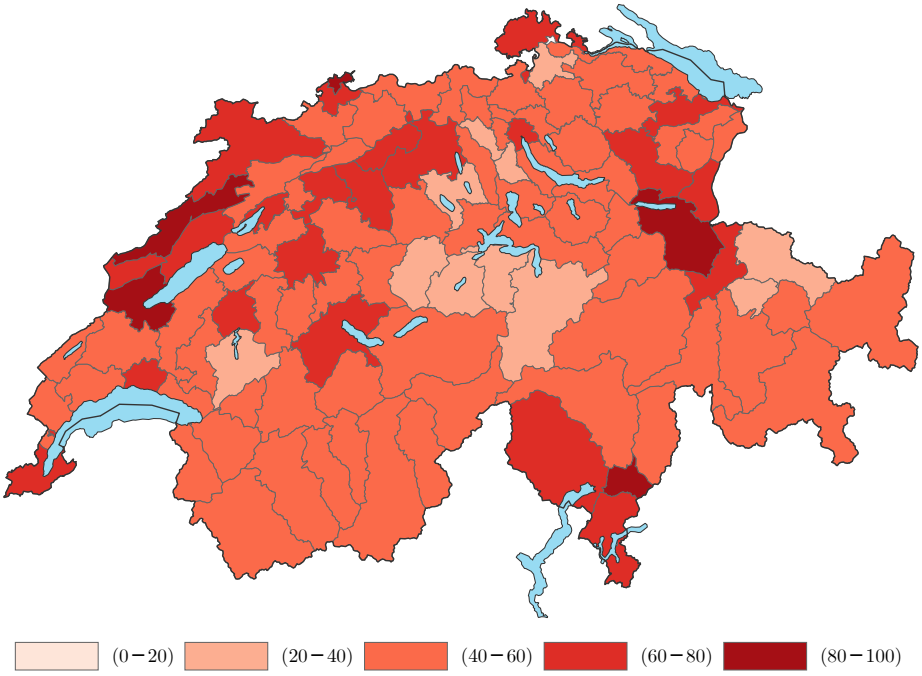
Terza, J. V., Basu, A., and Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, 27(3):531–543.

Wong, J., Motulsky, A., Abrahamowicz, M., Eguale, T., Buckeridge, D. L., and Tamblyn, R. (2017). Off-label indications for antidepressants in primary care: descriptive study of prescriptions from an indication based electronic prescribing system. *British Medical Journal*, 356:j603.

Figures and Tables



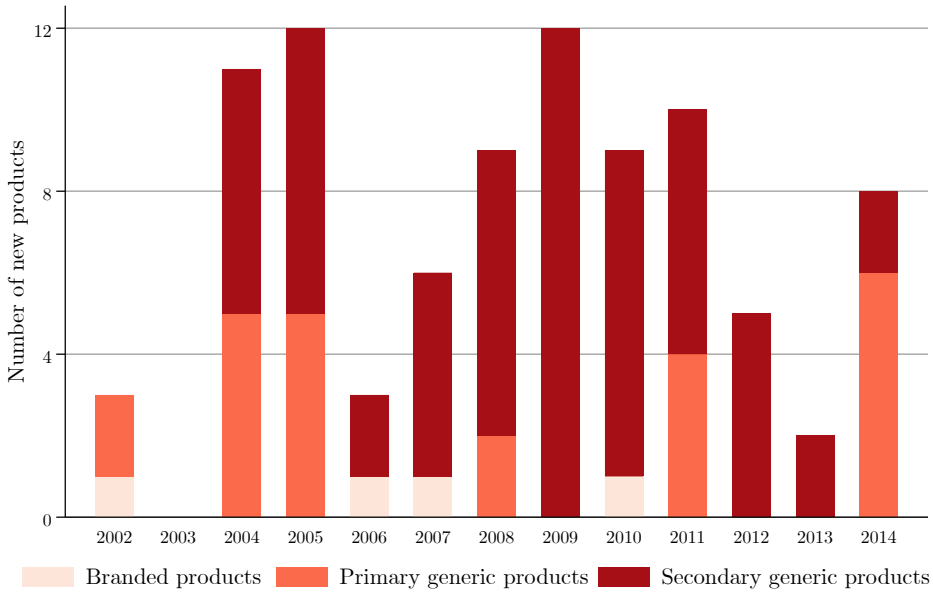
(a) Antidepressant sales in 2003



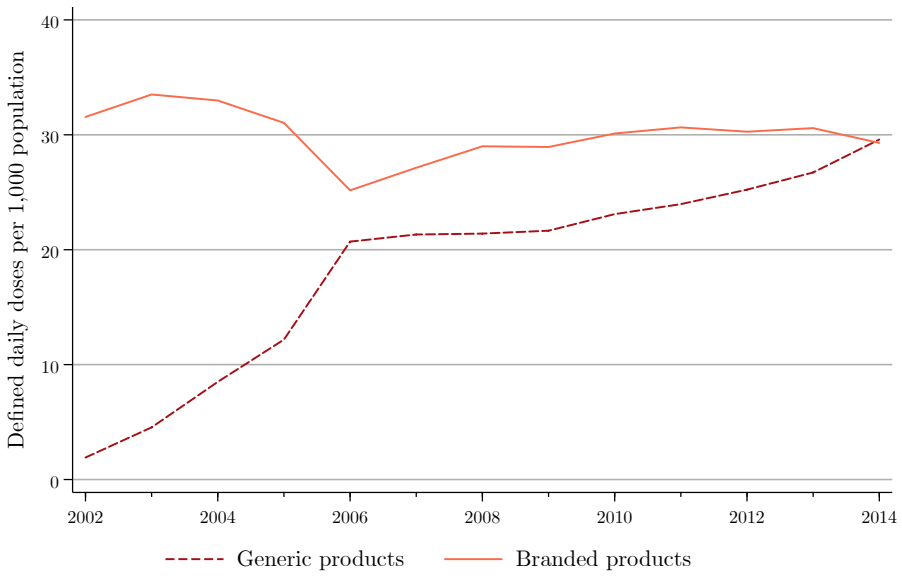
(b) Antidepressant sales in 2014

Figure 1: Antidepressant sales in Switzerland by small regions in 2003 and 2014

Notes – The figure examines antidepressant sales in Switzerland at the small region level for 2003 and 2014. To compare drug sales across regions, we classify the annual consumption according to five classes ranging from low to high where darker shades stand for higher consumption levels.



(a) Introduction of new anti-depression drugs in Switzerland



(b) Generic and branded antidepressant sales (DDD per capita)

Figure 2: Introduction of new antidepressant drugs in Switzerland from 2002 to 2014

Notes – The figure describes the antidepressant market in Switzerland over the period 2002–2014. Subfigure (a) shows the introduction of new antidepressant drugs in Switzerland. Light-red bars indicate brand name products, red bars primary generic products, and dark-red bars secondary generic products, respectively. Subfigure (b) shows antidepressants sales in DDD (Defined Daily Doses) per capita, distinguishing between branded and generic drugs.

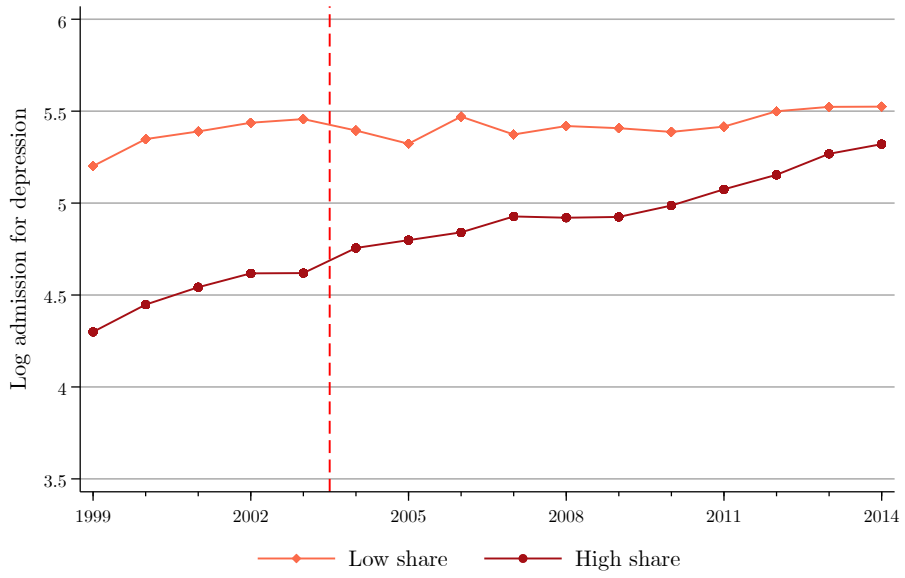


Figure 3: Trends in hospital admissions for depression in small regions with high and low Mepha-Teva market shares

Notes – The figure compares the average of log hospital admissions for depression in regions at the top quartile (high share) and the bottom quartile (low share) of the 2002 regional market share of Mepha Teva. The vertical red line between 2003 and 2004 marks the year of introduction of several new generic antidepressant drugs by Mepha-Teva in the Swiss Market.

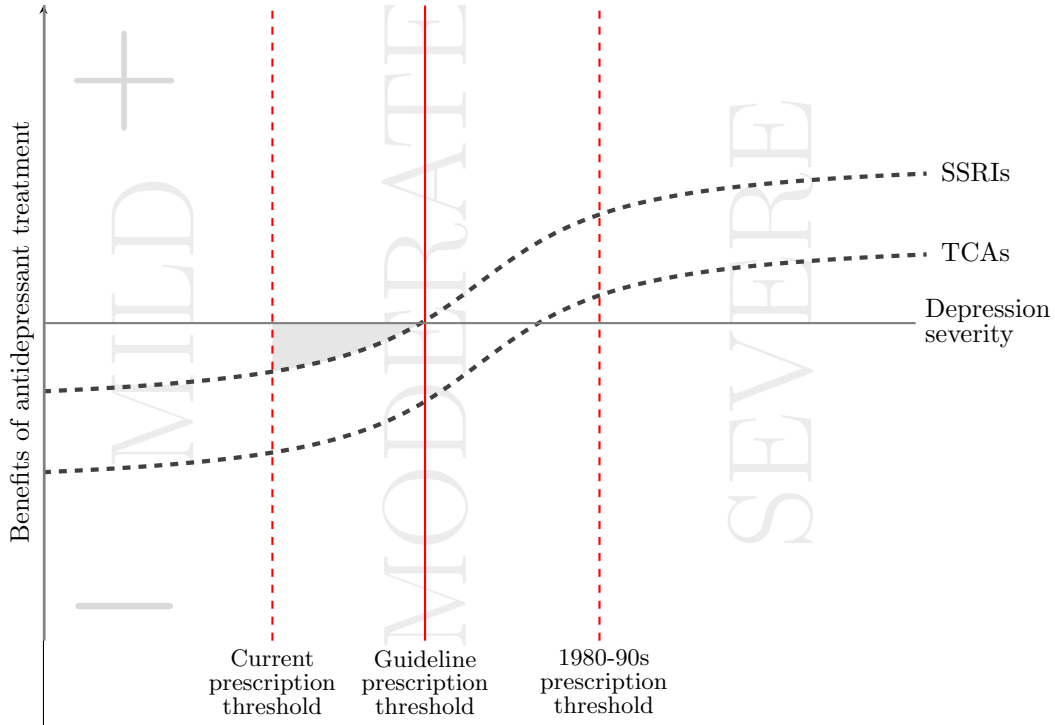
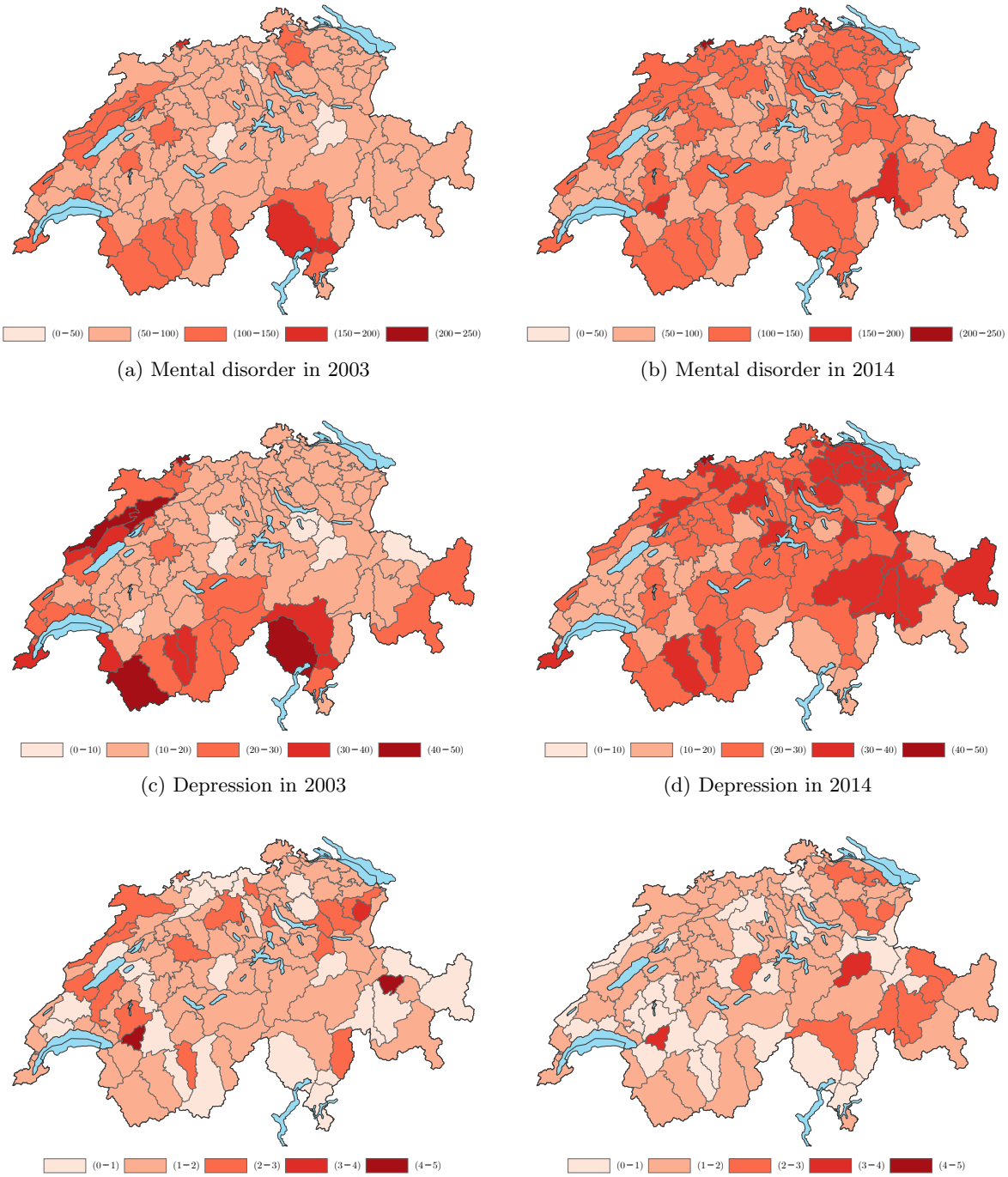


Figure 4: Conceptual framework for antidepressant benefits and prescription thresholds

Notes – The figure describes the potential relationship between benefits of antidepressant treatment and depression severity for old and new classes of antidepressants (i.e., TCA and SSRI, respectively). Vertical lines represent different prescription thresholds.



(a) Mental disorder in 2003

(b) Mental disorder in 2014

(c) Depression in 2003

(d) Depression in 2014

(e) Suicide in 2003

(f) Suicide in 2014

Figure 5: Mental health outcomes in Switzerland by small regions in 2003 and 2014

Notes – The figure classifies the prevalence of mental health disorders, depression, and suicide for 2003 and 2014. The health outcomes are categorized according to five classes ranging from low to high where darker shades stand for higher incidence.

Table 1: Descriptive statistics

		Standard deviation					
	Mean	Overall	Between	Within	Δ(2003/14)	Min	Max
Antidepressant sales	48.86	14.13	10.99	7.84	20.03	15.48	101.38
Mental disorders	102.98	29.62	22.73	13.65	16.54	22.16	219.08
Depression	22.09	7.66	5.02	5.35	6.12	2.89	56.07
Suicides	1.36	0.49	0.30	0.65	−0.19	0.00	6.68
Female share	50.83	0.87	0.81	0.26	−0.50	47.92	52.94
Below 15 share	15.57	1.68	1.56	0.97	−2.32	10.03	21.29
Between 15-65 share	67.69	1.54	1.68	0.56	−0.11	60.34	71.58
Foreigner share	22.04	8.01	7.03	1.49	3.69	3.65	40.95
German speaking share	63.19	36.80	37.42	0.01	−0.00	1.58	96.77
Unemployment rate (SLFS)	2.50	1.60	1.08	1.41	−0.89	0.00	10.98
Labor market participation (SLFS)	86.66	3.09	2.96	2.72	2.79	63.90	100.00
Annual total income (SLFS)	10.98	0.14	0.10	0.12	0.09	10.44	13.14
Annual working hours (SLFS)	9.36	1.07	0.85	0.55	0.52	6.63	11.36

Notes – The table reports descriptive statistics for the main variables. The statistics are obtained using annual data at the small region level for the period from 2003 to 2014. Antidepressant use is measured in terms of defined daily doses per 1,000 inhabitants per day. Hospital admissions for mental disorders and depression, and suicide attempts are expressed in terms of cases per 10,000 inhabitants. Specialists and general practitioners are measured by the density per 10,000 inhabitants. The annual total income (SFLS) and annual working hours (SLFS) are log transformed.

Table 2: Estimates of the effect of antidepressant sales on hospital admissions and suicides

	All admissions	Mental disorder	Depression	Admissions no mental	Suicide
	(1)	(2)	(3)	(4)	(5)
Fixed effects	0.001	0.002	0.012***	0.001	-0.003
	(0.001)	(0.002)	(0.005)	(0.001)	(0.004)
Second stage	0.004	0.018***	0.056***	0.003	0.010
	(0.003)	(0.007)	(0.019)	(0.003)	(0.011)
First stage	0.131***	0.131***	0.131***	0.131***	0.131***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Reduced form	0.001	0.002**	0.007***	0.000	0.001
	(0.000)	(0.001)	(0.002)	(0.000)	(0.001)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	15.637	15.637	15.637	15.637	15.637
Anderson-Rubin F (p-value)	0.153	0.026	0.000	0.274	0.363

Notes – For each health outcome, the table reports the parameter estimates for the fixed effect model (FE) and the instrumental variable model (2SLS) using the shift-share pharmaceutical market instrument. Column (1) shows the estimated effects on all hospital admissions, column (2) on hospital admissions related to mental disorders, column (3) on hospital admissions related to depression symptoms, column (4) on all hospital admissions excluding mental disorders, and column (5) on suicides (completed and attempted). All models control for year and region fixed effects and population characteristics, as reported in Table 1. We also include the corresponding first stage and the reduced form coefficients from the 2SLS model. We use cluster-robust standard errors at the region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table 3: Summary of Rotemberg weights

Panel A: Negative and positive weights

	Sum	Mean	Share	α-weighted β
Negative	-0.179	-0.030	0.132	-0.018
Positive	1.179	0.118	0.868	0.074

Panel B: Correlations of pharmaceutical company aggregates

	α_m	g_m	β_m	F_m	Var(z_m)
α_m	1				
g_m	0.528	1			
β_m	-0.023	0.263	1		
F_m	0.226	-0.127	-0.198	1	
Var(z_m)	0.255	0.117	0.356	0.225	1

Panel C: Top 5 Rotemberg weights by pharmaceutical company

	$\hat{\alpha}_m$	g_m	$\hat{\beta}_m$	95% CI	Market Share
Mepha-Teva	0.808	15.936	0.068	(0.035, 0.165)	4.107
Sandoz	0.096	12.541	0.072	(0.030, 0.255)	1.861
Vifor Pharma	0.096	3.612	0.007	(-0.035, 0.025)	4.546
MSD	0.095	5.541	0.064	(0.010, 0.315)	22.126
Sanofi-Aventis	0.053	2.036	0.081	(0.055, 0.130)	11.660

Notes – This table reports statistics about the Rotemberg weights aggregated at pharmaceutical company level (indexed by m) across years. Panel A shows the share and the sum of positive and negative weights. Panel B shows the correlation between the Rotemberg weights (α_m), the company national growth rate (g_m), the coefficient estimates for the effect of antidepressant sales on hospitalizations for depression (β_m), the first-stage F-statistic of the pharmaceutical company share (F_m), and the variation in the company shares across locations (Var(z_m)). Panel C shows the top-five pharmaceutical companies according to Rotemberg weights.

Table 4: Estimates of the effect of antidepressant sales on mental health outcomes using the practice spillover instrument

	All admissions	Mental disorder	Depression	Admissions no mental	Suicide
	(1)	(2)	(3)	(4)	(5)
Fixed effects	0.001	0.002	0.012***	0.001	-0.003
	(0.001)	(0.002)	(0.005)	(0.001)	(0.004)
Second stage	0.008**	0.018***	0.058***	0.008**	0.003
	(0.003)	(0.006)	(0.017)	(0.003)	(0.007)
First stage	0.413***	0.413***	0.413***	0.413***	0.413***
	(0.066)	(0.066)	(0.066)	(0.066)	(0.066)
Reduced form	0.003***	0.007***	0.024***	0.003***	0.001
	(0.001)	(0.002)	(0.004)	(0.001)	(0.003)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	38.816	38.816	38.816	38.816	38.816
Anderson-Rubin F (p-value)	0.002	0.001	0.000	0.003	0.718

Notes – The table reports the parameter estimates of the instrumental variable model (2SLS) using the practice spillover instrument. Column (1) shows the estimated effects on all hospital admissions, column (2) on hospital admissions related to mental disorders, column (3) on hospital admissions related to depression symptoms, column (4) on all hospital admissions excluding mental disorders, and column (5) on suicides (completed and attempted). We also include the corresponding first stage and reduced form coefficients from the 2SLS model. All models control for year and region fixed effects and population characteristics, as reported in Table 1. We use cluster-robust standard errors at the region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table 5: Estimates of the effect of antidepressant sales on labor market outcomes

	Unemployment	Labor force	ln (mean income)	ln (mean hours)
Fixed effects	-0.000	0.001***	-0.003***	-0.001***
	(0.000)	(0.000)	(0.001)	(0.000)
Second stage (IV-1)	-0.000	0.001	0.005	-0.001
	(0.001)	(0.001)	(0.003)	(0.001)
Kleijbergen-Paap F	14.261	14.261	14.261	14.261
Anderson-Rubin F (p-value)	0.505	0.428	0.088	0.256
Second stage (IV-2)	-0.000	0.001	0.001	-0.001
	(0.000)	(0.001)	(0.003)	(0.001)
Kleijbergen-Paap F	38.244	38.244	38.244	38.244
Anderson-Rubin F (p-value)	0.455	0.419	0.790	0.145
Observations	1,142	1,142	1,142	1,142

Notes – The table reports the parameter estimates of the effect of antidepressant sales on labor force outcomes using a fixed effects regression model (FE), the instrumental variable regression model based on the pharmaceutical industry market power (IV 1), and the instrumental variable regression model based on spatial spillovers (IV 2). Data are drawn from the Swiss Labor Force Survey and aggregated at the level of spatial mobility region. We use cluster-robust standard errors at the region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Happy pills? Mental Health Effects of the Dramatic Increase in Antidepressant Use

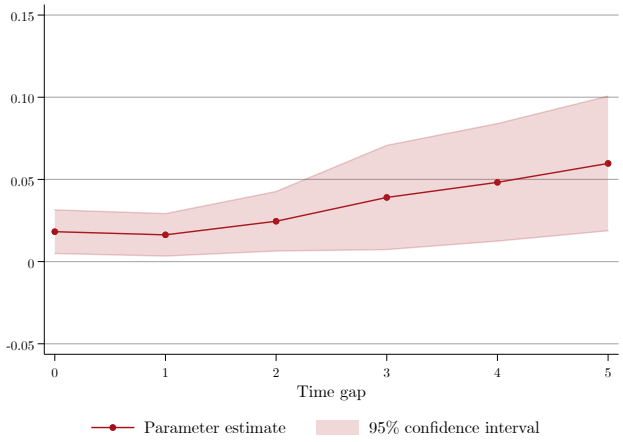
Online Appendix

Giuliano Masiero, Fabirizio Mazzona and Sandro Steinbach

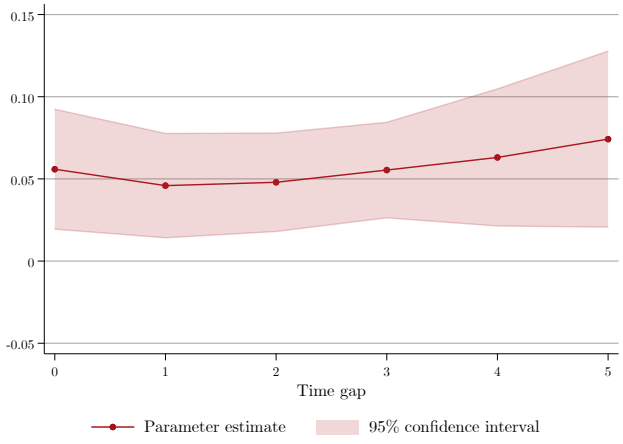
Data construction

We rely on the small region level (SMR - spatial mobility region) as the primary data aggregation level. There are 106 SMR in Switzerland, with each of them accounting for approximately 45,000 individuals. The SMR is a statistical subdivision of Switzerland based on economic activity around an agglomeration hub. Because the SMRs are based on the Swiss municipalities, we can aggregate municipality-level data at the SMR level. The antidepressant wholesale data are published at the pharmaceutical sales region (PSR) level for 2002 to 2014. There are 237 PSR regions in Switzerland that represent an aggregation of the postal codes. We use a Geographic Information System (GIS) to match postal codes to the SMRs. We obtained detailed information on the general population at the postal code level from the FSO. The approach was first suggested by Filippini et al. (2019). We use the population information to create spatial weights to recode the location information and obtain a match between PSR and SMR. We then reassign the antidepressant consumption data to the SMR level using population weights.

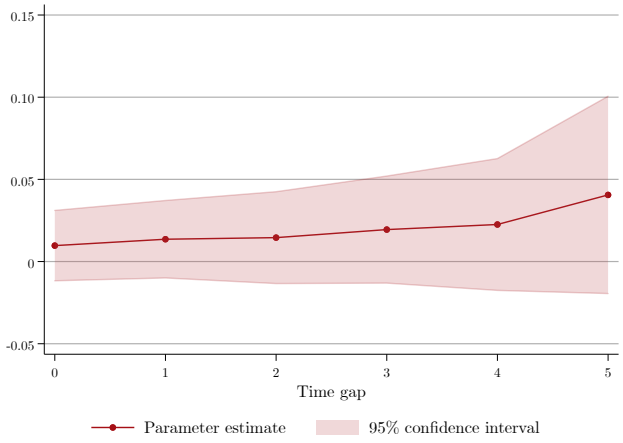
Figures and Tables



(a) Mental disorders



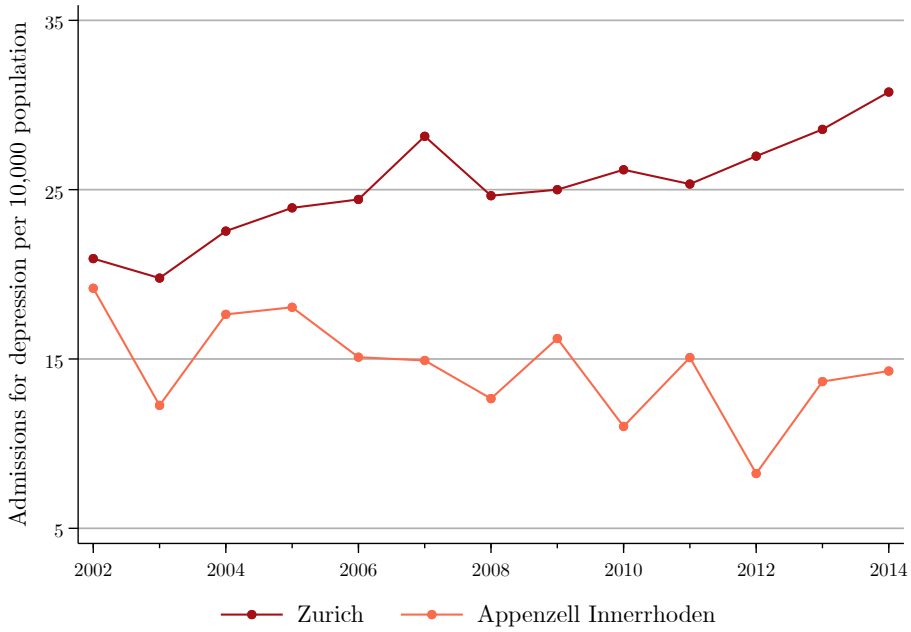
(b) Depression



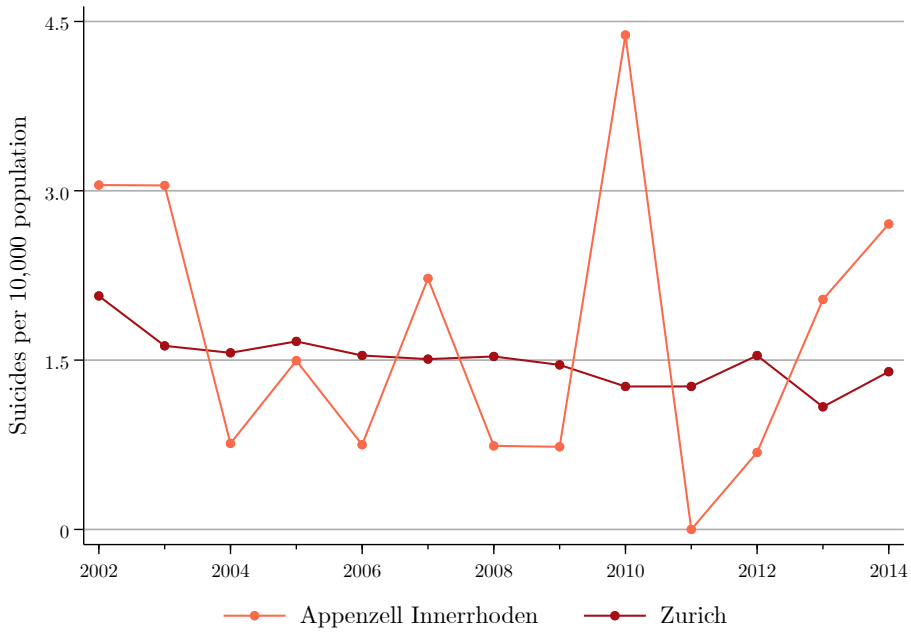
(c) Suicide

Figure A.1: Parameter estimates by health outcome relative to 2002

Notes – The figure shows IV estimates for mental disorders (a), depression (b), and suicide (c) when we increase the time gap used for the estimation relative to the base year 2002. We report the parameter estimates and the corresponding 95% confidence intervals.



(a) Depression



(b) Suicide

Figure A.2: Variability of mental health outcomes for two small regions between 2002 and 2014

Notes – The figure shows the variability of hospitalizations for depression (a) and suicide (b) for Zurich (a densely populated city) and Appenzell Innerrhoden (a scarcely populated rural area) between 2002 and 2014. Each health outcome is normalized using the rate per 10,000 inhabitants.

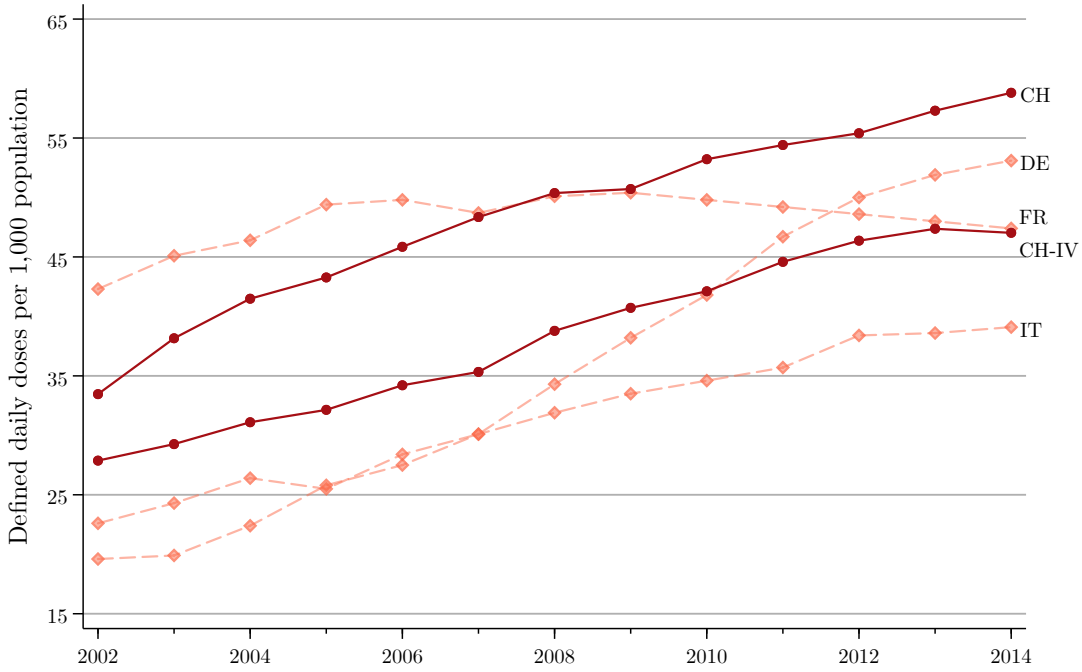
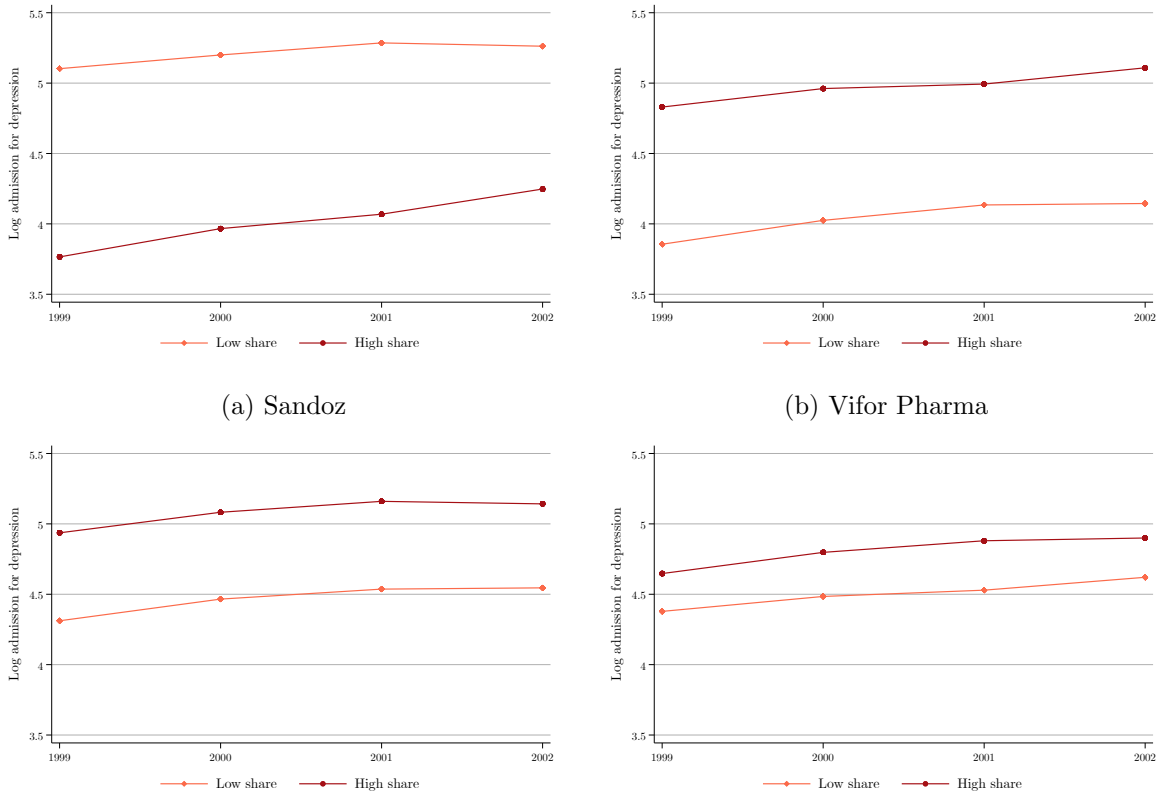


Figure A.3: Practice spillover instrument and changes in antidepressant use over time in Switzerland and in neighboring countries

Notes – The figure shows the trend of antidepressant use in Switzerland (CH), Germany (DE), France (FR), and Italy (IT), between 2002 and 2014, and for the practice spillover instrument (CH-IV) defined by the spatially-weighted average of antidepressant use in the main bordering countries of Switzerland (i.e., DE, FR and IT). Antidepressant use is measured in defined daily doses per 1,000 inhabitants.



(a) Sandoz

(b) Vifor Pharma

(c) MSD

(d) Sanofi-Aventis

Figure A.4: Trends in hospital admissions for depression in small regions with high and low market shares for the other pharmaceutical companies with top-five Rotemberg weights

Notes – For each pharmaceutical company, the figure compares the average of log hospital admissions for depression in regions at the top (high share) and the bottom quartile (low share) of the 2002 regional market share in the period 1999–2002.

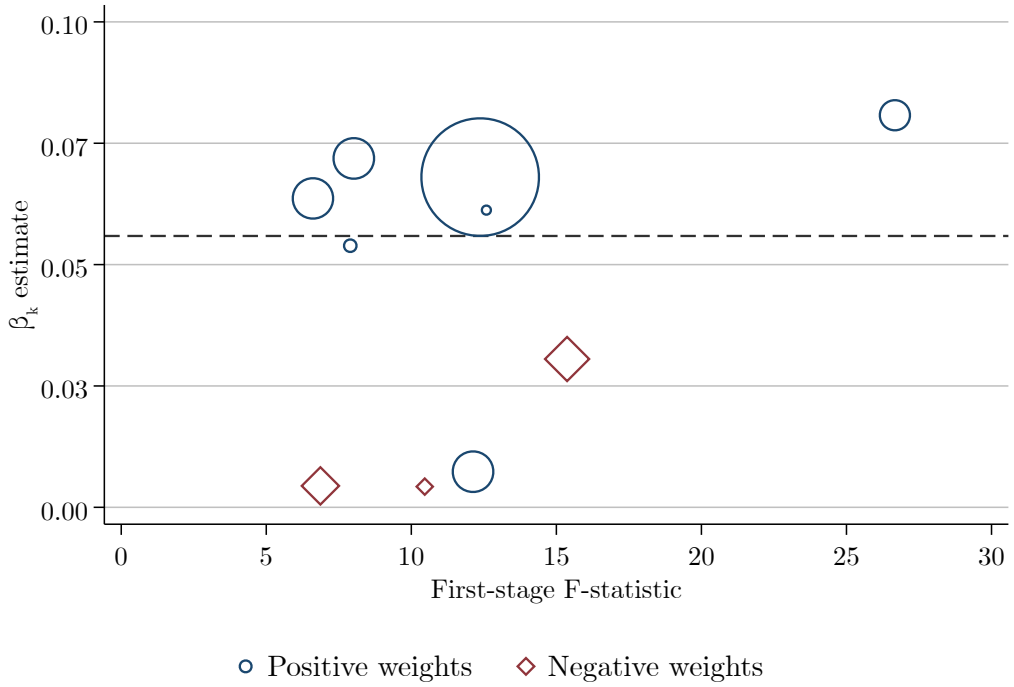


Figure A.5: Heterogeneity of β_m

Notes – The figure plots the relationship between the estimated β_m (on hospitalizations for depression) and their first-stage F-statistic. Each geometric shape indicates a separate instrument estimate. While the size of the shape represents the magnitude of the associated Rotemberg weights, the circles denotes positive weights and the diamonds negative weights. The horizontal dashed line denotes the overall β estimated in the main text by our IV strategy. We only report estimates with an F-statistic > 5.

Table A.1: Antidepressant molecules

ATC class	Molecules
N06A4	Citalopram, Escitalopram, Fluoxetine, Fluvoxamine, Paroxetine, Sertraline
N06A5	Duloxetine, Venlafaxine
N06A9	Agomelatine, Amitriptyline, Bupropion, Clomipramine, Dibenzepin, Dosulepin, Doxepin, Imipramine, Lofepramine, Maprotiline, Mianserin, Mirtazapine, Moclobemide, Nefazodone, Nortriptyline, Opipramol, Reboxetine, Trazodone, Trimipramine

Notes – The table reports antidepressant molecules included in the analysis. The ATC classes N06A4 (Selective serotonin re-uptake inhibitors - SSRI) and N06A5 (Serotonin norepinephrine re-uptake inhibitors - SNRI) represent recent drug classes, while the N06A9 class (tricyclic antidepressants and others) includes older drugs. We exclude the class N06A2 (herbal antidepressants) because defined daily doses cannot be calculated for herbal medicine.

Table A.2: Descriptive statistics of the practice spillover instrument

Country	Mean	SD	Δ(2003/14)	Min	Max
Switzerland	48.64	7.82	20.96	33.43	58.94
Switzerland IV	38.13	7.14	18.79	27.13	47.77
France	48.09	2.32	2.30	42.30	50.40
Germany	36.41	11.20	28.80	22.60	53.10
Italy	30.55	6.98	19.20	19.60	39.10

Notes – The table reports descriptive statistics of antidepressant consumption over the period 2003-2014 for Switzerland, and for each of the countries (i.e., Austria, France, Germany and Italy) used to construct the practice spillover instrument (Switzerland IV). Antidepressant use is measured in terms of defined daily doses per 1,000 inhabitants per day.

Table A.3: Introduction of brand name antidepressants

Pharmaceutical company	Active ingredient	Year
Lundbeck	Escitalopram	2002
Eli Lilly	Duloxetine	2006
GSK Pharma	Bupropion	2007
Servier	Agomelatine	2010

Notes – The table reports the introduction of new brand name antidepressants by pharmaceutical company and year. We do not include the introduction of a new mode of drug administration or package size.

Table A.4: Introduction of generic antidepressants

Pharmaceutical company	Active ingredient	Year
Sandoz	*Citalopram*	*2002*
Sandoz	*Moclobemide*	*2002*
Mepha-Teva	*Mianserin*	*2004*
Mepha-Teva	*Paroxetine*	*2004*
Sandoz	*Fluvoxamine*	*2004*
Sandoz	*Trimipramine*	*2004*
Spirig Healthcare	*Paroxetine*	*2004*
Acino Pharma	Fluoxetine	2004
Helvepharm	Citalopram	2004
Mepha-Teva	Citalopram	2004
Sandoz	Citalopram	2004
Spirig Healthcare	Citalopram	2004
Streuli Pharma	Citalopram	2004
Helvepharm	*Sertraline*	*2005*
Mepha-Teva	*Sertraline*	*2005*
Sandoz	*Sertraline*	*2005*
Spirig Healthcare	*Sertraline*	*2005*
Streuli Pharma	*Sertraline*	*2005*
Helvepharm	Paroxetine	2005
Mepha-Teva	Paroxetine	2005
Sandoz	Fluoxetine	2005
Sandoz	Paroxetine	2005
Streuli Pharma	Fluoxetine	2005
Streuli Pharma	Paroxetine	2005
Winthrop	Citalopram	2005
Mepha-Teva	Fluoxetine	2006
Sandoz	Sertraline	2006
Acino Pharma	Fluoxetine	2007
Actavis	Sertraline	2007
Helvepharm	Fluoxetine	2007
Mepha-Teva	Fluoxetine	2007
Sandoz	Sertraline	2007
Mepha-Teva	*Venlafaxine*	*2008*
Sandoz	*Venlafaxine*	*2008*
Actavis	Citalopram	2008
Adico Pharma	Fluoxetine	2008
Mepha-Teva	Citalopram	2008
Mepha-Teva	Fluoxetine	2008
Mepha-Teva	Sertraline	2008
Semo Trading	Citalopram	2008
Semo Trading	Sertraline	2008
1a Pharma	Citalopram	2009
1a Pharma	Paroxetine	2009

continues on next page

Table A.4 – Continued from previous page

Pharmaceutical company	Active ingredient	Year
1a Pharma	Sertraline	2009
Actavis	Fluoxetine	2009
Actavis	Paroxetine	2009
Actavis	Sertraline	2009
Actavis	Venlafaxine	2009
Axapharm	Fluoxetine	2009
Drossapharm	Venlafaxine	2009
Helvepharm	Venlafaxine	2009
Mepha-Teva	Sertraline	2009
Sandoz	Venlafaxine	2009
Actavis	Sertraline	2010
Helvepharm	Venlafaxine	2010
Mepha-Teva	Venlafaxine	2010
Pfizer	Sertraline	2010
Sandoz	Trimipramine	2010
Sandoz	Venlafaxine	2010
Spirig Healthcare	Fluoxetine	2010
Spirig Healthcare	Venlafaxine	2010
Helvepharm	*Mirtazapine*	*2011*
Mepha-Teva	*Mirtazapine*	*2011*
Sandoz	*Mirtazapine*	*2011*
Streuli Pharma	*Mirtazapine*	*2011*
Helvepharm	Citalopram	2011
Pfizer	Citalopram	2011
Pfizer	Sertraline	2011
Pfizer	Venlafaxine	2011
Sandoz	Trimipramine	2011
Sanofi-Aventis	Trimipramine	2011
Actavis	Mirtazapine	2012
Mepha-Teva	Venlafaxine	2012
Pfizer	Citalopram	2012
Sandoz	Mirtazapine	2012
Spirig Healthcare	Mirtazapine	2012
Actavis	Citalopram	2013
Actavis	Fluoxetine	2013
Actavis	*Escitalopram*	*2014*
Axapharm	*Escitalopram*	*2014*
Helvepharm	*Escitalopram*	*2014*
Mepha-Teva	*Escitalopram*	*2014*
Sandoz	*Escitalopram*	*2014*
Spirig Healthcare	*Escitalopram*	*2014*
Actavis	Citalopram	2014
Actavis	Venlafaxine	2014

Notes – The table reports the introduction of generic antidepressants by pharmaceutical company and year. First introducers are highlighted in italic.

Table A.5: Physician detailing per capita and pharmaceutical companies' market share and revenues

	(1)	(2)	(3)	(4)
Market shares	0.004***	0.005***	0.005***	0.002*
	(0.001)	(0.001)	(0.002)	(0.001)
Revenues	0.142***	0.111***	0.054***	0.064***
	(0.029)	(0.025)	(0.016)	(0.022)
Year FE	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
Canton FE	No	Yes	No	Yes
Region FE	No	No	Yes	No
Company FE	No	No	No	Yes
Observations	3,123	3,123	3,123	3,123

Notes – The table reports the parameter estimates of a series of OLS regressions of pharmaceutical companies' market share and revenues (in the Swiss market) on pharmaceutical companies' detailing per capita. Data aggregate yearly pharmaceutical sales for each company at small region level (SMR) and information on physician detailing for 11 companies from 2015 to 2017. In model (1) we only control for socio-demographic and year fixed effects; in model (2) we add canton fixed effects while in model (3) we add SMR fixed effects; model (4) includes pharmaceutical company and canton fixed effects. We use two-way clustered robust standard errors at region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.6: Estimates of the effect of antidepressant sales on hospital admissions and suicides (log-log specification)

	All admissions	Mental disorder	Depression	Admissions no mental	Suicide
	(1)	(2)	(3)	(4)	(5)
Fixed effects	0.103	0.211	0.940***	0.098	-0.097
	(0.067)	(0.136)	(0.229)	(0.064)	(0.180)
Second stage	0.231**	0.802***	2.517***	0.183	0.295
	(0.117)	(0.274)	(0.667)	(0.115)	(0.410)
First stage	0.954***	0.954***	0.954***	0.954***	0.954***
	(0.180)	(0.180)	(0.180)	(0.180)	(0.180)
Reduced form	0.221**	0.765***	2.401***	0.175	0.282
	(0.111)	(0.297)	(0.491)	(0.108)	(0.389)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	28.031	28.031	28.031	28.031	28.031
Anderson-Rubin F (p-value)	0.053	0.012	0.000	0.113	0.475

Notes – For each health outcome, the table reports the parameter estimates for the fixed effect model (FE) and the instrumental variable model (2SLS) using a log-log specification. Column (1) shows the estimated effects on all hospital admissions, column (2) on hospital admissions related to mental disorders, column (3) on hospital admissions related to depression symptoms, column (4) on all hospital admissions excluding mental disorders, and column (5) on suicides (completed and attempted). All models control for year and region fixed effects and population characteristics. We also include the corresponding first stage and reduced form coefficients from the 2SLS model. We use cluster-robust standard errors at the region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.7: 2SLS estimates of the effect of antidepressant sales on mental-health related outcomes by sex and age groups using the market share instrument

Outcomes (ln):	Mental disorder	Depression	Suicide
Males	0.016**	0.054**	0.003
	(0.007)	(0.022)	(0.014)
Females	0.020***	0.056***	0.049*
	(0.007)	(0.017)	(0.027)
Age $<$ 20	0.013	0.069**	0.021
	(0.011)	(0.032)	(0.057)
Age 20 – 65	0.021***	0.057***	0.019
	(0.007)	(0.021)	(0.015)
Age $>$ 65	0.015*	0.047***	-0.048
	(0.008)	(0.015)	(0.034)
Observations	1,272	1,272	1,272
Kleijbergen-Paap F	15.637	15.637	15.637
Anderson-Rubin F (p-value)	0.094	0.000	0.113

Notes – The table reports the second-stage IV estimates for each health outcome, separately by gender or age. We control for year and region fixed effects and population characteristics. We use cluster-robust standard errors at region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.8: OLS regressions of pre-trends in outcome variables on the regional market share of the top-five companies and instrument growth

Outcomes ($\Delta_{1999-2002}$):	Mental disorders		Depression		Suicide	
	(1)	(2)	(3)	(4)	(5)	(6)
Market share No. 1	-11.314**		-9.427		-11.642	
	(5.135)		(6.993)		(9.614)	
Market share No. 2	20.360*		13.887		36.978*	
	(11.930)		(15.013)		(21.476)	
Market share No. 3	8.564		8.073		-11.395*	
	(7.106)		(8.521)		(6.324)	
Market share No. 4	-3.619		-5.631		5.511*	
	(3.612)		(4.207)		(2.887)	
Market share No. 5	10.652		4.330		5.933	
	(7.583)		(10.586)		(13.883)	
Δ IV 2003/2014		-0.005**		-0.004		0.004
		(0.002)		(0.003)		(0.003)
Observations	106	106	106	106	106	106

Notes – In this table we regress changes in hospitalization outcomes (mental disorders and depression) and suicides in the (pre-) period 1999-2002 on the regional non-antidepressant market shares of the top-five weight companies in 2002, reported in Table 3 (column 1 and 3), and on changes in the market share instrument (ΔIV 1) over the period 2003-2014 (column 2 and 4). All models include canton fixed effects. Heteroskedasticity-robust standard errors are reported in parenthesis. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.9: 2SLS estimates using the market share instrument on placebo outcomes

	All elective hospital admissions	Elective admissions for depression	Infectious diseases	Bone fractures	Pregnancy and childbirth
	(1)	(2)	(3)	(4)	(5)
Fixed effects	0.004	0.013*	0.002	-0.004**	0.001
	(0.003)	(0.007)	(0.002)	(0.002)	(0.001)
Second stage	-0.003	0.020	-0.007	-0.005	0.003
	(0.009)	(0.017)	(0.007)	(0.006)	(0.004)
First stage	0.131***	0.131***	0.131***	0.131***	0.131***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Reduced form	-0.000	0.003	-0.001	-0.001	0.000
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	15.637	15.637	15.637	15.637	15.637
Anderson-Rubin F (p-value)	0.771	0.320	0.325	0.449	0.500

Notes – The table reports the second-stage IV estimates of the effect of antidepressant sales on all elective hospital admissions (column 1), elective hospital admissions for depression-related symptoms (column 2), infectious and parasitic diseases (column 3), bone fractures (column 4), and pregnancy, childbirth and the puerperium (column 5). In the regression models, we control for year and region fixed effects and population characteristics. We use cluster-robust standard errors at region level. Significance at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.10: Poisson estimates of the effect of antidepressant sales on mental health outcomes

	Mental disorder		Depression		Suicide	
Model:	(IV 1)	(IV 2)	(IV 1)	(IV 2)	(IV 1)	(IV 2)
Poisson	0.002	0.002***	0.015***	0.015***	0.001***	0.001***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
2nd stage Poisson	0.035***	0.019***	0.054***	0.046***	0.008	0.007
	(0.010)	(0.006)	(0.014)	(0.009)	(0.013)	(0.008)
2nd stage Residuals	-0.037***	-0.026***	-0.046***	-0.049***	-0.009	-0.010
	(0.011)	(0.008)	(0.015)	(0.009)	(0.016)	(0.010)
1st stage Instrument	0.131***	0.413***	0.131***	0.413***	0.131***	0.413***
	(0.033)	(0.066)	(0.033)	(0.066)	(0.033)	(0.066)
Observations	1,272	1,272	1,272	1,272	1,272	1,272

Notes – The table reports the parameter estimates for each mental health outcome (mental disorders, depression, and suicide) for the Poisson model and the Poisson instrumental-variable regression model. Each regression controls for year and region fixed effects, and includes covariates for demographics. We report the estimates for the shift-share instrument in specification (IV 1) and for the practice spillover instrument in specification (IV 2). We use cluster-robust standard errors at the region level. The clustered standard errors are bootstrapped with 1,000 replications and replacement for the 2nd stage Poisson regression. Significance levels at 10%, 5%, and 1% indicated by *, **, and ***, respectively.

Table A.11: Estimates of the effect of antidepressant sales on hospital admissions and suicides at the pharmaceutical company level

	All admissions	Mental disorder	Depression	Admissions no mental	Suicide
	(1)	(2)	(3)	(4)	(5)
Panel A: all pharmaceutical companies					
Second stage	0.004	0.010	0.040**	0.004	-0.003
	(0.004)	(0.007)	(0.016)	(0.003)	(0.006)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	9.605	9.605	9.605	9.605	9.605
Anderson-Rubin F (p-value)	0.001	0.000	0.000	0.001	0.663
Panel B: only Mepha-Teva					
Second stage	0.006	0.012	0.068***	0.006	-0.001
	(0.005)	(0.009)	(0.023)	(0.005)	(0.009)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	13.516	13.516	13.516	13.516	13.516
Anderson-Rubin F (p-value)	0.146	0.147	0.000	0.160	0.929
Panel C: without Mepha-Teva					
Second stage	0.004	0.010	0.034**	0.004	-0.005
	(0.004)	(0.007)	(0.014)	(0.003)	(0.006)
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	8.845	8.845	8.845	8.845	8.845
Anderson-Rubin F (p-value)	0.001	0.000	0.000	0.001	0.793

Notes – For each health outcome, the table reports the parameter estimates for the Limited Information Maximum Likelihood estimator (LIML) using 16 instruments, one for each pharmaceutical company. Column (1) shows the estimated effects on all hospital admissions, column (2) on hospital admissions related to mental disorders, column (3) on hospital admissions related to depression symptoms, column (4) on all hospital admissions excluding mental disorders, and column (5) on suicides (completed and attempted). All models control for year and region fixed effects and population characteristics. We use cluster-robust standard errors at the region level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.