I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13727

Happy Pills? Mental Health Effects of the Dramatic Increase of Antidepressant Use

Giuliano Masiero
Fabrizio Mazzonna
Sandro Steinbach

SEPTEMBER 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13727

Happy Pills? Mental Health Effects of the Dramatic Increase of Antidepressant Use

Giuliano Masiero
DIGIP, University of Bergamo, IdEP and USI

Fabrizio Mazzonna
IdEP, USI, MEA and IZA

Sandro Steinbach
University of Connecticut

SEPTEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Happy Pills? Mental Health Effects of the Dramatic Increase of Antidepressant Use*

Despite the growing skepticism regarding the efficacy of antidepressants, global consumption has increased at an unprecedented rate with unknown implications for society. We estimate the causal effect of this increase on mental health outcomes using an instrumental variable strategy that exploits pharmaceutical company local market power and the availability of detailed drug sales data from Switzerland between 2002 and 2014. Our main instrument, a modified version of the popular shift-share instrument, relies on the national growth in antidepressant sales for pharmaceutical companies (the shift) – mainly due to product innovation – and assigns it locally using regional non-antidepressant market shares. Our estimates show that an increase in antidepressant sales causes a sharp increase of hospital admissions related to depression symptoms. An alternative instrument, which exploits prescribing practice spillovers from neighboring countries, leads to very similar point estimates providing further evidence about the validity of our results.

JEL Classification: I12, I18
Keywords: depression, antidepressant treatment, suicides, mental health

Corresponding author:
Fabrizio Mazzonna
Università della Svizzera Italiana
Institute of Economics
via Buffi 13, 6900 Lugano
Switzerland
E-mail: fabrizio.mazzonna@usi.ch

* We thank Olaf Verbeek for his initial research support on this work and Osea Giuntella and Nicolas Ziebarth for insightful comments on early drafts of this paper.

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). In the US antidepressants are the third most prescribed class of drugs, with 12% of people aged 12 and over who reported to take antidepressants within the last month (National Health and Nutrition Examination Surveys 2011–2014). This study investigates the health consequences of the recent upsurge in antidepressant use in Switzerland showing causal evidence of adverse mental health effects. 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 50%. In 2016, the prevalence rate of antidepressant prescription 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 pharmaceutical industry represents almost 5% of Swiss GDP.

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) (Lane et al., 1995). Given the well-know economic burden of depression, which is associated with increasing disability and absenteeism (e.g., James et al., 2018) and lower productivity (Stewart et al., 2003), a marked growth in depression treatment can be seen as a positive outcome. However, there is expanding evidence that only part of the recent development can be attributed to a rise in depression prevalence or in medical treatment for people with major depressive symptoms. 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 (Mojtabai and Olfson, 2011; Wong et al., 2017; Haller et al., 2019). More generally, prescription threshold shifted towards milder forms of depression (Moore et al., 2009).

There is also rising concern over the efficacy of these drugs, the occurrence of adverse events and the potential side-effects related to long-term use and overprescription. Although the evidence on the efficacy of antidepressants is mostly grounded on randomized controlled trials (RCT), several meta-studies have questioned the results. Critics point to methodological issues, such as the short duration of the RCT, small sample size and under-reporting of adverse events during the trials and questionable clinical significance (e.g., Fournier et al., 2010; Jakobsen et al., 2017). Overall, these studies find evidence of clinical efficacy only for the most severely

depressed and in most cases the placebo effects account for about 80% of the total measured effect (Currie and MacLeod, 2019). Moreover, their use is associated with an increase in suicide risk among pediatric patients (Cipriani et al., 2016; Stone et al., 2009), and with several adverse health outcomes among the elderly (Coupland et al., 2011). 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.[1] Not less important are the unknown consequences of off-label prescriptions and use among people with mild depression symptoms 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). Moreover, inappropriate use of antidepressants, especially SSRI, is also associated to withdrawal symptoms that in some cases might be particularly severe (Davies and Read, 2019).[2]

These concerns fostered an extensive literature on the spatial correlation between antidepressant consumption and suicides to evaluate the population health effects in the "real world". 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 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, the large part of antidepressants consumed are SSRI, and the marginal patient treated with antidepressant drugs likely suffers from a milder form of depression as compared to the marginal patient at the beginning of the '80s when SSRI were introduced, even without considering off-label use. Note also that 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.

This paper provides new insights into the analysis of mental health consequences of the surge in antidepressant use. We can exploit the large geographical and time variation in antidepressant sales at product level and individual hospital admissions and suicide events for 13 years, between 2002 and 2014, for the whole Swiss country (see Figure 1).

An important contribution of this paper is the focus on the effect of antidepressants on hos-

[1] 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.

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

pitalization for mental health conditions and depression rather than relying only on suicides as in previous literature (e.g., Ludwig et al., 2009). Suicide is a very low probability event and, therefore, its statistical power tends to be low (Currie and MacLeod, 2019). Moreover, hospitalizations have a very large impact on health care costs.

We address the endogeneity concern that arises in ecological studies using two 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 AD sales in a region is a weighted average of the national manufacturer (pharmaceutical company) growth rates (the shifts), but the weights depends on the manufacturer 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 manufacturers in different Swiss regions, while the variation in the growth rates (net of year fixed effects) mainly comes from the introduction of new products in the market (as reported in Figure 2). 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. In other words, our instrument implies that whenever a pharmaceutical company introduces a new antidepressant in the market, it is more likely to sell the new product in regions where it has a larger market share in the non antidepressant market, compared to its competitors. The only fact that our instrument has predicting power represents another important contribution of this paper because it 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 their product among physicians.[3]

Using the before described research design, we find that an increase in consumption of one defined daily dose per 1,000 inhabitants (roughly 3% of 2003 sales) increases hospital admissions for mental disorders by 2%, driven by an even larger increase in hospitalizations for depression (6.5%). The evidence on suicides in noisier although point estimates are generally positive as for hospital admissions.

[3] There is large literature in economics on the causal link between pharmaceutical marketing and prescription drug utilization including antidepressants. Most of this literature focuses on direct to consumer advertising (e.g., Shapiro, 2018; Sinkinson and Starc, 2019; Shapiro, 2020). As explained in Section 2, direct-to-consumer advertising is not allowed, price are set at federal level, so pharmaceutical companies compete over physician detailing.

Following the recent and growing literature that investigates the formal conditions underlying the validity of the Bartik instrument (Borusyak et al., 2018; Goldsmith-Pinkham et al., 2020), we provide 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 manufacturer (interacted with time fixed effects) as instruments in a weighted GMM estimation. We calculate the weights on these instruments, the so-called "Rotemberg weights", that allow us to identify the manufacturers that account for the largest share of the identifying variation. From this calculation, we show that roughly half of our identifying variation comes from one company which experiences a dramatic increases in AD sales in the Swiss market over the observation window, driven by the introduction of a series of new generics drugs. As shown in Figure 4, this allows us to demonstrate that there are no differential pre-trends in mental health outcomes across regions with high and low market shares of this company. Moreover, in a series of placebo estimates aimed to test the exogeneity of our instrument indirectly, we do not find any evidence of relevant correlations between the increase in antidepressant sales induced by our instrument and hospital admissions for diseases that should not be affected by the increase in antidepressant use.

As additional robustness check, we also 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 should be especially true for Switzerland, where over 25% of the doctors studied in one of the four (main) neighboring countries (Germany, Italy, France and Austria). Under the assumption that the magnitude of these spillovers is inversely related to geographical distance, we build our instrument using spatially weighted averages of antidepressant sales in neighboring countries and assign them to local areas. Reassuringly, the results using this alternative estimation strategy lead to very similar point estimates.

The remainder of this paper is organized as follows. In the next section, we introduce the Swiss institutional setting in which we conduct our analysis. In Section 3, we describe the data used for the analysis and the data aggregation process. In Section 4, we discuss our empirical strategy while in Section 5 we present our results. Finally, Section 6 concludes.

2. Institutional setting

We use data for Switzerland to study the effects of antidepressant use on mental health outcomes. 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. Moreover, the deductible is quickly exhausted by physician visits and psychotherapy consultations.[4] 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.

[4] 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.

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.[5] These market characteristics suggest that drug manufacturers 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, manufacturers can only influence their sales through physician detailing.

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 mental health outcomes for Switzerland covering the period from 2002 to 2014. The data is aggregated at the small area level (SMR - spatial mobility region). This level divides Switzerland into 106 SMR regions, each of them accounting for approximately 45,000 individuals. A 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

[5] 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.

additional source of measurement error may arise from the use of wholesale data (from the manufacturer 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.

To account for confounding factors, we supplement our primary datasets on antidepressant sales and mental health outcomes with data on essential covariates for each region and year. These variables are the distribution of the population across gender and age, the share of German-speaking people, the share of foreigners, and the average municipal unemployment rate. Lastly, we obtained access to anonymized data by the Swiss Medical Association (FMH), which allowed us to calculate the share of antidepressant prescribing specialists (Neurologists and Psychiatrists) and GPs per 10,000 inhabitants.

3.1 Antidepressant sales and mental health outcomes

We obtained data on antidepressant sales from IMS Health Switzerland. 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 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. We also use an accessory dataset with annual sales for the universe of all other drugs aggregated at the manufacturer level by pharmaceutical sales region for the period from 2002 to 2014.[6]

We obtained individual-level data on mental health outcomes 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

[6] A negligible number of drugs with a retail price greater than 5,000 CHF is also not included in the analysis.

used by the FSO to anonymize individual-level hospital admission data.[7] 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 us to build a final dataset with comparable spatial data on both antidepressant consumption and mental health outcomes. We express mental health outcomes in terms of annual prevalence per 10,000 inhabitants by SMR throughout our analysis.

We consider three different mental health outcomes in our analysis. To capture the impact on suicide, we create a measure of completed suicides and hospital admissions for suicide attempts (Intentional Self-harm - X60-X84). We also account for hospital admissions for depression (depressive episode - F32, and Recurrent Depressive Disorder - F33), and hospitalizations for other mental health conditions (Chapter V).[8] Data on mortality are from the official Swiss mortality statistics, and hospitalization data are from the Swiss hospital statistics.

Finally, to construct our second instrument based on spatial spillovers in prescription practice, we use OECD data on antidepressants consumption in DDD for Austria, France, Germany and Italy from 2003 and 2014 (see Table A.5 for descriptive statistics).

3.2 Descriptive evidence

We summarize the main variables used in our analysis in Table 1. The average antidepressant consumption across SMR regions 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 in all regions. However, most of the variation is across small areas, and another important source of variation comes from the introduction of new products. Figure 2 shows the number of newly introduced antidepressants per year.[9] We will discuss the use of

[7] 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. We follow the approach developed by Filippini et al. (2019) We obtained detailed information on the general population at the postal code level for 2010 from the FSO. We use this information to create weights and recode the location information to obtain a match between the new and the old definition. 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.

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

[9] Additional details (manufacturer, active ingredient and year of introduction) for new brand name products and generic drugs are provided respectively in Table A.2 and Table A.3.

this source of variation to construct our instrument for antidepressant consumption in Section 4. 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 3. 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. The prevalence of both mental disorders and depression appear to increase in the North-Eastern areas compared to 2014 according to subfigures (b) and (d). 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.

4. Empirical strategy

Following previous literature on the effect of antidepressant sales or consumption on suicides (e.g. Ludwig et al., 2009), our estimates of the effect of antidepressant sales on mental health outcomes relies on the following empirical model:

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

where y_{rt} is the natural log of the mental health outcome (number of hospital admissions for mental health problems, depression or number of suicides)[10] in region r at time t; AD_{rt} represents antidepressants sales (expressed in DDD per 1'000 inhabitants in year t) in the same region at the same time; 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), the level of unemployment in a region, and the density of antidepressant prescribing physicians; ϑ_r are region fixed effects, ϑ_t are time fixed effects, and ϵ_{rt} is an idiosyncratic error term. Following the previous literature, we use the natural log of mental health outcomes to approximate a

[10] 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.

normal distribution for our data.[11] As a robustness check, we also estimate (1) assuming a data generating process that mimics the Poisson distribution.

We use the fixed effects (FE) estimator as our benchmark model to estimate Equation (1). Given the substantial scale differences between small areas (we move from almost half a million 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 areas show a higher signal-to-noise ratio, which is an issue, especially when dealing with a low-frequency outcome such as suicides.[12]

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. In particular, 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 create awareness of depression and decrease stigma in the general population and among health care practitioners.[13] We would expect such policy interventions to have a positive effect on antidepressant sales since they encourage uptake of antidepressant treatment, and possibly a negative impact on our mental health outcomes, introducing a downward bias in our estimates. To account for these issues, we rely on an instrumental variable strategy to estimate the impact of antidepressant sales on mental health outcomes.

4.1 Instrumental variable approach

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 manufacture are assigned to regions using the supply-driven increase in antidepressant sales. More specifically, we allocate the annual sales of antidepressant drugs of each manufacturer using its

[11] We take care of zeros for suicides by adding a small constant to each outcome.

[12] Comparing Zurich (a densely populated city) and Appenzell Innerhoden (a scarcely populated rural area), we observe a large variability in mental health outcomes (see Figure A.1).

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

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 manufacturers are likely to sell more in areas where the mental health conditions of the population are poor. The use of market shares allows us to overcome this problem since we consider the sales of each manufacturer 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 manufacturer in a region and to overcome the residual concerns regarding the potential endogeneity of our shares. In practice, we exploit the fact that manufacturers tend to sell more in areas 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 manufacturers (net of the overall yearly change in AD 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).[14] The robustness of our findings to the increasing time gap confirms our main results.

More formally, our instrument is constructed as follows. Let the annual national stock of antidepressant for each manufacturer be represented by $AD_{mt} = \sum_r AD_{mrt}$, where AD_{mrt} represents the antidepressant sales for manufacturer m in region r at time t. This stock is used to calculate

[14] 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.

the shifts, i.e. the variation in AD sales over time. The shares instead are calculated as:

$$\widetilde{S}_{mr2002} = \frac{MS_{mr2002}}{\sum_r MS_{mr2002}}, \tag{2}$$

where $\text{MS}_{mr2002} = \frac{v_{mr2002}}{\sum_m v_{mr2002}}$ is the market share from wholesales of non-antidepressants (v_{mrt}) for manufacturer m, in region r, in the base year 2002.[15]. This allows us to redistribute the national stock AD_{mt} as follows:

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

where the regional variation in $\widetilde{AD}_{mrt}$ comes from variation in manufacturer non-antidepressant market shares in the base year, and the temporal variation comes from the national growth rates of the manufacturer. 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 τ to be positive if manufacturers are more successful in pushing their sales in areas where they have a larger market share as compared to other areas.

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 manufacturers can only compete on quantity by physician detailing. As such, our instrument captures the potential influence that a manufacturer can exert in a certain area with respect to its competitors in that area. τ captures how much this market power actually influences antidepressant sales.

[15] To construct these market shares we use total sales based on ex-factory prices because we do not have information on single products, and hence on final prices.

The estimate obtained from the first stage is then used in the second stage as

$$y_{rt} = \beta_0 + \beta_1 \widehat{AD}_{rt} + \beta_2 X_{rt} + \vartheta_r + \lambda_t + \epsilon_{rt}, \tag{6}$$

where $\widehat{AD}_{rt}$ is the predicted antidepressant consumption obtained from (5). β can then be interpreted as the effect of a change in the exogenous share of antidepressant consumption on mental health outcomes and, therefore, as the causal effect of antidepressant consumption on mental health outcomes. One should bear in mind that, in 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 mental health outcomes for those who consumed antidepressants because of the manufacturer market power but would not have used antidepressant had market power been absent.

As already mentioned in the Introduction, 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 a single share as an instrument. The (Rotemberg) weights on these instruments are then used to make clearer our identifying variation and to provide further tests about the validity of our research design.

Given the count nature of our original outcomes (before the log transformation), we evaluate the robustness of our results to the use of non-linear estimation methods based on a Poisson count data model. Following Terza et al. (2008), we take the endogeneity of antidepressant sales into account in a non-linear model using the so-called "two-stage residual inclusion", 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.

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% of doctors practicing in Switzerland have foreign qualifications and almost all of them (25%) studied in one of the four big neighboring countries, namely Germany, Austria, France and Italy. In Figure A.2 we report the average DDD of antidepressant consumption per 1000 inhabitants from 2002 to 2014 for these 4 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 local areas as follows:

$$\widetilde{AD}_{rt} = \frac{\sum_c w_{cr} AD_{ct}}{\sum_c w_{cr}}, \tag{7}$$

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. In other words, 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 language area.

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 spillover in prescription practice (exclusion restriction). The inclusion in our specification of a full set of region fixed effect, that accounts for other time invariant cultural confounders, makes this assumption more reasonable. However, we cannot exclude that other unobserved treatment practices correlated with changes in antidepressant prescription might violated the exclusion restriction. For this reason, we use this instrument mainly as a robustness check.

5. Results

2SLS estimates using pharmaceutical industry market power

Table 2 shows the baseline results for our analysis using our main instrument based on the pharmaceutical industry market power. For each mental health outcome (hospital admissions for mental disorders, hospital admissions for depression, and suicides), we present the results of two model specifications. Model (1) includes demographic controls and year and region fixed effects. This model is our preferred specification because it includes only regional demographic characteristics, which should not be affected by changes in AD consumption. Model (2) includes physician density and unemployment rate as additional regressors, which might be affected by

changes in antidepressant consumption (potentially biasing our estimates). Reported standard errors are robust and two-way clustered (Cameron et al., 2011) at region and year level, so robust to arbitrary within-panel autocorrelation (region) and to contemporaneous cross-panel correlation (year).

As the benchmark, we report the results obtained from FE in the first row of Table 2. The FE estimates indicate that one DDD increase in antidepressant sales in a region is associated with 1.3% more hospital admissions for depression, while there is no evidence of a significant correlation with hospitalizations for mental disorders in general. The estimates of antidepressant sales on suicides are also statistically insignificant.

The second row of Table 2 reports the results of our 2SLS estimates, while we provide the respective first stage and reduced form results in the third and fourth rows. Even when we exploit the arguably exogenous variation in antidepressant sales due to differences in market shares of pharmaceutical companies, we find that an increase in antidepressant sales substantially increases hospital admissions for mental disorder and depression. The estimated effects are also larger in magnitude than those estimated with FE. In particular, we find that a one-unit increase in DDD of antidepressant sales per 1'000 inhabitants leads to a change of hospital admissions for mental health disorders of 2%. This effect that is mainly driven by the increase in hospitalizations for depression (+ 6.5%).[16] As for the estimated effect of antidepressant sales on suicides, we find no evidence for a positive relationship and the point estimates suggest a slightly negative impact (less than one percentage point).

We report the first-stage results in the third row of Table 2. The estimates show that our instrument predicts antidepressant sales well. Although the value of the Kleijbergen-Paap F-statistic for weak instruments is only slightly above ten for model specification (2), it is worth noting that we have almost saturated the model with a full set of fixed effects and control variables (some which also potentially endogenous) and reported the most conservative standard errors. Reassuringly, the fourth row of Table 2 also provides evidence for the presence of significant reduced form effects, especially for hospital admissions related to depression.

In Table 3 we open up the black box of our Bartik-type instrument reporting the summary statistics about the Rotemberg weights collapsed at manufacturer level (as in Goldsmith-Pinkham

[16] The estimates of the impact on all mental disorders excluding depression are smaller and not statistically significant.

et al., 2020).[17] While Panel A shows the presence of negative weights on some manufacturers, these weights are on average very small and, if ever, they reduce the estimated effect of AD on hospitalizations for depression (see the values in the α-weighted β column). The correlations between manufacturer aggregates reported in Panel B indicates that the weights (α_m) are largely correlated with national growth rates (g_m) and, to a lower extent, with the variation in shares across regions (Var(z_m)). In other words, this indicates that the variation exploited by our estimator derives mostly (although not exclusively) from differences in product innovation across manufactures in the observation window than from the variation in market shares across regions. This is reflected also in Panel C, where the top 3 manufacturers accounts for 87% of the positive weights. The top weight manufacturer, Mepha-Teva, explains alone 55% of the positive weights. Indeed, Mepha-Teva experienced an extraordinary growth (the value of g_m in the second column), progressing from an average of one DDD to 13 DDD per 1000 inhabitants (which represents almost one quarter of the 2014 antidepressant market), thanks to the introduction of several new generic drugs starting from 2004 (see also Table A.3). Since most of the variation in our research design comes from this company, in Figure 4 we compare trends in average hospital admissions for depression in regions with high and low market shares of Mepha-Teva, as in a difference-in-difference design (DiD). We use the top and the bottom quartiles of non-AD regional market shares in 2002 as a sort of measure of exposure to the shock, which is represented by the introduction of new generics in the Swiss market. It is reassuring to observe that the trends in hospital admissions in the two groups of regions between 1998 and 2003 are parallel.[18] Conversely, the two trends clearly diverge starting from 2004, with high-share regions experiencing a larger increase in hospital admissions. It is worth noting that although most of the variation is driven by this company our estimate can be generalized also to the antidepressants sold by other companies. In the Appendix (Figure A.3), we show that the point estimates associated with the other major companies are all positive (although there is some evidence of heterogeneity) except for one company that has a very low value for the F-statistic on the excluded instrument suggesting very low power and precision.

In the Appendix, we presents the estimation results obtained when we exploit the heterogeneity across gender and age groups (A.4). In the case of hospitalizations for depression, we find some

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

[18] the overall increasing trends in hospital admission between 1998 and 2001 is partially due to the increasing coverage of Swiss hospitals in the administrative records that became full only in 2001.

evidence of heterogeneity across age groups, with somewhat larger effects on depression among teenagers (+8%) and smaller among the elderly (+5.5%). Larger heterogeneity is found for suicides. In particular, we estimate a significant increase in suicides for women and people aged 20–65. However, since we do not find similar evidence of heterogeneity for the other mental health outcomes and given the risk of a "false positive" when breaking down estimates for such a rare event like suicides is high, we avoid speculating on these differences. Moreover, while our outcomes vary by gender and age, antidepressant sales and the instrument do not.

2SLS estimates using spatial spillovers

In Table 4 we replicate our main results using the alternative instrument based on spatial spillovers in prescription practice. It is remarkable that our estimates based on this alternative instrument leads to very similar point estimates. For instance, in the case of depression the estimated effect of one DDD increase in antidepressant sales is 6.7% vs. 6.5% using our main instrument. It is also worth noting that this instrument is particularly powerful allowing to reasonably reject the hypothesis of weak instrument bias (see the value of the Kleibergen-Paap Wald F statistic).

5.1 Robustness checks

In our baseline regression model we use 2002 as the base year, and estimate the treatment effect for the period between 2003 and 2014. As a robustness check, we keep the base year at 2002 and, in steps of one year, we use incrementally fewer data to estimate the treatment effects. In practice, we increase the time gap between the base year used to construct the share and our observation window. In this way, we effectively decrease the sample size and, therefore, the variation that we can exploit. Nonetheless, for all the three outcomes we find consistent point estimates and standard errors for all the three outcomes (Figure 5). Subfigure (a) shows a stronger treatment effect for mental health disorders when we use increasingly more recent data. The estimates for depression appear to become even more precise with a decreasing sample size, as indicated in Subfigure (b) of Figure 5.

In Table A.6 we report 2SLS estimates of the effect of an increase in antidepressant sales on all the other hospital admissions —distinguishing between emergency and elective admissions— and a set of placebo outcomes —namely infectious diseases, bone fractures and pregnancy— selected to match the mean and standard deviation of our primary outcomes. Since patient with mental health crisis are usually advised to proceed to the nearest ER for assessment, we

can expect some residual effect on non-mental health related emergency hospital admissions but not on elective. Although not statistically significant we estimate an increase of almost 1% in emergency admission while the point estimate on elective is basically zero. Moreover, none of the three placebo outcomes show a significant (and relevant) correlation with the instrumented antidepressant sales.

Table A.7 shows that the 2SLS results are similar when we deal with the count nature of the outcome variables. For each instrument, we apply an IV strategy to Poisson models using the control function approach. For mental health and depression related hospitalizations, point estimates are very similar to those reported in the main text using the log-linear regression model, while for suicides are somewhat larger and even statistically significant at 1% level. This might be due to the fact that count data nature of suicides is better capture by the Poisson model.

6. Conclusion

This research sheds light on the mental 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 disorder by 2% and for depression by 6.5%. Our evidence on suicides is not always statistically significant but points out 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. Almost identical results are found when we use a second instrument which exploits spatial spillovers in antidepressant prescribing practices from the four neighboring countries.

Our results are in contrast with earlier evidence by Ludwig et al. (2009), who find that the increase in SSRI use decreases suicide mortality by 5%. These authors compare SSRI use across countries and over time in the 80s and the 90s. At that time, SSRI were promoted as being more efficient than TCA, particularly in terms of reduced side effects. Because the analysis includes the introduction period of SSRI, the study probably captures the initial impact of their uptake. The current market, however, is dominated by 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 two very different research design might help to generalize our result at least in the light of the current levels of treatment with antidepressants. Evidence suggests that 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 put measures in place to ensure adherence to the prescription guidelines and emphasize the importance of alternatives to pharmacotherapy.

Our research does not shed light on the cause of over-treatment with pharmacotherapy. For instance, over-treatment could be the result of undercapacity of psychotherapists. Psychotherapy is a form of treatment more time-consuming than prescribing antidepressants. Decreasing stigma and increased 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. However, previous evidence on extensive off-label prescription practice and our evidence on pharmaceutical company local market power suggest that the influence of pharmaceutical company over doctor prescription practice might be one of the cause of over-treatment.

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.

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.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

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. *BMJ*, 343:d4551.

Currie, J. M. and MacLeod, W. B. (2019). Understanding physician decision making: The case of depression. *Econometrica*, forthcoming.

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.

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*, 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.

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.

Lane, R., Baldwin, D., and Preskorn, S. (1995). The SSRIs: Advantages, disadvantages and differences. *Journal of Psychopharmacology*, 9(2):163–178.

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. *BMJ*, 339:b3999.

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

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.

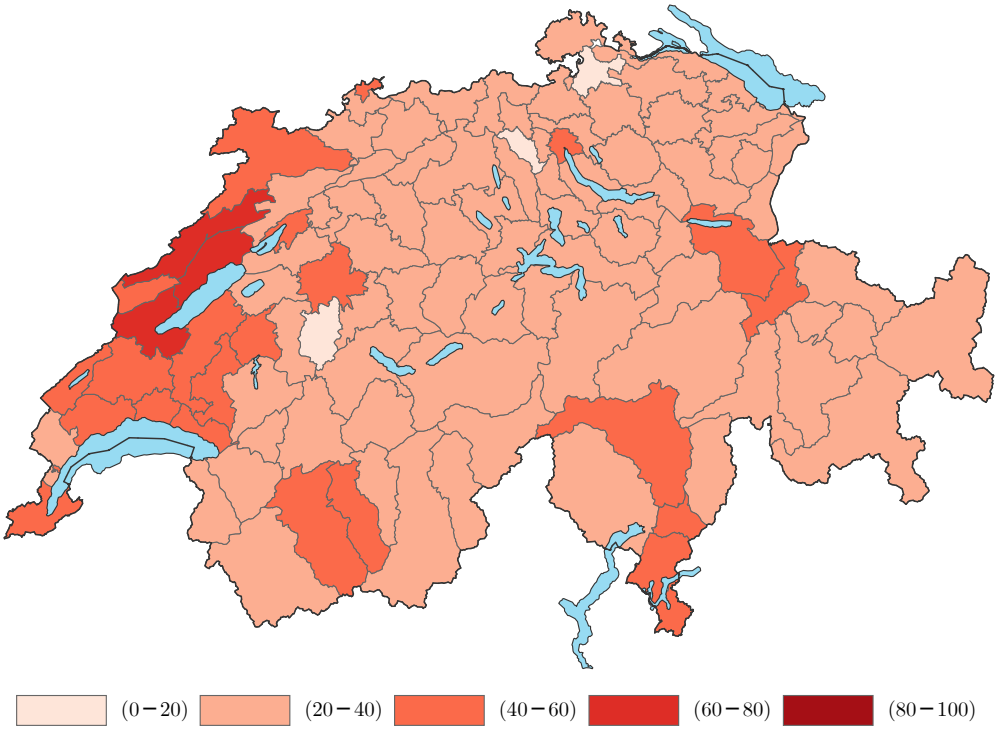
Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.

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. *Jama*, 289(23):3135–3144.

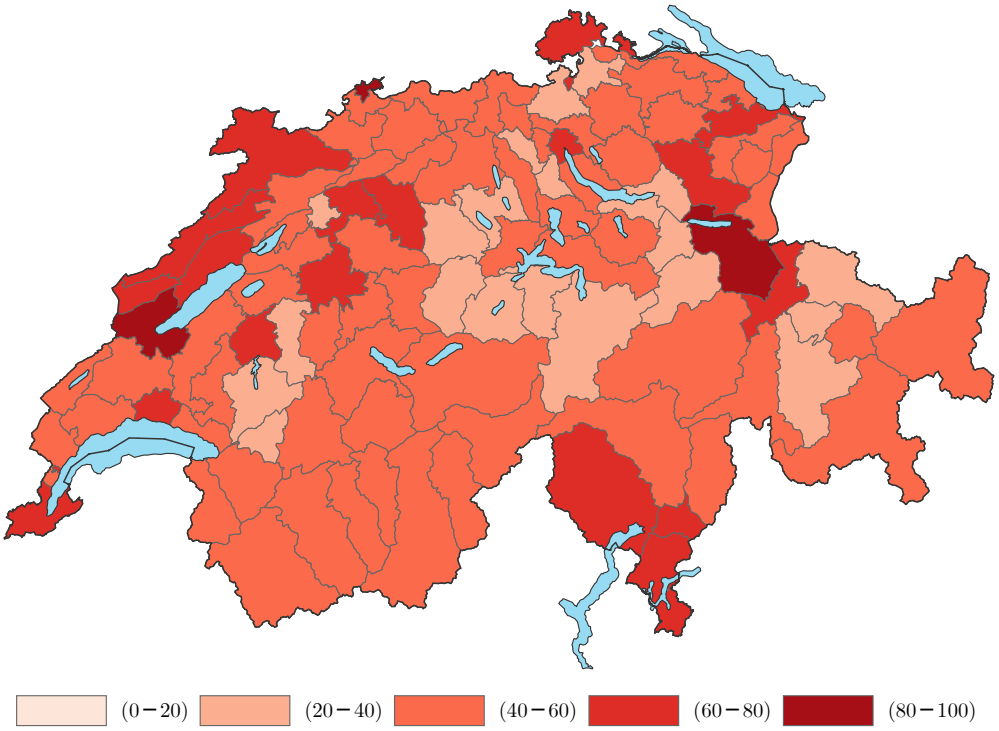
Stone, M., Laughren, T., Jones, M. L., Levenson, M., Holland, P. C., Hughes, A., Hammad, T. A., Temple, R., and Rochester, G. (2009). Risk of suicidality in clinical trials of antidepressants in adults: Analysis of proprietary data submitted to US Food and Drug Administration. *BMJ*, 339:b2880.

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. *bmj*, 356:j603.



(a) Antidepressant sales in 2003



(b) Antidepressant sales in 2014

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

Notes – The figure examines antidepressant sales in Switzerland at the small area 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.

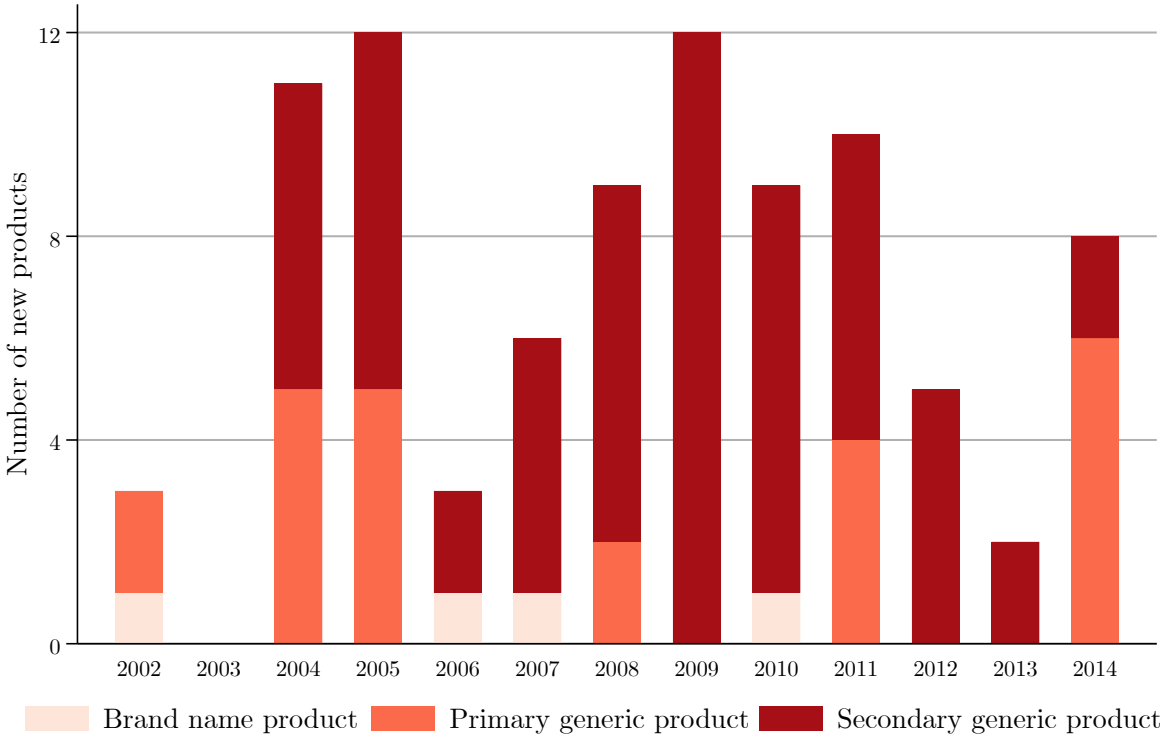
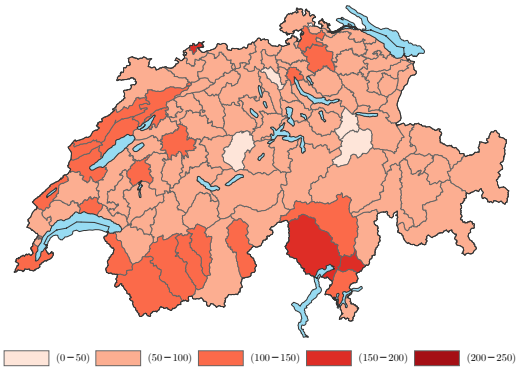
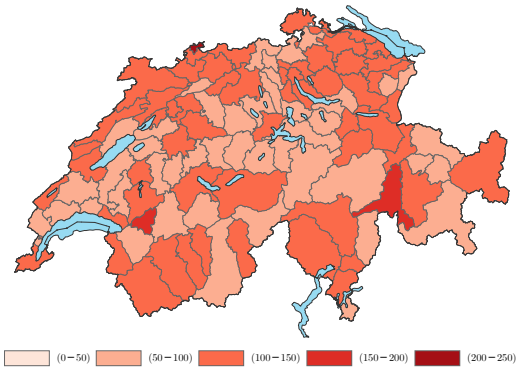


Figure 2: Introduction of new anti-depression drugs in Switzerland from 2002 to 2014

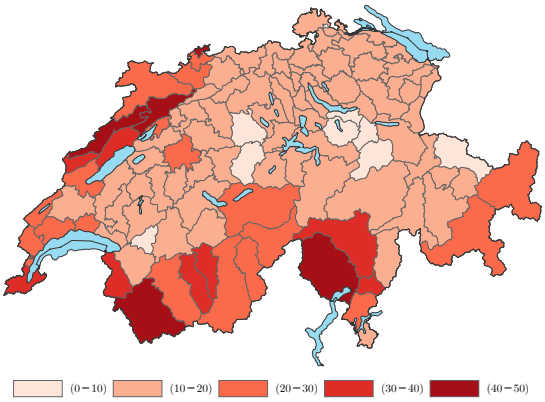
Notes – The figure shows the introduction of new antidepressant drugs in Switzerland from 2002 to 2014. Light red bars indicate brand name products, red bars primary generic products, and dark red bars secondary generic products, respectively.



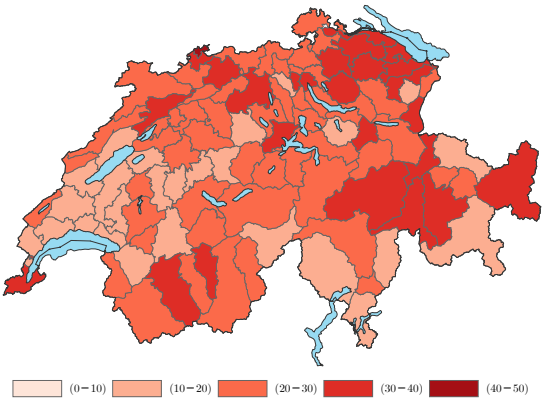
(a) Mental disorder in 2003



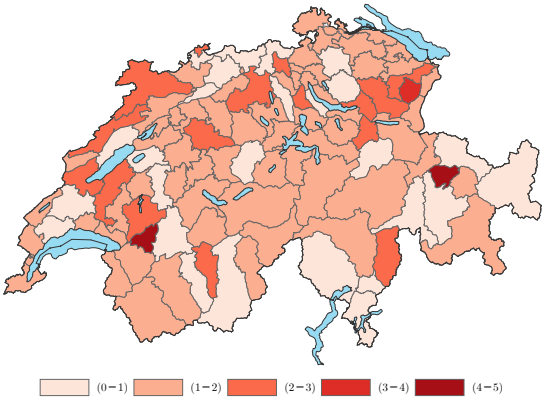
(b) Mental disorder in 2014



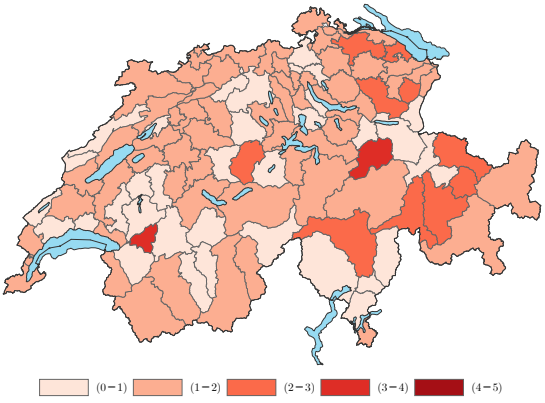
(c) Depression in 2003



(d) Depression in 2014



(e) Suicide in 2003



(f) Suicide in 2014

Figure 3: Mental health outcomes in Switzerland by small areas in 2003 and 2014

Notes – The figure classifies the prevalence of mental health disorder, 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.

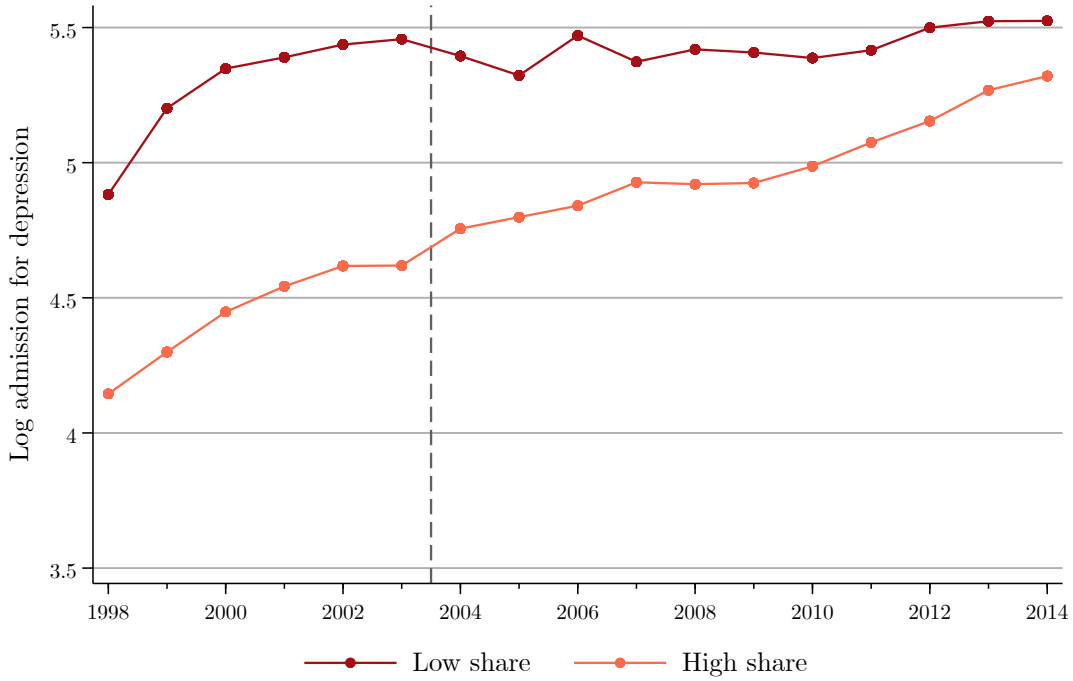
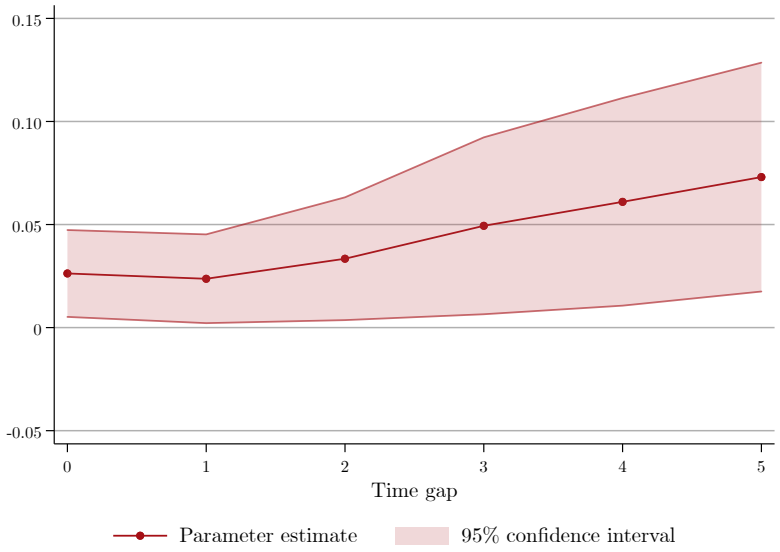
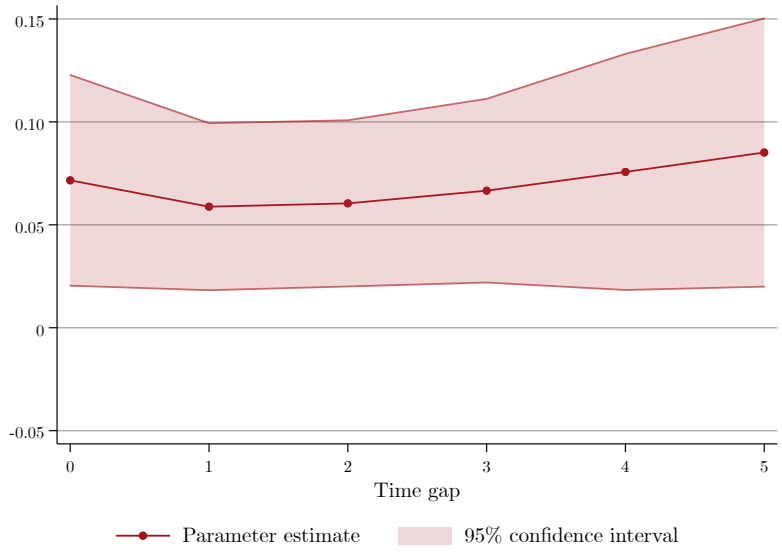


Figure 4: Trends in hospital admissions for depression in areas with high and low Mepha-Teva market shares

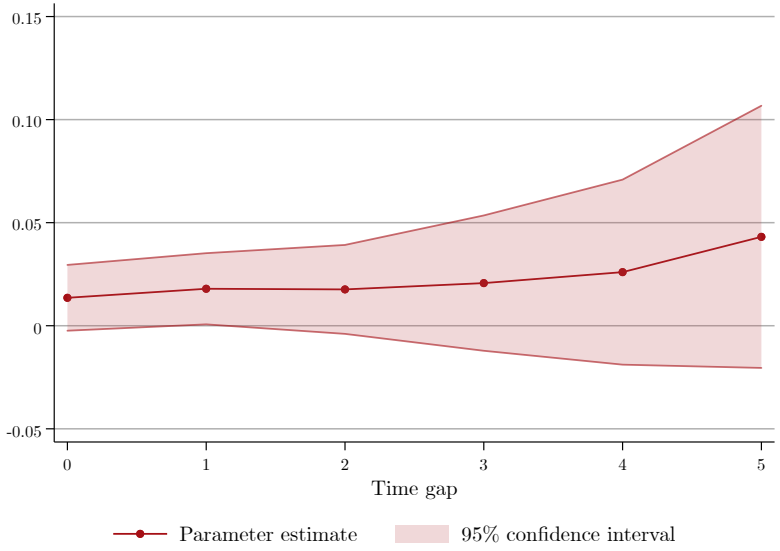
Notes – The figure compares the average of log hospital admissions for depression in regions at the top (high share) and at the bottom quartile (low share) of the 2002 regional market share of Mepha Teva. The vertical red line between 2003 and 2004 marks the year of introduction of several new generic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



(a) Mental disorder



(b) Depression



(c) Suicide

Figure 5: Parameter estimates by health outcome relative to 2002

Notes – The figure shows IV estimates for mental disorder (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.

Table 1: Descriptive statistics

		Standard deviation					
	Mean	Overall	Between	Within	Δ(2003/14)	Min	Max
Antidepressant sales	42.57	13.13	11.07	7.14	17.61	15.16	96.13
Mental disorders	94.23	26.43	22.73	13.65	16.54	22.16	219.08
Depression	20.54	7.32	5.02	5.35	6.12	2.89	56.07
Suicides	1.40	0.71	0.30	0.65	-0.19	0.00	6.68
Unemployment rate	2.63	1.16	1.07	0.47	-0.23	0.51	6.86
Below 15 share	15.84	1.83	1.56	0.97	-2.32	10.03	21.29
Between 15-65 share	67.22	1.77	1.68	0.56	-0.11	60.34	71.58
Over 65 share	16.93	2.43	2.26	0.90	2.43	10.52	24.56
Female share	50.48	0.85	0.81	0.26	-0.50	47.92	52.94
Foreigner share	18.56	7.16	7.03	1.49	3.69	3.65	40.95
German speaking share	64.30	37.25	37.42	0.01	-0.00	1.58	96.77
Specialists	2.30	2.34	2.28	0.59	1.08	0.00	14.98
General practitioners	6.65	1.52	1.38	0.66	0.22	2.63	11.99

Notes – The table reports descriptive statistics for the main variables. The statistics are obtained using annual data at the small area level for the period from 2003 to 2014. Antidepressant use is measured in terms of defined daily doses per 1,000 inhabitants per day. Mental disorder and depression are expressed in terms of hospital admissions per 10,000 inhabitants. Specialists and general practitioners are measured by the density per 10,000 inhabitants.

Table 2: Estimates of the effect of antidepressant sales on mental health outcomes

Outcomes (ln):	Mental disorder		Depression		Suicide	
Model	(1)	(2)	(1)	(2)	(1)	(2)
FE	.002	.002	.012**	.011**	-.001	-.001
	(.002)	(.002)	(.005)	(.004)	(.004)	(.002)
2SLS	.021**	.022**	.065***	.065***	.011	.010
	(.009)	(.010)	(.023)	(.024)	(.008)	(.008)
1st stage	.112***	.108***	.112***	.108***	.112***	.108***
	(.031)	(.033)	(.031)	(.033)	(.031)	(.033)
Reduced form	.002**	.002**	.007***	.007***	.001	.001
	(.001)	(.001)	(.002)	(.002)	(.001)	(.001)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Physician density	No	Yes	No	Yes	No	Yes
Unemployment rate	No	Yes	No	Yes	No	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	12.690	10.639	12.690	10.639	12.690	10.639

Notes – The table reports the parameter estimates for each health outcome (mental disorder, depression, and suicide) for the linear regression model and the instrumental variable regression model. We control for year and region fixed effects and population characteristics in the specification (1), and include additional control variables in the specification (2). We use two-way clustered robust standard errors at region and year 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	-0.279	-0.035	0.179	-0.027
Positive	1.279	0.160	0.821	0.093

Panel B: Correlations between manufacturers aggregates

	α_m	g_m	β_k	F_m	Var(z_m)
α_m	1				
g_m	0.547	1			
β_m	-0.063	0.121	1		
F_m	0.322	-0.243	-0.396	1	
Var(z_m)	0.419	0.054	0.274	0.337	1

Panel C: Top 5 Rotemberg weights by manufacturer

	$\hat{\alpha}_m$	g_m	$\hat{\beta}_m$	95 % CI	Market Share
Mepha-teva	0.706	10.914	0.084	(0.05;0.24)	4.110
MSD	0.240	1.148	0.051	(0.01;0.16)	22.329
Sanofi-Aventis	0.167	-3.423	0.086	(0.06;0.14)	11.708
Vifor Pharma	0.073	-1.866	0.009	(-0.03;0.03)	4.550
Sandoz	0.065	6.709	0.104	(0.04;0.97)	1.861

Notes – This table reports statistics about the Rotemberg weights aggregated at manufacturer 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 national manufacturer 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 manufacturer share (F_m), and the variation in the manufacturer shares across locations (Var(z_m)). Panel C shows the Top 5 manufacturers according to the Rotemberg weights.

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

Outcomes (ln):	Mental disorder		Depression		Suicide	
Model	(1)	(2)	(1)	(2)	(1)	(2)
2SLS	.021***	.021***	.067***	.066***	.003	.001
	(.008)	(.008)	(.020)	(.020)	(.006)	(.006)
1st stage	.357***	.354***	.357***	.354***	.357***	.354***
	(.062)	(.062)	(.062)	(.062)	(.062)	(.062)
Reduced form	.007***	.008***	.024***	.023***	.001	.000
	(.002)	(.002)	(.005)	(.005)	(.002)	(.002)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Physician density	No	Yes	No	Yes	No	Yes
Unemployment rate	No	Yes	No	Yes	No	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	33.603	32.625	33.603	32.625	33.603	32.625

Notes – The table reports the parameter estimates for each health outcome (mental disorder, depression, and suicide) for the linear regression model and the instrumental variable regression model. We control for year and region fixed effects and population characteristics in the specification (1), and include additional control variables in the specification (2). We use two-way clustered robust standard errors at region and year level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Happy pills? Mental health effects of the dramatic increase of antidepressant use

Supplementary Materials

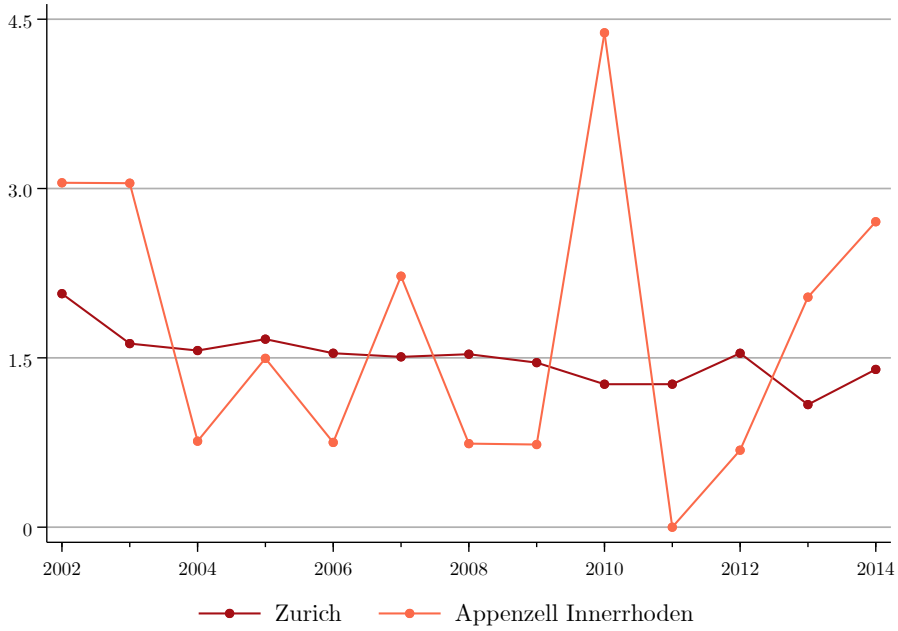
Giuliano Masiero, Fabirizio Mazzona and Sandro Steinbach

Appendix A: Data construction

We rely on the small area 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 region 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 region level using population weights.

Appendix B: Figures and tables

(a) Depression



(b) Suicide

Figure A.1: Variability of mental health outcomes (normalized using the rate per 10,000 inhabitants) for Zurich and Appenzell Innerrhoden

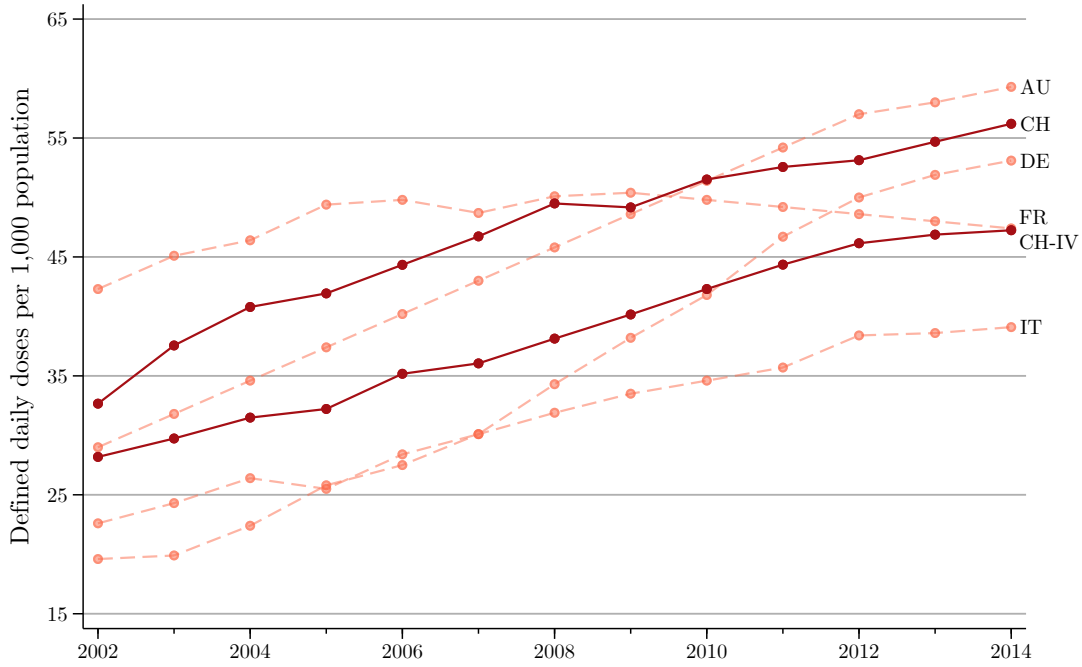


Figure A.2: Practice spillover instrument and changes in antidepressant use over time

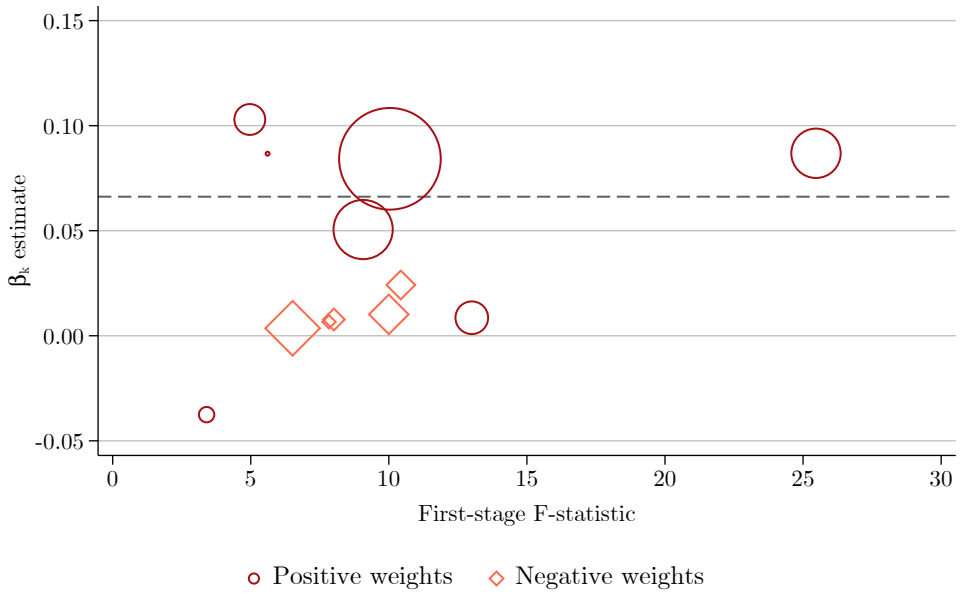


Figure A.3: Heterogeneity of β_m

Notes – The figure plots the relationship between the estimated β_m (on hospitalizations for depression) and their first-stage F-statistics. Each point represents a separate instrument estimate. While the size of each point represents the magnitude of the associated Rotemberg weights, the circle denote positive weights while the diamonds negative weights. The horizontal dashed line denote the overall β estimated in the main text by our IV strategy.

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 the 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: 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 manufacturer and year. We do not include the introduction of a new mode of drug administration or package size.

Table A.3: 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.3 – 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 manufacturer and year. First introducers are highlighted in italic.

Table A.4: Estimates of the effect of antidepressant sales on mental health outcomes by sex and age groups

Outcomes (ln):	Mental disorder	Depression	Suicide
		Males	
2SLS	.018*	.063**	.004
	(.010)	(.027)	(.013)
		Females	
2SLS	.023**	.066***	.057***
	(.009)	(.021)	(.017)
		Age < 20	
2SLS	.015	.080**	.024
	(.015)	(.038)	(.037)
		Age $20 - 65$	
2SLS	.024**	.066**	.023***
	(.010)	(.027)	(.008)
		Age > 65	
2SLS	.017	.055***	-.057
	(.011)	(.015)	(.043)
Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Demographics	Yes	Yes	Yes
Observations	1,272	1,272	1,272
Kleijbergen-Paap F	11.99	11.99	11.99

Notes – The table reports the second stage IV estimates for each health outcome (mental disorder, depression, and suicide) separated by gender and age. We control for year and region fixed effects and population characteristics. We use two-way clustered robust standard errors at region and year level. Significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.5: Descriptive statistics of practice spillover instrument

Country	Mean	SD	Δ(2003/14)	Min	Max
Switzerland	46.98	7.13	18.64	32.66	56.20
Switzerland IV	38.31	6.74	17.51	28.19	47.24
Austria	45.41	10.32	27.50	29.00	59.30
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 for the practice spillover instrument. The statistics are obtained using annual data at the small area level for the period from 2003 to 2014. Antidepressant use is measured in terms of defined daily doses per 1,000 inhabitants per day.

Table A.6: Placebo estimates using alternative hospitalization outcomes

Outcomes (ln):	Emergency Hospitalization excl. mental health	Elective Hospitalization excl. mental health	Infectious diseases	Bone fractures	Pregnancy and childbirth
2SLS	.009	-.003	-.008	-.006	.003
	(.006)	(.011)	(.008)	(.008)	(.005)
Year FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes
Physician Density	Yes	Yes	Yes	Yes	Yes
Observations	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	12.69	12.69	12.69	12.69	12.69
Mean	924.76	637.96	35.55	150.97	111.36
Within SD	69.48	72.78	9.41	23.34	11.28
Between SD	100.50	103.11	6.97	27.91	13.07

Notes – The table reports the second-stage IV estimates for the placebo outcomes. We control for year and region fixed effects and population characteristics in the regression models. The placebo outcomes (ICD10 codes) are neoplasms (C00-C97 & D00-D09 & D10-D36 & D37-D48), certain infectious and parasitic diseases (A-B), bone fractures (S), sexually transmitted diseases (A50-A64), and pregnancy, childbirth and the puerperium (O). We use two-way clustered robust standard errors at region and year level. Significance at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table A.7: Poisson estimates of the effect of antidepressant sales on mental health outcomes

	Mental disorder		Depression		Suicide	
Model:	(1)	(2)	(1)	(2)	(1)	(2)
Poisson	0.004**	0.004**	0.014***	0.014***	0.002***	0.002***
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)
Second stage Poisson	0.033**	0.020**	0.074***	0.059***	0.021***	0.011***
	(0.014)	(0.008)	(0.020)	(0.016)	(0.016)	(0.010)
Second stage Residual	-0.031**	-0.021**	-0.064***	-0.058***	-0.020***	-0.011***
	(0.014)	(0.009)	(0.020)	(0.016)	(0.017)	(0.011)
First stage Instrument	0.104***	0.595***	0.104***	0.595***	0.104***	0.595***
	(0.032)	(0.087)	(0.032)	(0.087)	(0.032)	(0.087)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Physician Density	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,272	1,272	1,272	1,272	1,272	1,272
Kleijbergen-Paap F	8.42	17.40	8.42	17.40	8.42	17.40

Notes – The table reports the parameter estimates for each mental health outcome (mental disorder, depression, and suicide) for the Poisson model and the Poisson instrumental variable regression model. Each regression controls for year and region fixed effects. We also include covariates for demographics and physician density as in our preferred regression specification. We report estimates for the shift-share instrument in the specification (1) and the practice spillovers instrument in the specification (2). We use two-way clustered robust standard errors at the region and year level. The cluster standard errors are bootstrapped with 1,000 replications and replacement. Significance levels at 10%, 5%, and 1% indicated by *, **, and ***, respectively.