Medical Marijuana Laws and Mental Health in the United States

Jörg Kalbfuß*, Reto Odermatt† and Alois Stutzer‡

February 13, 2019

Abstract

The consequences of legal access to medical marijuana for individual welfare are a matter of controversy. We contribute to the ongoing discussion by evaluating the impact of the staggered introduction and extension of medical marijuana laws across US states on self-reported mental health. Our main analysis is based on BRFSS survey data from more than six million respondents between 1993 and 2015. We find that medical marijuana laws lead to a reduction in the self-reported number of days with mental health problems. The reductions are largest for individuals with high propensities to consume marijuana for medical purposes and likely pain sufferers.

Keywords: medical marijuana laws, cannabis regulation, mental health, chronic pain, prescription drug monitoring

JEL classification: H75, I12, I18, I31, K42

*University of Cambridge and University of Basel, Center for Research in Economics and Well-Being, Peter Merian-Weg 6, 4002 Basel, Switzerland. Email: jk704@cam.ac.uk.

†University of Chicago Booth School of Business and University of Basel, Center for Research in Economics and Well-Being. Email: reto.odermatt@unibas.ch.

‡University of Basel, Faculty of Business and Economics and Center for Research in Economics and Well-Being, Peter Merian-Weg 6, 4002 Basel, Switzerland. Phone: +41 (0)61 207 33 61, email: alois.stutzer@unibas.ch.

1 Introduction

The legal status of cannabis has become successively more liberal in many countries in recent years. In the United States, 33 states had eased access to cannabis via decriminalization, medical programs or recreational allowances by 2019. Yet, the new laws remain contentious.[1] Major controversies revolve around the long-term consequences of cannabis consumption. To date, these consequences are not well understood, partly because strict regulations in the past also encumbered medical research. Besides disagreement regarding the therapeutic value of marijuana, there is also no consensus about the potential negative externalities as well as internalities due to addiction. The medical marijuana movement is thus concurrently understood as an attempt to bring back marijuana for therapeutic purposes to help people with chronic cancer pain, spasticity, nausea, or loss of appetite, and as a Trojan horse for the legalization of recreational abuse (Kilmer and MacCoun 2017).

We contribute to this discussion with a comprehensive evaluation of the effect of US medical marijuana laws (MML) on self-reported mental health. Our metric tries to capture welfare differences due to variation in regulations in a preferably encompassing way.[2] To identify the policy effects on people's mental well-being, we exploit the staggered introduction of all MMLs in the United States until the end of 2015. The basis for our analysis is repeated cross-sectional data from the Behavioral Risk Factor Surveillance System (BRFSS) from 1993 until the end of 2015, covering all US states and the District of Columbia. The data comprise a total of around 6.3 million observations. We address the concern of potential endogeneity of the introduction of MMLs in various ways. First, we make treatment and control states more comparable by dropping states which never introduced any MML. Second, we include dummies capturing lead effects of the policy introductions. Third, we test the robustness of the results based on placebo tests with randomly chosen introduction dates for the MMLs. Lastly, we apply a regression-adjusted matching estimator to study effect heterogeneity. Beside state- and time-fixed effects, our empirical strategy considers state-level institutional information on beer taxes, cigarette taxes, Medicaid expenditures, politic control, and minimum wages to control for possibly confounding institutional variation across states and over time.

Additionally, we make use of the National Survey on Drug Use and Health (NSDUH), which provides information on individual cannabis consumption behavior. We use this information in a triple difference approach to study how people are differently affected who are likely to consume under a MML regime for medical or recreational purposes. This strategy has not been used, as far as we are aware, in previous studies to deal with potential time-*variant* factors that affect the outcome variables at the state level and are simultaneously correlated with the introduction of MMLs. The triple difference interpretation thus allows us to identify

1. This holds, for example, for the most important control factors (Caulkins et al. 2012) and even for the goals of the law (Room 2014, Richter and Levy 2014).

2. A similar approach has been applied, for example, by Gruber and Mullainathan (2005) and Odermatt and Stutzer (2015) to evaluate tobacco control policies based on reported subjective well-being.

lower bound estimates under rather weak assumptions about confounding factors. The NSDUH also enables us to study the effect on people who likely suffer from chronic pain.

Overall, we observe a reduction in the number of bad mental health days with the adoption of a MML. In the analysis of the impact of MMLs on different subgroups, we do not find any evidence that presumed risk groups, such as young adults, are negatively affected by liberal regimes. We find the most pronounced beneficial effects in terms of reduced mental health problems for likely pain sufferers and medical marijuana consumers. The effect size for the latter group amounts to approximately 0.7 fewer days with self-reported bad mental health per month, where the group mean amounts to approximately 6.4 days. Studying the heterogeneity in MMLs, we find that the improvements primarily arise in states that have the most liberal medical marijuana regimes (specifically, those that allow the prescription of cannabis for unspecific pain), as was the case in California. For less liberal regulations (adopted mainly in later periods), we do not find any systematic effects, on average. Moreover, an event study shows that the benefits of the MMLs seem most prevalent in the first three years after their adoption. In addition, we find spatial spillover effects, i.e., a reduction in mental health problems owing to the introduction of MMLs in neighboring states, as well as smaller treatment effects if neighboring states already have a MML in place.

In a supplementary analysis, we study the interaction of MMLs with Prescription Drug Monitoring Programs (PDMPs). These are state-level institutions that require physicians to check their patients' medical histories when prescribing them potentially addictive drugs. Our analysis shows that mental health is only positively affected by PDMPs if they are introduced in a state that has an effective MML. This suggests that MMLs might be seen as complements to PDMPs – patients who formerly abused prescription drugs such as opioids, but are now kept from doing so, might use cannabis as an easily accessible but regulated substitute.

The remainder of this paper is structured as follows. In Section 2, we summarize potential consequences of MMLs and discuss the previous literature in more detail. In Section 3, we describe the institutional data on the regulation of medical marijuana as well as the individual data on mental health, cannabis consumption and pain. In the same section, we also present and qualify our empirical strategy. In Sections 4, 5, and 6, we present our results. Section 7 offers concluding remarks.

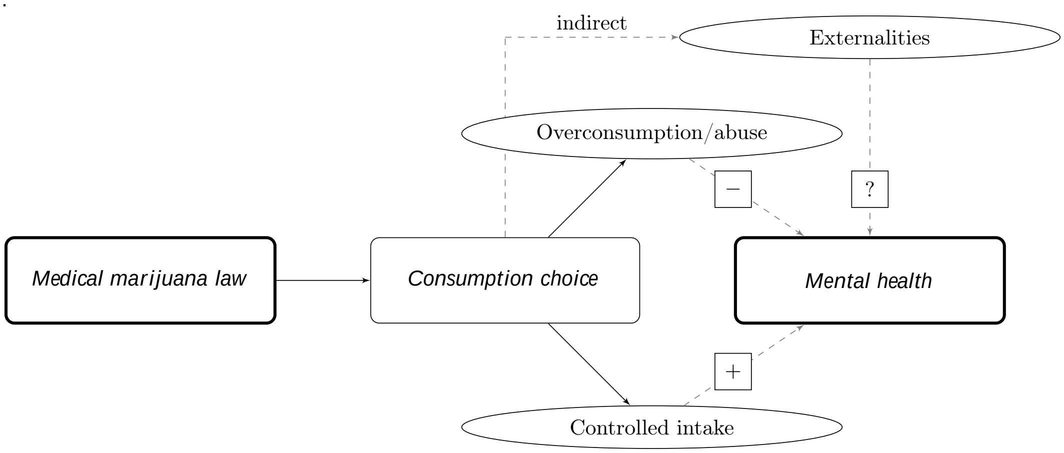
2 Potential consequences of medical marijuana laws

The introduction of a MML might affect people's mental health through various channels. Many effects are probably mediated by the impact on consumption behavior.[3] However, there

3. A short discussion of the environmental and individual-level covariates of consumption can be found in Appendix C.1.

are also potential general equilibrium effects related, for example, to a redeployment of police forces. We distinguish between the consequences of MMLs on the controlled consumption of cannabis for medical or recreational purposes, and uncontrolled consumption, i.e., the overconsumption or abuse of cannabis. It is this trade-off between the potential beneficial value of marijuana as a new therapeutic option and the detrimental risk of overconsumption or abuse which also stands in the focus of the public debate. We furthermore consider the possible side-effects on non-consumers in the form of externalities. Figure 1 depicts these links.

Figure (1) - Representation of potential channels by which MMLs might influence mental health.



There are several studies that report the therapeutic value of marijuana under controlled consumption. Meta-analyses show analgesic and other medicinal benefits of cannabis compared to placebo treatments (Martin-Sanchez et al. 2009, Iskedjian et al. 2007, Whiting et al. 2015). In contrast, the risks associated with marijuana consumption are less clear. Examples of potential harmful effects are neurological decline (Meier et al. 2012), cardiovascular diseases (Hall and Degenhardt 2009) and schizophrenia (Semple, McIntosh, and Lawrie 2005). However, the verdict on whether facilitated access to medical marijuana is deemed predominantly beneficial or detrimental will depend, ceteris paribus, on its comparative advantages over alternative treatments. In the context of chronic pain, these would be, for example, the opioid-based drugs codeine or OxyContin. In view of the well-documented side-effects caused by opioids, controlled cannabis intake for medical purposes can be seen as an efficacious alternative to established analgesics. This is in line with studies reporting lower opioid prescriptions (Ozluk 2017) and lower opioid-related fatalities after the introduction of MMLs (Bachhuber et al. 2014, Powell, Pacula, and Jacobson 2018). Moreover, controlled intake might be used to cope with stressful life events, decreasing the prevalence of suicide (Anderson, Rees, and Sabia 2014).

MMLs facilitate access to cannabis not only for medical use, but also for recreational consumption (Jacobi and Sovinsky 2016). Any welfare effects of potential diversion are hard to judge. They depend on the consumption value of cannabis to consumers, the concomitant risk of non-rational dependency and the degree to which the diverted marijuana is consumed as a complement or a substitute to other substances. Wen, Hockenberry, and Cummings (2015) report that the implementation of MMLs lead to an increase in the probability of past-month marijuana use, regular marijuana use, and dependence among adults aged 21 or above. With regard to adolescents – who are often claimed to be a major risk group – they observe an increase in marijuana use initiation, but not a higher probability of dependence. Pacula et al. 2015 further find that states which legally protect dispensaries face increased recreational marijuana use and dependence for both adults and youth. Moreover, MML exposure tends to reduce the highschool graduation rate (Plunk et al. 2016), expected labour earnings for young males (Sabia and Nguyen 2018), and academic performance (Marie and Zölitz 2017), particularly in the case of comparatively weak students. In contrast, Wall et al. (2016) find no systematic change in marijuana use in response to the introduction of MMLs among youths and in a systematic review and meta-analysis, Sarvet et al. (2018) come to the same conclusion. Regarding potential benefits, Sabia, Swigert, and Young (2017) find that states which adopted a MML exhibit a lower prevalence of obesity among the young as well as increased physical mobility among the elderly.

With regard to externalities, the literature reports a multitude of effects. Examples of these effects are decreased absenteeism from work (Ullman 2017), accidental ingestion of cannabis by young children (Wang, Roosevelt, and Heard 2013), a negative environmental impact of local cultivation (Carah et al. 2015), or additional tax revenues and a decrease in crime related to drug trafficking (Gavrilova, Kamada, and Zoutman 2018). Furthermore, there are several studies that report systematic relationships between MMLs and the rate of traffic fatalities. For example, Anderson, Hansen, and Rees (2013) and Reiman (2009) find substantial decreases in traffic fatalities, due to a conjectured substitutional relationship between marijuana and alcohol. This is in line with Baggio, Chong, and Kwon (2018) who find a negative impact of MMLs on alcohol sales based on scanner data. Smart (2015) reports asymmetric impacts on traffic fatalities conditional on age: While the young cause more drug-related traffic accidents, the reverse holds for older adults who drink less alcohol due to the availability of medical cannabis. This suggests that omitted treatment response heterogeneity could be one of the causes of the mixed findings in the literature.

Given the various ways in which easier access to cannabis might affect individual welfare, the net welfare effects of MMLs are difficult to identify. In particular, observable consumption behavior is insufficient to evaluate policies as it does not consider potential negative externalities and internalities as well as network effects such as the impact on the consumption value of cannabis for peers. In the light of this reasoning, we strive towards an analysis of net welfare effects using mental health as a proxy measure.

3 Data description and empirical strategy

3.1 Medical marijuana regulations in the United States

The regulation of medical marijuana differs widely across US states and ranges from laws that provide only minimal access to laws that permit an almost unrestricted supply of cannabis for medical as well as recreational use. While marijuana was effectively illegal before 1996 in all states, California pioneered the United States' first MML in November 1996.[4] By December 31, 2015, 23 states had followed suit in liberalizing access to medical marijuana. Figure 2 presents a map of the United States showing the legislation on cannabis for each state, including Washington D.C., at the end of 2015: It shows whether an MML was in place, as well as whether recreational use and possession were legal. Furthermore, the figure indicates whether or not a state was entitled to impose a jail sentence for first-time consumption or small-scale possession of cannabis.[5]

Figure 3 shows the distribution of the changes in the regulation of marijuana over time. In total, there are 24 different dates on which MMLs were introduced, which we can exploit in our empirical analysis. Furthermore, eight states abolished the punishment of imposing a jail sentence on a first-time offender for cannabis consumption and small-scale possession during our sampling period. Regarding recreational use, however, we only observe five changes in legislation from 2012 onwards. While we include these two latter regime changes as control variables in our analysis, the estimates of their effect should be interpreted with caution due to the limited variation across time. As our treatment indicator, we consider the date when a MML became effective (rather than when it was passed). An overview of the respective dates can be found in Appendix G. In addition, we capture and classify law heterogeneity, such as different qualifying medical conditions that give patients legal access to medical marijuana. However, this is not a trivial task. Several taxonomies designed to capture relevant distinctions in the law and their timing have been proposed (Pacula, Boustead, and Hunt 2014, Chapman et al. 2016, Williams et al. 2016). We follow the practice of recent analyses and consider the legislation that protects individuals who possess cannabis for medical purposes from prosecution, the authorization of home cultivation, the presence of operational dispensaries, as well as unspecific pain as a qualifying condition. The rationale behind these choices will be explained in Section A.2, when estimates for the effects of the different policy dimensions are discussed.

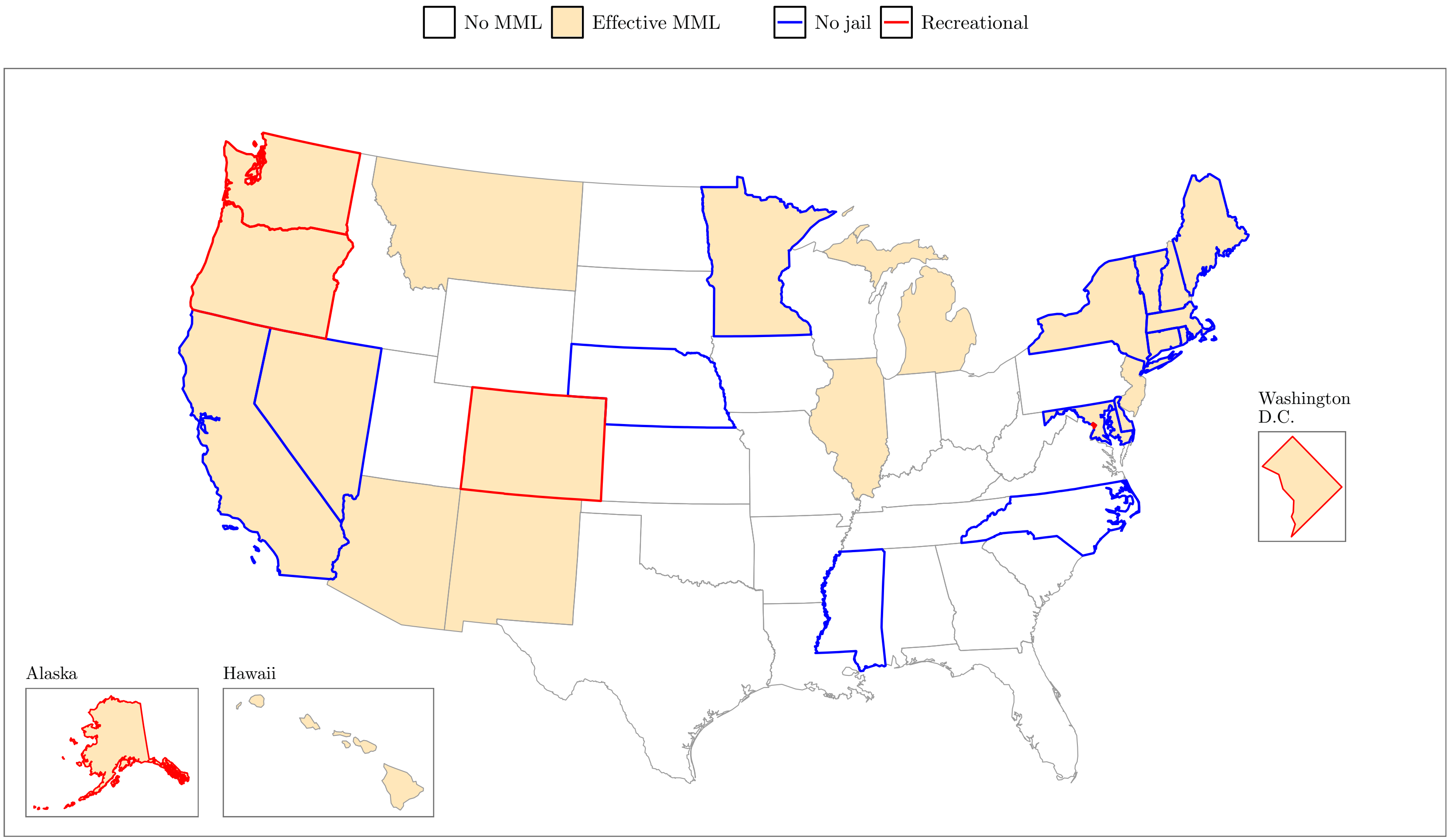
3.2 Individual-level data

Our main data source is the Behavioral Risk Factor Surveillance System (BRFSS). It consists of repeated cross-sections of telephone surveys which target US residents above the age of

4. We thereby ignore minor concessions, such as the Alaska law case *Ravin v. State* in 1975 which declared that small possessions of marijuana at home would be protected by privacy laws.

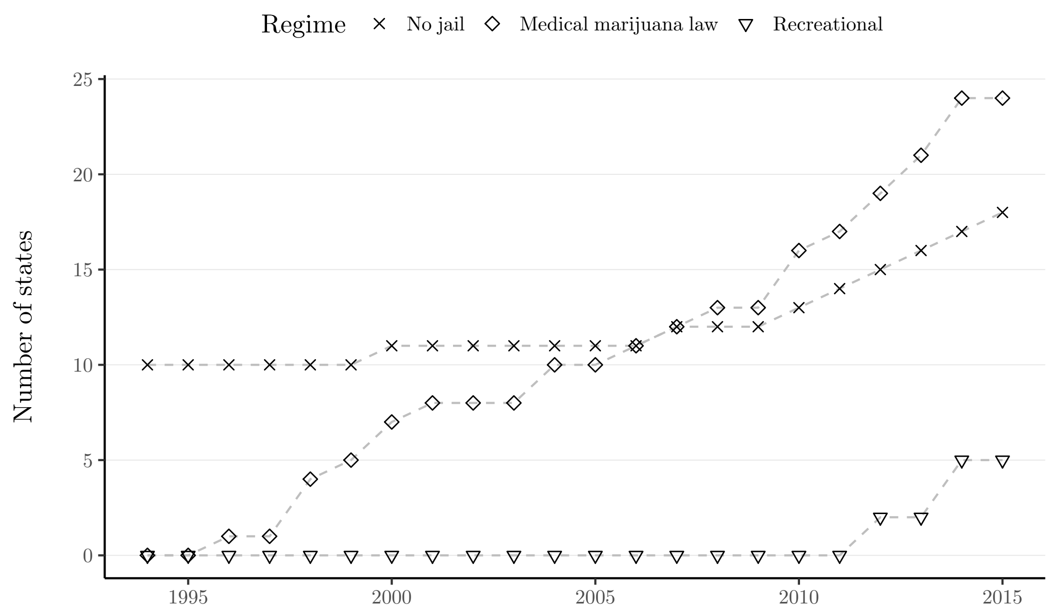
5. One visible pattern is the geographical clustering of laws. States in the west and northeast of America seem much more likely to adopt medical marijuana laws than the states in other regions of the country. In particular, the conservative so-called "Bible Belt" region in the south-east appears to be reluctant to liberalize or decriminalize marijuana in any form.

Figure (2) - Regulation of (medical) marijuana across the US states at the end of 2015.



Notes: No jail indicated by a blue border shows whether first-time consumption and small-scale possession of cannabis in violation of the law are punishable by a jail sentence or not. *Recreational* shows whether recreational use and possession is legal in the respective state. For a comprehensive table of legislation introduction dates, see Appendix B. *Data source:* Own compilation.

Figure (3) - Timeline of cannabis regime adoptions in US states.



Data source: Own compilation.

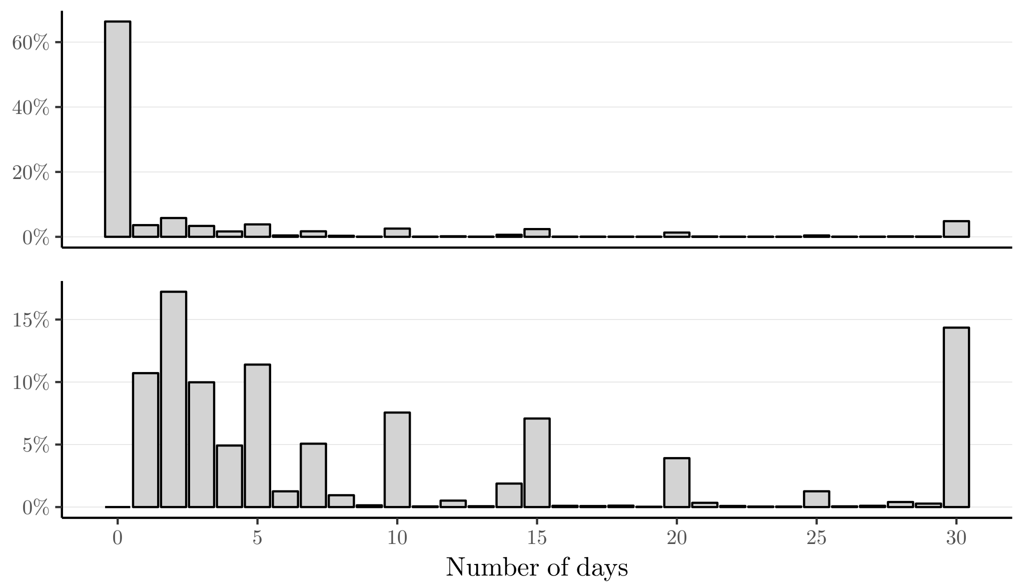
18. In every year, the respondents answer the following question regarding their mental state of health: "Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?". In our analysis, we focus on the answers to this question as our primary outcome variable. The relevance of this measure is shown by various studies which report that self-reported mental health is, for example, a good predictor of help-seeking behavior (Hunt and Eisenberg 2010), suicide (Bramness et al. 2010) or psychological functioning and mortality (Lee 2000). This metric is available for almost all individuals in all states and years with an item non-response of approximately 2%.[6]

Looking at the sample distribution of bad mental health days in Figure 4, the strong spike at zero indicates that almost 70 percent of the respondents did not experience bad mental health on any of the days during the previous month at the time of their interview. In the bottom plot of the figure, we see that, conditional on having at least one bad mental health day, a majority of people report experiencing between one and five days with bad mental health during the previous month. Furthermore, in higher ranges of the distribution, we observe bunching at five-point intervals. Figure 5 shows the evolution of these extensive and intensive components of average reported bad mental health days over the last 23 years. Diverging trends for these components are apparent in general. While the population reports

6. For a more general discussion of self-reported health and well-being measures in policy evaluations, see, e.g., Dolan, Layard, and Metcalfe (2011) and Odermatt and Stutzer (2018). A possible objection to our main outcome variable is the risk of simultaneity. People with mental health problems might want to self-medicate using cannabis, and therefore advocate MMLs or sort into states which have such a regime in place. However, medical research does not support this critique (Harris and Edlund 2005, Van Ours et al. 2013).

Figure (4) - Distribution of the number of bad mental health days during the last 30 days.

Bottom plot: Conditional on at least one bad mental health day



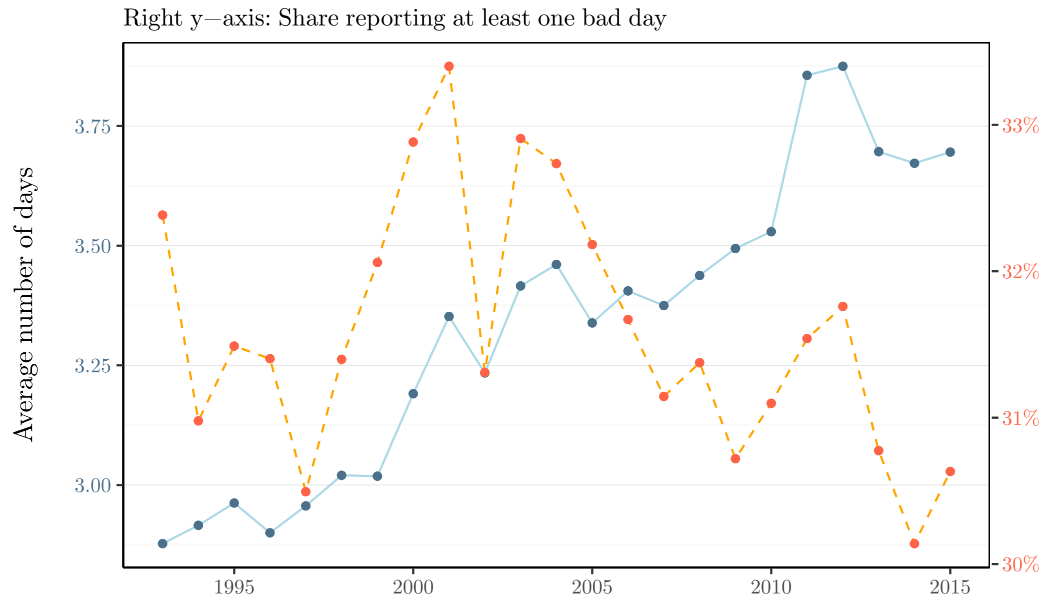
Data source: BRFSS. Calculated using survey weights.

more days with mental health problems over time, on average this increase does not seem to be driven by an increase in the share of afflicted individuals. Instead, conditional on having problems, the number of people's reported bad mental health days grew progressively over the last two decades. This suggests that distributional effects should also be considered in a policy evaluation. We analyze such effects in Section 4.1.[7]

For our analysis, it would be valuable to know about individual cannabis consumption behavior. However, in the BRFSS this question is only available from 2014 onwards. In order to study the policies' potentially heterogeneous effects conditional on individual propensities to consume (medical) marijuana, we make use of a second major data source, namely the National Survey on Drug Use and Health (NSDUH). The NSDUH offers national data on the use and abuse of addictive drugs in the US population aged 12 and older. It is frequently used as the basis for estimating the national prevalence of and state trends regarding, for example, opioid dependence. We are primarily interested in two questions contained in the survey. First, in every wave respondents are asked to answer the question, "During the past 30 days, on how many days did you use marijuana or hashish?". Based on the answers to this question, Figure 6 shows that the general trend in marijuana consumption has increased since 1994. The picture is consistent with the successive liberalization and decriminalization of marijuana that we observe over time. However, the descriptive patterns cannot tell us to what extent

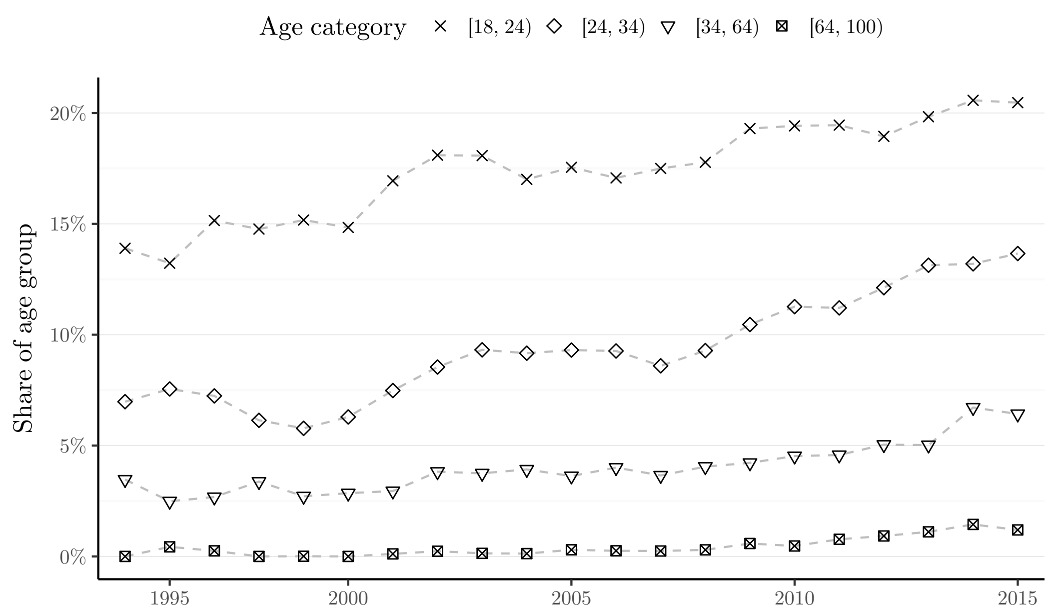
7. In 2011, the BRFFS landline interviews were complemented with mobile phone sampling. Additionally, a more sophisticated weighting method was introduced in compliance with the new sampling scheme, providing representativeness regarding additional socio-demographic variables (Centers for Disease Control and Prevention 2012).

Figure (5) - Time series of the extensive and intensive margin of self-reported bad mental health days in the United States.



Data source: BRFSS. Calculated using survey weights.

Figure (6) - United States national cannabis consumption rates over time by age group.



Note: An observation is classified as a consumer if the respondent used cannabis on at least five days within the last 30 days. *Data source:* NSDUH. Calculated using survey weights.

the trends are driven by changes in the legal status of medical marijuana. Second, from 2013 onwards, the NSDUH also asks the survey participants whether some or all of their cannabis consumption is recommended by a doctor. Additionally, the information whether the respondent lives in a state with some form of legal medical cannabis program is provided.[8] These variables allow us to augment the BRFSS data with consumption propensity scores based on a predictive model for consumer status. Details of the procedure will be discussed in Section 5.1.

3.3 Empirical strategy

Most of the econometric analyses are based on the following estimation specification:

$$\begin{aligned} y_{ist} = {} & \beta\, \mathrm{mml}_{st} && \textit{treatment dummy} \\ & + \gamma Z_{st} + \omega X_{ist} && \textit{state \& individual controls} \\ & + \alpha_s + \theta_t + t\lambda_s && \textit{time \& state fixed e ects and trends} \\ & + \epsilon_{ist} && \textit{error (clustered at the state level)} \end{aligned}$$

The dependent variable y_{ist} is the self-reported number of bad mental health days in the last 30 days of individual i living in state s in year t.[9] Our primary explanatory variable of interest is mml_{st}, a binary treatment dummy indicating whether in state s at time t a MML is in place or not. The estimated parameter β, thereby, represents the average treatment effect (ATE) of the policy to the extent that our control strategy captures all observable and unobservable differences between treatment and control states. We use the exact interview and MML introduction dates to determine the treatment status for every observation on a daily basis. Furthermore, we capture potential lead effects of the policy introductions with dummies indicating observations one and two years before the respective policy introductions. In general, we would expect the lead coefficients to be either zero or of the same sign as the treatment effect if anticipatory effects occur (for example, due to changes in law enforcement).

In the light of the geographic clustering in Figure 2 and the results of Bradford and David Bradford (2017), who conclude that border diffusion and voter ideology are important drivers of MML adoption, a careful strategy to control for potential confounding factors is required. We address such issues by including various control variables. X_{ist} is a vector of variables at the individual level, controlling for differential socio-demographic compositions across states which might be correlated with the adoption of the policy. Specifically, we control for age, education, marital status, employment, income, the acquisition of a health plan, sex, race, and the number of children who live in the household (in categories capped at three). However,

8. Due to the new regulations of the institution SAMHDA, which publishes the NSDUH, access to state identifiers cannot be provided to researchers who are not resident in the US.

9. The scaling of our outcome variable has some technical implications. As it is a censored count variable, non-linear estimators such as Tobit or Poisson regressions might be considered. We refrain from applying them as there is no consensus regarding the handling of clustered errors in maximum-likelihood frameworks. Furthermore, the possibility of negative predictions for some sets of covariates does not affect the validity of our conclusions, since we are interested in net overall effects only.

state-level controls are arguably more important when striving for causal interpretations of the effect of the state-level policies. We therefore consider the vector Z_{st} of state variables including the unemployment rate, the beer tax, the cigarette tax, an indicator for urbanity on the county-level, the minimum wage, indicators for the parties holding political control in a state, as well as expenditures per capita for the Medicaid and the Temporary Assistance for Needy Families (TANF) programs. All monetary values are in real terms. We also take into account whether a state abolished jail sentences for first-time offenders for cannabis consumption and whether it legalized cannabis for recreational consumption. Additionally, we include interactions of MMLs with neighboring states' laws in the form of a dummy which equals one if at least in one adjacent state an effective MML is in place. A separate dummy captures whether a neighboring state allows cannabis for recreational consumption. Recent work by Hao and Cowan (2017) and Hansen, Miller, and Weber (2017) point towards the importance of such spatial controls. Descriptive statistics and sources for the respective variables can be found in Appendix B. Finally, we include state as well as time fixed effects and state-specific linear trends in order to control for some forms of unobserved heterogeneity across states.

4 Main results

4.1 Overall effects

Our specifications in Table 1 show the overall effect of an MML on bad mental health days based on different samples for the control states. The main variable of interest is the dummy variable "Overall MML" which captures the net effect for all years after the adoption of the law. Two further dummy variables capture the potential lead effects of the policy introductions.

The results in column (1) show a reduction of approximately 0.18 in the number of bad mental health days per month when a state introduces an MML. Hence, adults in MML states experience, on average, approximately two fewer bad mental health days per year due to the law. However, the effect is only statistically significantly different from zero by a small margin. In column (2), we allow MMLs to both spill over into neighboring states (row "Border MML") and to interact with the MMLs that these neighboring states might have in place (row "Border $\times$ MML"). By implementing this spatial control, the overall MML effect now indicates the effect of an MML if no neighboring state already operates an MML. In such a case, the effect size is slightly bigger and more precisely estimated. If a neighboring state has already implemented an MML, the introduction of a medical marijuana law still exhibits a positive effect on mental health on top of the pure neighborhood effect. Yet, the sum of the main law effect and the interaction with a bordering MML, amounting to -0.145, is not significant at the usual statistical levels.

Table (1) - Overall effect of medical marijuana laws (MML) on the number of days per month with bad mental health.

	All states		Any law	Effective law
	No. of days	No. of days	No. of days	No. of days
	(1)	(2)	(3)	(4)
Two years before MML	−0.106	−0.108	−0.111	−0.087
	(0.078)	(0.077)	(0.077)	(0.074)
One year before MML	−0.081	−0.087	−0.088	−0.073
	(0.113)	(0.110)	(0.109)	(0.093)
Overall MML	−0.182*	−0.227**	−0.230**	−0.211**
	(0.106)	(0.106)	(0.106)	(0.083)
Border × MML	–	0.082	0.087	0.138**
		(0.061)	(0.062)	(0.061)
Border MML	–	−0.074	−0.076	−0.172**
		(0.068)	(0.072)	(0.075)
Sample means	3.361	3.361	3.364	3.359
Time FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
State trends	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Clusters	51	51	45	29
Observations	6,349,173	6,349,173	5,540,433	3,631,697
Adjusted R^2	0.088	0.088	0.088	0.084

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: In column (3), only states are included which introduce or have in place some form of liberalized cannabis regulation during the considered time period. The sample in column (4) is further restricted to states which at some point introduced an effective medical marijuana measure as categorized in the Marijuana Policy Project (2016). Clustered standard errors are reported in parentheses. Due to few clusters, p-values in (3) and (4) refer to T(#Cluster -1) distributions (see Cameron, Gelbach, and Miller 2008). The row *sample means* reports the average bad mental health days (dependent variable) of the respective sample. *Data source:* BRFSS. Calculated using survey weights.

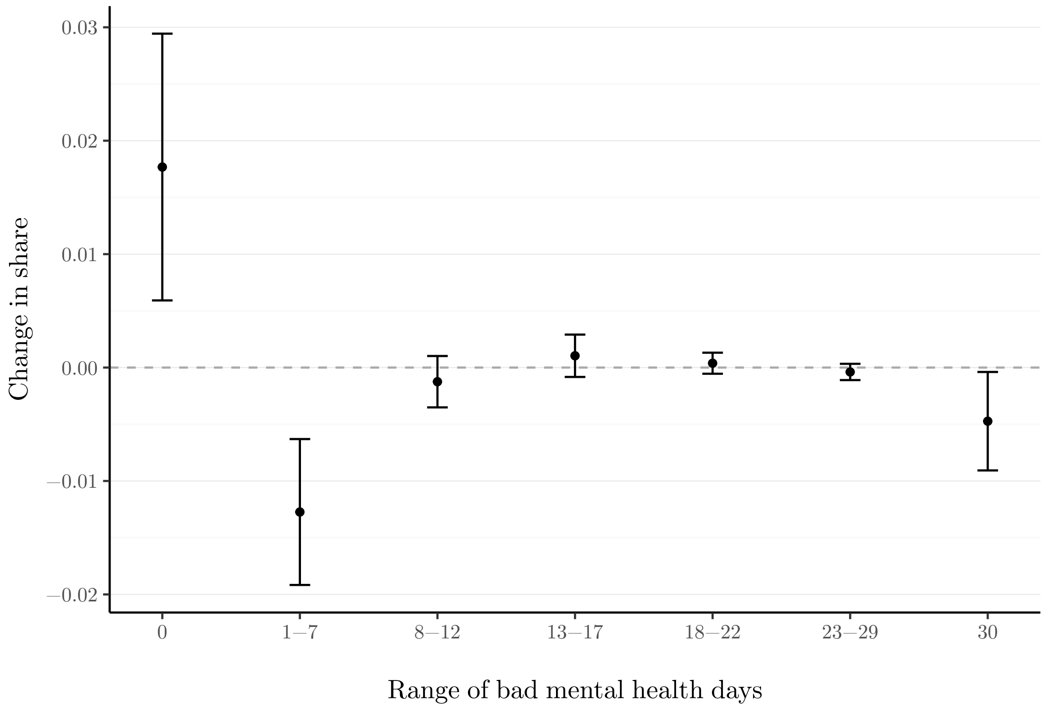
In columns (3) and (4), we consider the possibility that the states which do not change any of their cannabis regulations in our sample might form an inappropriate control group. For example, tight cannabis regulations could be related to unobserved characteristics of states, such as a puritan culture, which are systematically related to the impact of an MML. In order to make the states more comparable, we first restrict the sample to states which have adopted some form of regulation in column (3). Compared to column (2), no qualitative differences become visible. In column (4), we restrict the sample further by only considering those states that adopted an effective MML up until 2017, as classified in the report of the Marijuana Policy Project (2016). The changes in the main effect owing to the altered sample definition are, again, rather small. However, the spill-over effects of bordering MMLs are stronger. We find a significant positive spill-over effect of MMLs on mental health in neighboring states which have not yet introduced a MML themselves. Furthermore, MML introduction is observed to have a reduced effect in an environment where at least one neighboring state has an MML in place.

Figure 7 shows the distributional changes induced by MMLs regarding different initial levels of mental health. Technically, the plot reports the results of a conditional density estimation where certain ranges of bad mental health days are collected into bins. For each of the seven intervals shown on the x-axis of the figure, we first code a dummy which equals one if an individual's reported number of bad mental health days falls within it. We then use this dummy as the dependent variable in a linear probability model. The figure suggests that the improvements reported so far are driven by a general shift of the distribution towards fewer bad mental health days. Thereby, the biggest shift of mass is from the category reporting one to seven days into the one reporting none. The probability of falling into this latter category increases by almost two percentage points. Importantly, the probability also declines that somebody falls into the category at the upper end of the scale with people suffering a maximum number of days from bad mental health. The adoption of a MML thus seems not to lead to an overall polarization in mental health.

4.2 Robustness

We test the robustness of the results in Table 1 based on a placebo test that randomly allocates treatment dates. Specifically, for every column, we re-estimate the specification 5,000 times, whereby treatment dates are assigned at random to states in every run. The spatial relationships regarding bordering MMLs at specific times are adapted concomitantly. Figure 8 shows the densities of the calculated pseudo-effects for the four specifications. The continuous lines mark the estimated treatment effects from Table 1, whereas the dashed ones indicate the 5% quantiles of the respective pseudo-distributions. We find that every effect besides the one for specification (4) lies below the indicated quantile, indicating that they pass, roughly speaking, a one-sided t-test at the 5% significance level. The effect for column (4) only passes at the 10% level. In this specification, 22 states are dropped for the effect's estimation, which is associated with a considerable loss of precision. Considering

Figure (7) - Overall distributional effect of medical marijuana laws.



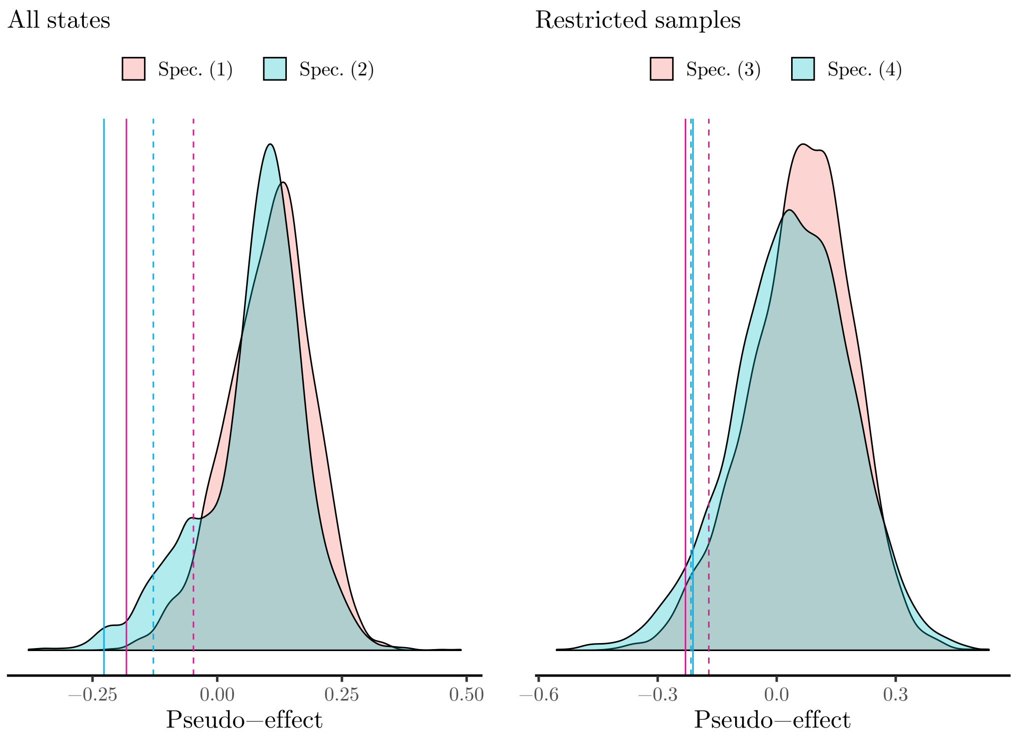
Notes: Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

the joint likelihood for the outcome over all four specifications, the figure indicates that our estimated treatment effects cannot simply be explained by spurious heterogeneous non-linear state-trend components.

4.3 Supplementary analyses

In a series of supplementary analyses, we explore many more refined aspects of heterogeneity in our data pool. However, given the number of legal changes, statistical power is limited and we often cannot draw strong conclusions regarding statistically significant differential effects from these analyses. Still, many results are suggestive of interesting effect heterogeneity along different dimensions. First, we study the heterogeneity of the reform effects across states and also look for influential states. The results are reported in Appendix A.1. They show that California seems to contribute substantially to the overall effect. We therefore study the effects for likely consumers and pain sufferers in Section 5 for the full sample as well as for a restricted sample excluding observations from California. Second, we investigate effect heterogeneity due to differences in the laws. This analysis is reported in Appendix A.2. It turns out that medical marijuana regimes that allow the recommendation of cannabis for unspecific pain (i.e., pain without a clearly diagnosed cause) are associated with the largest reductions in the number of bad mental health days. The large effects of these regimes might be driven by the patients' easier access to the drug. However, they could equally result from

Figure (8) - Distributions of pseudo-effects for columns (1)-(4) in Table 1.



Notes: Each specification has been re-estimated 5,000 times, whereby treatment dates are assigned at random at the day level in every run. Continuous lines refer to the treatment effects estimated in Table 1, whereas dashed lines indicate the 5% quantile of the pseudo-distribution. Note that the earliest date for pseudo-treatment assignment has been set to 01/06/1996. If earlier dates were allowed, states might drop out since not enough pre-treatment periods would be available to support the linear state-trend.

recreational use as this policy might come along with more opportunities for diversion. In Section 5.3, we provide some evidence that the first driver presumably dominates, which is based on an analysis of the population of likely pain sufferers. Third, we examine the dynamic effects of MMLs in a flexible way. The results are illustrated in Appendix A.3. Together with the other evidence so far, they indicate that MMLs are, on average, not harmful. If anything, our results suggest that they are distinctly beneficial, at least in the short- and medium-run. Fourth, we assess assertions that are prominent in the public debate about effect heterogeneity across demographic groups. The findings reported in Appendix A.4 reveal that the mental health of young women and men, and of white middle-aged men in particular, is not adversely affected in a systematic fashion by the introduction of a MML.

5 Effects for likely cannabis users and pain sufferers

5.1 Hypothetical propensity estimation

Since MMLs target patients who might benefit from the treatment option, we want to allow for differing effects of cannabis regulations for this group, i.e., pain sufferers and medical marijuana consumers. Another subsample of interest are recreational marijuana users. While these people would not be affected directly by the law, this latter group might still be influenced by diversion, cultural change or impacts on illicit supply. Formally, we do this by augmenting the baseline equation with dummies for the respective consumer groups, group specific linear time trends and interactions with MMLs, the no jail policies and recreational regimes, as well as neighbourhood effects. However, the consumer status is not reported in the BRFSS for the studied time horizon. In order to impute the missing information, we use an auxiliary data set from the NSDUH to predict consumption propensities for each observation in the BRFSS in a first step. These estimated propensities will then allow us to partition the sample into likely abstainers, recreational users and medical users. To derive the propensities, we use the following two questions asked in the NSDUH:

(1) "*During the past 30 days, on how many days did you use marijuana or hashish?*"[10]

(2) "*Was any of your marijuana use in the past 12 months recommended by a doctor or other health care professional?*"

Regarding the first question (which is available for the years 1994-2015), we first recode the variable as a dummy which equals one if consumption occurred on at least five days during the past month (i.e. marijuana was consumed on a weekly basis). We then fit a predictive model using only variables which are reported in both the BRFSS and the NSDUH. Beside basic socio-demographics and year effects, we include smoking status, the number

10. Some interviewees also reported the number of consumption days during the past year. We largely reproduce our results using this alternative metric, but at the cost of precision. We further abstract away from the *type* of marijuana that has been consumed. Potency, purity and the mix of strains such as *sativa*, *indica* or *ruderalis*, have a substantive impact on the effect of cannabis. For example, the connection between the incidence of psychosis and marijuana usage is highly dependent on the consumed mixture (Schubart et al. 2011, Di Forti et al. 2009, Mehmedic et al. 2010). However, our data does not allow us to differentiate along this dimension.

of days a person has consumed alcohol during the past thirty days and their Body Mass Index.

In order to derive the consumption propensities, we employ stochastic gradient boosting with decision trees as the base learners. This non-parametric boosting approach "learns" the functional form of the data generating process which predicts the outcome best according to some metric (see, e.g., Friedman 2002).[11] Our motivation for applying this specific method is three-fold. First, as our time period spans 23 years, cohort effects are likely to play a role – seniors in 1995 respond differently to an MML than seniors in 2015. Trying to incorporate such effects parametrically would either increase the number of coefficients exponentially (evoking the *curse of dimensionality*) or require arbitrary parameter restrictions. Second, this method is able to exploit the rich available variation on the individual-level to the fullest. Lastly, stochastic gradient boosted decision trees routinely head comprehensive machine learning rankings (Caruana and Niculescu-Mizil 2006, Caruana, Karampatziakis, and Yessenalina 2008). Performance diagnostics for our predictions can be found in Appendix C.2.

Using the model fitted on the NSDUH data in the first step, we then predict individual propensities to consume marijuana for medical and recreational reasons in the BRFSS in a second step. Here, we have to address the fact that MMLs plausibly induce selection effects, as cannabis regulations are likely to change the pool of users and the quantities people choose to consume. Hence, a simple regression on the factual propensities (i.e., the estimated likelihood to consume cannabis given the actual treatment status) will lead to biased estimates. For our evaluation of the effects of MMLs, we need to compare those people who consume under an MML with the respondents from the untreated states who *would* consume cannabis if an MML were in place. In our propensity regression, we thus replace the factual propensities for the control observations with the *counterfactual* ones, i.e., those propensities which we would observe if they were to live under an MML. We call the propensities derived from this replacement the *hypothetical propensities*. Furthermore, to equalize treated and untreated (potential) consumers regarding differential compositional time trends, we also enforce predictions to be made as if everyone lived in the year 2015. Figure 9 shows the distributions of both the factual and hypothetical propensities. The final classification of observations as either likely abstainers, recreational users or medical users was determined according to the procedure described in Appendix C.3.

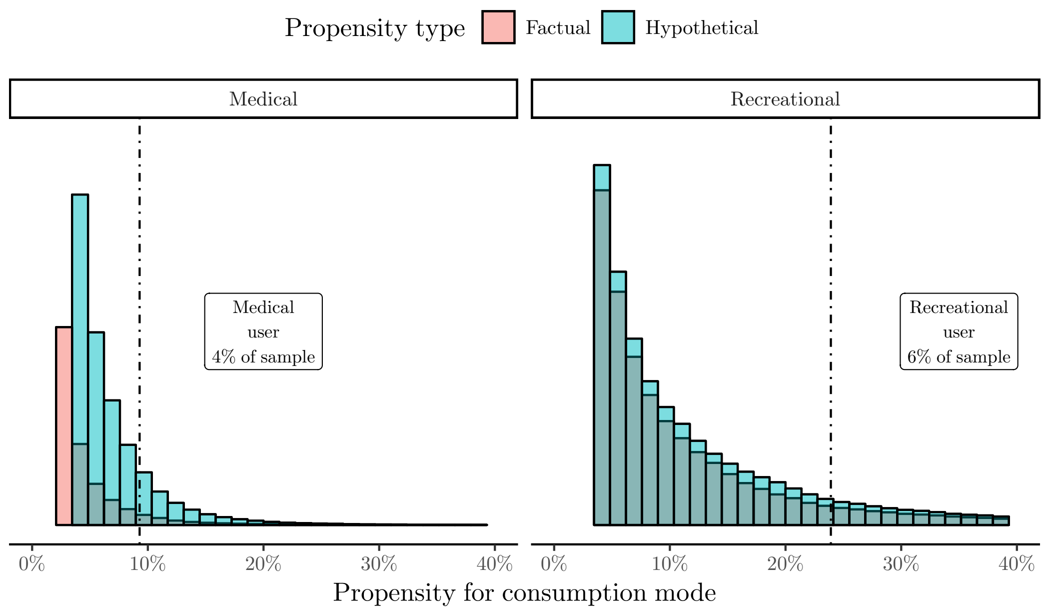
5.2 Estimates for likely recreational and medical marijuana users

In the following, we use the described hypothetical propensity estimates to analyze the differential MML effects for the subgroups with likely differing modes of consumption. Table 2 reports OLS estimates for the three groups of likely abstainers, recreational users and medical users.[12] We find big differences regarding the impact of MMLs on the three groups.

11. We use the implementation in the R package `xgboost` (Chen and Guestrin 2016).

12. The regressions which involve propensity scores require an adjustment of the standard errors, since we use an estimated explanatory variable which is itself subject to sampling variability. We use a two-stage

Figure (9) - Distribution of factual and hypothetical propensities for recreational and medical cannabis use.



Notes: For counterfactual predictions, the MML dummy is set to one for control observations and the year is fixed at 2015 for everyone. A respondent is classified as a recreational consumer if intake occurred on at least five days during the past 30 days. The thresholds are chosen so that they reproduce the hypothetical national prevalence of marijuana consumption in the United States. The x-axes are capped at 40% to improve the display. *Data source:* BRFSS. Calculated using survey weights.

For likely medical users, the effect size increases more than three-fold compared to previous results for the whole population. For both likely recreational users and likely abstainers, we observe comparable negative effects on bad mental health days. However, the effects for these latter groups will be qualified below based on a re-estimation of the same specification using a regression-adjusted matching estimator. Overall, the results in Table 2 indicate that the effects of MMLs differ widely across groups of the population. From a technical perspective, it appears to be important that the estimation approach allows for effect heterogeneity.

In a robustness check, we pursue an approach that goes one step further where we re-estimate the same specification using a regression-adjusted matching estimator.[13] This matching approach additionally enables us to estimate the average treatment effect on the treated (ATT) and the average treatment effect on the controls (ATC) in a straightforward way. Technically, the estimation proceeds in two steps following Ho et al. (2007). First, we esti-

bootstrapping approach to correct for this. We have generally set $N = 500$, based on an empirical investigation of convergence rates of standard errors.

13. Another motivation for this approach is Słoczyński (2017), who shows that both the exclusion of outliers regarding control variables and the flexible allowance for heterogeneous treatment responses conditional on covariates can have a substantial effect on the estimates. For example, given that very different people with diverse ailments might be classified as medical users, it is reasonable to check for the impact of allowing for such heterogeneity.

Table (2) - Effects of medical marijuana laws (MML) on bad mental health of likely abstainers, recreational users and medical users.

	No. of days		
	Abstainer	Recreational user	Medical user
Two years before MML	−0.096	−0.284	−0.145
	(0.080)	(0.296)	(0.338)
One year before MML	−0.090	0.204	−0.492*
	(0.122)	(0.242)	(0.273)
Overall MML	−0.215*	−0.200	−0.710***
	(0.113)	(0.201)	(0.243)
Border × MML	0.046	0.451***	0.810***
	(0.076)	(0.129)	(0.143)
Border MML	−0.068	−0.036	−0.399***
	(0.083)	(0.137)	(0.145)
No jail	0.133*	0.180	0.061
Sample means	3.158	5.01	6.358
Time FE	✓	–	–
State FE	✓	–	–
State trends	✓	–	–
Controls	✓	–	–
Clusters	51	–	–
Observations	6,029,498	202,351	117,324
Adj. R^2	0.091	–	–

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Columns (1)-(3) belong to the same regression. Treatment and neighbourhood effects are satiated regarding the propensity group; i.e., coefficients per column can be interpreted independently of the other columns belonging to the same regression. Propensity thresholds have been set to reproduce hypothetical NSDUH rates in the weighted sample. The row *sample means* reports the average number of bad mental health days (dependent variable) in the respective sample. Standard errors have been bootstrapped using 500 replications. *Data source:* BRFSS. Calculated using survey weights.

mate stabilized inverse probability weights which balance individual and some institutional covariates across the treated and the control observations (Austin and Stuart 2015). For the prediction of the treatment exposure, we employ stochastic gradient boosted decision trees (the same method that we used to predict consumption propensity scores). In the second step, we conduct an OLS estimation specified like the base equation and augmented with consumer status dummies. With this procedure, we allow treatment responses to vary over individual characteristics, the indicator for urbanity, Medicaid expenditures per capita, and the state unemployment rate, as well as all first-order interactions between these variables.[14]

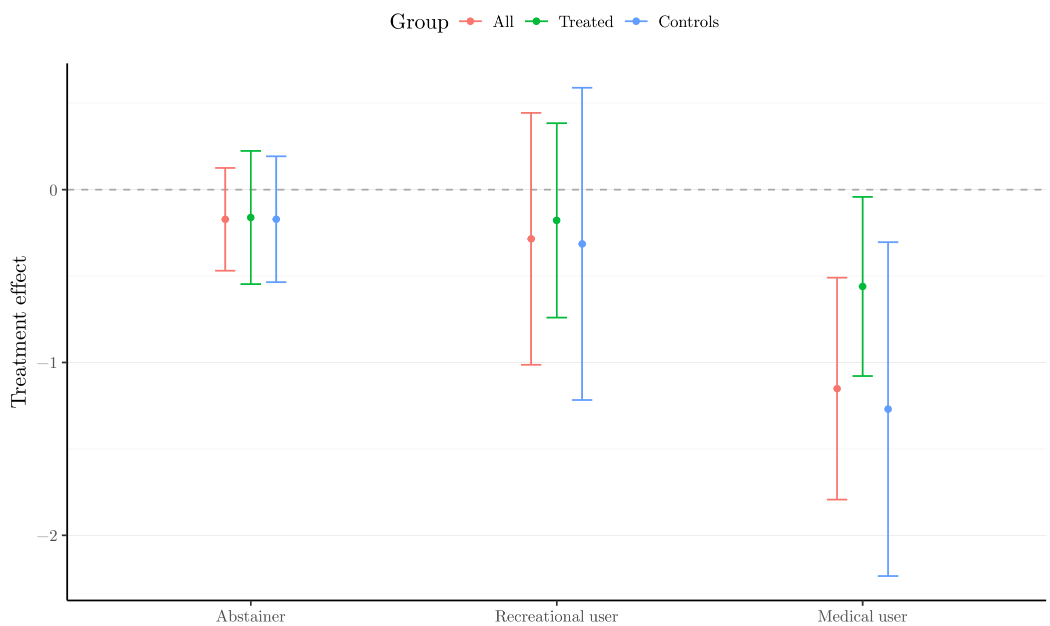
Figure 10 shows to what extent the ATE estimate changes with this more flexible specification, and additionally reports estimates for the average effect on the treated (ATT) and on the controls (ATC). It is revealed that the point estimates for both abstainers and recreational users remain largely unaffected, while the precision of the estimates diminishes substantially. Notably, the effect on abstainers cannot be distinguished from zero anymore. In contrast, the effect size of the average treatment estimate for medical users increases by approximately 50%. Thereby, the ATT estimate is approximately half the size of the ATE estimate, implying that effect heterogeneity in the group of medical users is substantial.[15][16] Overall, allowing for a flexible weighting of conditional treatment responses seems to be a sensible robustness check. Moreover, the results suggest that some of the mixed findings in the literature might be due to insufficient allowance for effect heterogeneity.

14. The reusage of control variables from the first stage would not be necessary if we were completely convinced that our data generating process concerning treatment exposure was correctly learned by our non-parametric estimator. Yet in addition to a further reduction of residual covariate imbalances which the boosting estimator might not have been able to eradicate, the re-estimation with all controls yields the so called "double-robustness" property. The latter ensures that only either one of the selection specification or the outcome specification needs to be correct for consistent estimation. This effectively shields our estimations against both potential over- or underfitting of the data regarding the true functional form of the data-generating process (Słoczyński and Wooldridge 2018).

15. In particular, the decomposition of the ATE reveals that the latter is disproportionately driven by an increased incidence of covariate combinations in control states which would have predicted a substantial benefit from the introduction of an MML in a treated state. This suggests that hypothetical users in states without access to medical marijuana would benefit even more than medical users in treated states. Such conclusions, however, need to be drawn with caution. First, the width of the confidence intervals indicates substantial uncertainty in the effect sizes. Second, interpreting the ATT and the ATC estimates as being causal requires the fulfillment of distinct assumptions. For the ATT, it requires that if a treated state had *not* been treated, expected mental health conditional on covariates would be the same as in the control states. In contrast, a causal interpretation of the ATC estimate requires that if a control state *had* been treated, expected conditional outcomes would coincide with those in a de facto treated state. Since the ATE is composed of both the ATT and the ATC, *both* assumptions are necessary to interpret the ATE estimate in a causal way. Since unobserved heterogeneity in control states might mediate the effect of an MML in ways which we cannot gauge in our data analysis, we suggest that the ATT estimate is the more reliable anchor for the expected effect of MMLs which might be introduced in the future.

16. This is consistent with the results of the same analysis excluding observations from California. While the ATT for medical users remains similar in magnitude, the ATE and the ATC do no longer differ from the ATT. The point estimates for the effects on the mental health of abstainers and recreational users are close to zero.

Figure (10) - Regression-adjusted matching estimates for the effect of medical marijuana laws on bad mental health, i.e., the average treatment effect (ATE), the average treatment effect on the treated (ATT) and the average treatment effect on the controls (ATC) for different cannabis consumer groups.



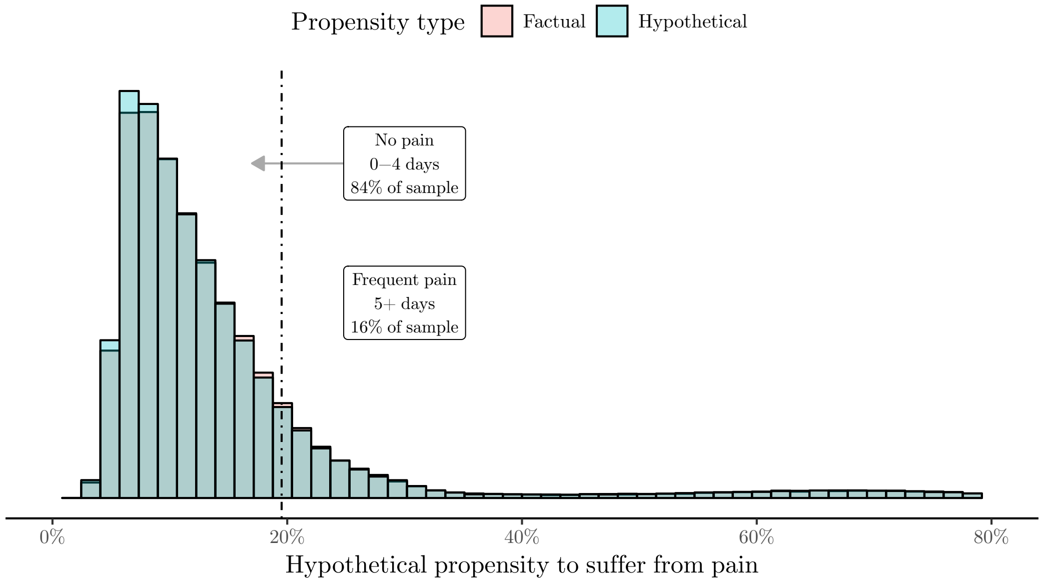
Notes: Standard errors have been obtained by bootstrapping using 500 replications. Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

5.3 Estimates for likely pain sufferers

As discussed in Section A.2, unspecific pain as a qualifying medical condition for the therapeutic use of marijuana is a strong predictor of the potential beneficial effects of MMLs on mental health. In order to gain further insights about this finding, we repeat the analysis that we used for different modes of consumption with respect to the propensity to suffer from frequent pain. We use the following question from the BRFSS to predict hypothetical pain propensities: "*During the past 30 days, for about how many days did pain make it hard for you to do your usual activities such as self-care, work or recreation?*" We categorize respondents who suffered for a minimum of five days as being frequent pain sufferers. Figure 11 shows the distributions of both the factual and the hypothetical propensities. Here, the hypothetical propensities are calculated for the situation in which *no* MML is in place. The group of likely pain sufferers thus reflects the ex ante situation including also those people who may no longer suffer from pain after a treatment with medical marijuana becomes legally available. The propensity threshold for a positive prediction was set at the 84% quantile to reproduce the prevalence of chronic pain in the United States in 2012 (Nahin 2015).
Figure 12 summarizes the results for the matching approach.[17] We first discuss the ATE estimates which are represented by the red elements in the graph. The effect of an MML

17. The corresponding OLS results can be found in Appendix B, Table 9.

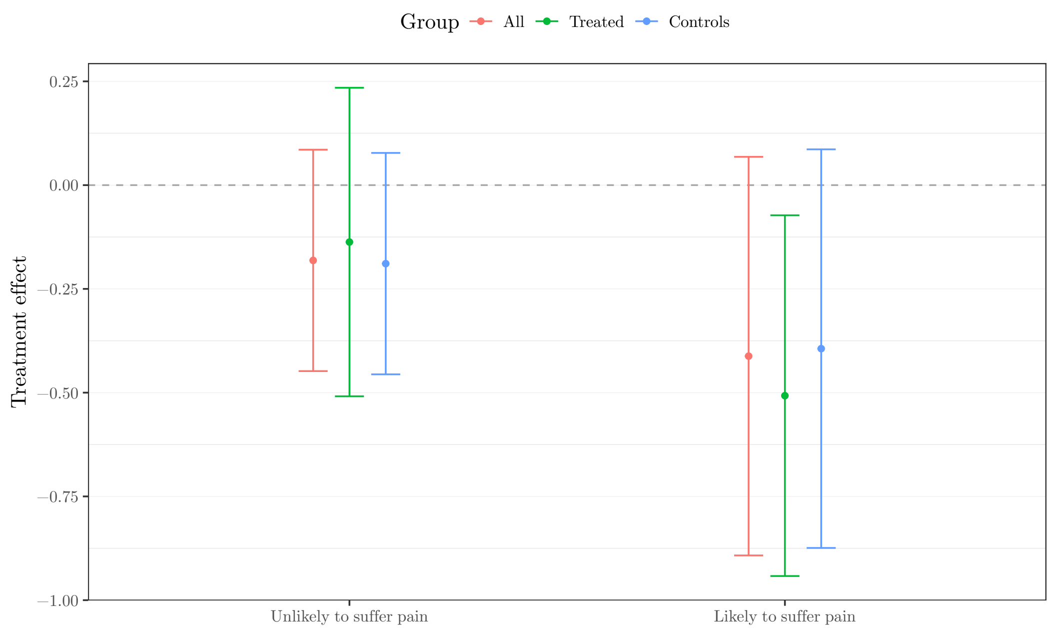
Figure (11) - Distribution of the factual and hypothetical propensities to experience frequent pain.



Notes: For counterfactual predictions, the year has been fixed at 2015. The pain status for this category as people who have experienced pain on at least fifteen days during the past 30 days. The threshold has been chosen to reproduce the national prevalence of chronic pain in 2012 (Nahin 2015). ***Data source:*** BRFSS. Calculated using survey weights.

on those individuals who are unlikely to suffer from pain is both smaller and less precisely estimated than the global effect reported in Table 1. In line with the results for MMLs which allow unspecific pain as a qualifying condition in Section A.2, the improvements in mental health for people likely to suffer pain are more than twice as large as they are for the remaining population and amount to almost half a day less of bad mental health per month. Yet, for both groups, the ATE estimate comes with a large standard error and is statistically insignificant. When we look at the results for the ATT estimates (which are represented by the green elements), substantial and systematic benefits are revealed for those likely to suffer pain. The reported ATT for likely pain sufferers amounts to around six days per year less of bad mental health, and is thus close to the effect on the treated using medical cannabis presented earlier in the discussion. This finding is robust to the exclusion of observations from California. The significant effects identified for the groups of people likely to use medicinal marijuana and the group likely to suffer from frequent pain suggest that a substantial part of the therapeutic value of cannabis might stem from its capacity to alleviate pain with fewer side-effects compared to established drugs.

Figure (12) - Regression-adjusted matching estimates for the effect of medical marijuana laws on bad mental health, i.e., the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the average treatment effect on the controls (ATC) for people who are unlikely or likely to suffer from pain.



Notes: Standard errors have been obtained by bootstrapping using 500 replications. Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

6 Interaction with Prescription Drug Monitoring Programs

A plausible channel through which MMLs influence mental health are potential substitution effects. In addition to those people who start consuming cannabis because access becomes easier, there might well be others who *replace* their current drug of choice by cannabis. Such potential substitution effects seem particularly important given the contemporary upward trends in prescription drug abuse (Dart et al. 2015). A legal supply of quality-controlled marijuana (for medical purposes) might well curb the more damaging illicit intake of other substances for recreational purposes or self-medication. Empirically, the relevance of such considerations is demonstrated by the consequences of the nationwide introduction of abuse-deterrent OxyContin in the United States in August 2010. Alpert, Powell, and Pacula (2017) show that almost every death prevented by decreased OxyContin abuse was replaced by one due to a heroine overdose. This evidence suggests that prohibitive interventions might be ineffective or even backfire when the availability of potential substitute drugs is not taken into account. At least for opioid addicts who genuinely suffer from pain, we argue that substitution using marijuana is likely to be preferable to synthetic opioids such as heroin or the highly potent newcomer Fentanyl. We test this argument on the basis of another supply-side intervention endorsed widely across the United States, i.e., the introduction of so-called Prescription Drug Monitoring Programs (PDMPs).

PDMPs are state-level policies built on a statewide database containing the prescription histories of patients. The specific PDMP regimes show considerable variation across states regarding physicians' obligations: While participation is voluntary in some designs, others require physicians to check the database every time they prescribe drugs that are potentially addictive. When focusing on interventions that strictly require checkups when prescriptions are issued, it is likely that the introduction of a PDMP will immediately disrupt illicit drug supply. Previous evidence indicates that PDMPs are, in fact, associated with significant drops in opioid prescriptions as well as death rates (Bao et al. 2016, Patrick et al. 2016). Yet, judgments about net welfare effects are difficult to deduce from these findings alone. For those people who are not confronted with life-or-death choices, it remains unclear whether they simply gravitate towards other substances and how their mental health is affected. We hypothesize that there is greater potential for a PDMP's introduction to have positive welfare effects if an MML is in place, because cannabis would be more easily accessible as a relatively safe substitute (compared to other substances).

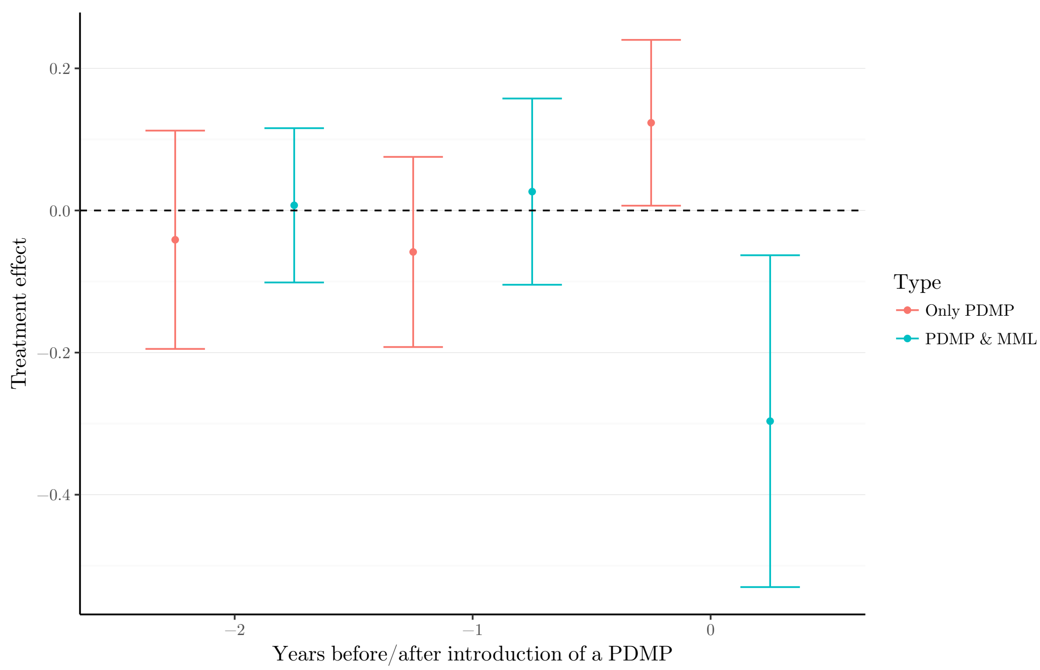
Using the same data as before, we test whether the states which introduce mandatory PDMPs experience more positive effects if an MML is in place at the time of the introduction. The same empirical strategy is applied as previously with MMLs. The estimands are the effects of PDMP treatment dummies interacted with a dummy indicating the presence or absence of an MML. As the earliest PDMP introduction date is June 2012 in West Virginia, we restrict our sample to observations from 2009 onwards. In this way, we include a sufficient number of pre-treatment periods to estimate leads and state-trends, while reducing the incidence of confounding institutional variation and cohort effects. Additionally, we restrict the set of states to those whose MML was in place before 2012. Furthermore, we drop the five states that initiated recreational programs up until the end of 2015. In total, these restrictions result in estimates which are based on 11 treated states, of which 4 ran a parallel MML, and 29 control states.

Figure 13 summarizes the results. The estimates suggest that, on average, a PDMP in a state with an MML caused around 0.4 fewer bad mental health days per a month than one in a state without a MML.[18] This is a substantial difference that is statistically significant at the 1% level (test not shown). Surprisingly, the estimation results suggest that PDMPs in states without a medical marijuana supply might even have led to a short-term net *decrease* in mental health. However, we are cautious in the interpretation of our findings. Especially, the estimate of the harmful effect of PDMPs in the absence of an MML is only marginally statistically significant, and we cannot analyze the post-treatment dynamics due to a shortage

18. The results remain qualitatively the same when we retain another pre-treatment year and control for quadratic state-trends. We also conducted a placebo test where PDMP introduction dates were reshuffled at random among treated states 5,000 times. Our treatment effects both passed the 1% significance hurdle determined by the distribution of pseudo effects.

of data. Furthermore, it is not clear how the potentially endogenous relationship between PDMPs and the restrictiveness of the specific MML implementations affects the estimates. While strict PDMPs are more likely to be introduced in states which disproportionately suffer from abuse problems, the latter also tend to introduce the tightest MMLs (see, e.g., New York). It is not clear a priori how these countervailing factors should balance out. Yet, especially given the substantial effect size, we argue that our estimate is a first indication that MMLs can be understood as justifiable complements to restrictive supply-side interventions on the drug market.

Figure (13) - Lead and main effects of mandatory Prescription Drug Monitoring Programs (PDMPs) on bad mental health conditional on the presence of a MML.



Notes: The data were restricted to 2009 onwards. The states IL, MD, MN, NH, MA and NY are excluded as their MMLs were introduced after 2012. Additionally, CO, WA, OR, DC and AK are also excluded due to their introduction of recreational allowances. State-specific trends are included. Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

7 Conclusions

The consequences of legal access to medical marijuana for individual welfare are a matter of controversy. We contribute to the ongoing discussion by evaluating the impact of the staggered introduction and extension of MMLs across US states on self-reported mental health. Our analysis is based on individual-level data with more than six million observations and exploits 24 interventions over 23 years on the state-level. Employing two-way fixed effects as well as matching estimators, we present and discuss net effects on mental health outcomes for the population as a whole and relevant subgroups.

We find evidence of systematic positive effects on mental health due to the liberalization of medical marijuana for the US population aged 18 and above. Our average treatment effect estimate indicates a reduction of approximately two bad mental health days a year for the representative citizen. We further show that the estimated average improvement in mental health in the general population is not accompanied by an increase in the tail of severe mental health problems. Examining substantive differences between cannabis laws, states which list unspecific pain as a qualifying condition for access to medical marijuana exhibit the largest benefits. This indicates that easier access for patients might overcompensate other potentially adverse effects, such as increased harmful diversion. Most of the benefits seem to accrue in the first three years after the adoption of a MML. Differentiating between the effects across states shows that California contributes disproportionately to the overall positive effect on mental health. Yet allowing for state-specific MML effects and averaging over the interventions, the overall effect remains a systematic reduction of around 1 day a year.

Furthermore, we find large differential responses to MMLs conditional on cannabis consumption modes. While we cannot identify statistically significant effects for abstainers and recreational users, medical users experience systematic gains in terms of their mental well-being. For the latter group, our matching estimate for the average treatment effect indicates that individuals report on average around one day less a month with bad mental health under a liberal cannabis regime. This effect size is approximately one half of the negative impact of frequent alcohol binge drinking on US adults (Okoro et al. 2004). In an alternative partition, we concentrate on people who are likely to suffer from chronic pain. We estimate a reduction of approximately one half of a bad mental health day per month for this group if an MML is in place. Combined with the result for medical users with an effect of similar magnitude, the findings suggest that direct consumption effects are the main drivers behind these benefits.

Finally, we present evidence that supply-side interventions aiming to curb the illicit supply of (non)-medical drugs might be complementary to MMLs. Our results indicate that Prescription Drug Monitoring Programs (PDMPs) only had a positive impact on mental health in those states which also had a medical marijuana law in place.

Overall, our results are in line with the hypothesis that MMLs benefit those categories of individuals for which they are nominally designed without harming other groups systematically. Whether the results carry over to further liberalizations needs additional research, though, and should be carefully considered when deciding about the regulatory regime for marijuana in the future.

References

Alpert, Abby, David Powell, and Rosalie L. Pacula. 2017. "Supply-side drug policy in the presence of substitutes: Evidence from the introduction of abuse-deterrent opioids." NBER Working Paper No. 23031. Cambridge (MA): National Bureau of Economic Research.

Anderson, D. Mark, Benjamin Hansen, and Daniel I. Rees. 2013. "Medical marijuana laws, traffic fatalities, and alcohol consumption." *Journal of Law and Economics* 56 (2): 333–369.

Anderson, D. Mark, Daniel I. Rees, and Joseph J. Sabia. 2014. "Medical marijuana laws and suicides by gender and age." *American Journal of Public Health* 104 (12): 2369–2376.

Austin, Peter C., and Elizabeth A. Stuart. 2015. "Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies." *Statistics in Medicine* 34 (28): 3661–3679.

Bachhuber, Marcus A., Brendan Saloner, Chinazo O. Cunningham, and Colleen L. Barry. 2014. "Medical cannabis laws and opioid analgesic overdose mortality in the United States, 1999-2010." *JAMA Internal Medicine* 174 (10): 1668–1673.

Baggio, Michele, Alberto Chong, and Sungoh Kwon. 2018. "Marijuana and alcohol evidence using border analysis and retail sales data." Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3063288.

Bao, Yuhua, Yijun Pan, Aryn Taylor, Sharmini Radakrishnan, Feijun Luo, Harold A. Pincus, and Bruce R. Schackman. 2016. "Prescription drug monitoring programs are associated with sustained reductions in opioid prescribing by physicians." *Health A airs* 35 (6): 1045–1051.

Bradford, Ashley C., and W. David Bradford. 2017. "Factors driving the diffusion of medical marijuana legalisation in the United States." *Drugs: Education, Prevention and Policy* 24 (1): 75–84.

Bramness, Jørgen G., Fredrik A. Walby, Vidar Hjellvik, Randi Selmer, and Aage Tverdal. 2010. "Self-reported mental health and its gender differences as a predictor of suicide in the middle-aged." *American Journal of Epidemiology* 172 (2): 160–166.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *Review of Economics and Statistics* 90 (3): 414–427.

Carah, Jennifer K., Jeanette K. Howard, Sally E. Thompson, Anne G. Short Gianotti, Scott D. Bauer, Stephanie M. Carlson, David N. Dralle, Mourad W. Gabriel, Lisa L. Hulette, Brian J. Johnson, et al. 2015. "High time for conservation: Adding the environment to the debate on marijuana liberalization." *BioScience* 65 (8): 822–829.

Cartier, Rachel A. 2010. "Federal marijuana laws and their criminal implications on cultivation, distribution, and personal use in California." *San Joaquin Agricultural Law Review* 20:101–115.

Caruana, Rich, Nikos Karampatziakis, and Ainur Yessenalina. 2008. "An empirical evaluation of supervised learning in high dimensions." In *Proceedings of the 25th International Conference on Machine Learning,* 96–103. ACM.

Caruana, Rich, and Alexandru Niculescu-Mizil. 2006. "An empirical comparison of supervised learning algorithms." In *Proceedings of the 23rd International Conference on Machine Learning,* 161–168. ACM.

Caughey, Devin, Christopher Warshaw, and Yiqing Xu. 2015. "The policy effects of the partisan composition of state government." MIT Political Science Department Research Paper No. 2015-24.

Caulkins, Jonathan P., Beau Kilmer, Robert J. MacCoun, Rosalie L. Pacula, and Peter Reuter. 2012. "Design considerations for legalizing cannabis: Lessons inspired by analysis of California's Proposition 19." *Addiction* 107 (5): 865–871.

Centers for Disease Control and Prevention. 2012. "Methodologic changes in the Behavioral Risk Factor Surveillance System in 2011 and potential effects on prevalence estimates." *Morbidity and Mortality Weekly Report* 61 (22): 410–413.

Chapman, Susan A., Joanne Spetz, Jessica Lin, Krista Chan, and Laura A. Schmidt. 2016. "Capturing heterogeneity in medical marijuana policies: A taxonomy of regulatory regimes across the United States." *Substance Use & Misuse* 51 (9): 1174–1184.

Chen, Tianqi, and Carlos Guestrin. 2016. "Xgboost: a scalable tree boosting system." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 785–794. ACM.

Cole, Stephen R., and Miguel A. Hernán. 2008. "Constructing inverse probability weights for marginal structural models." *American Journal of Epidemiology* 168 (6): 656–664.

Dart, Richard C., Hilary L. Surratt, Theodore J. Cicero, Mark W. Parrino, Stevan G. Severtson, Becki Bucher-Bartelson, and Jody L. Green. 2015. "Trends in opioid analgesic abuse and mortality in the United States." *New England Journal of Medicine* 2015 (372): 241–248.

Di Forti, Marta, Craig Morgan, Paola Dazzan, Carmine Pariante, Valeria Mondelli, Tiago Reis Marques, Rowena Handley, Sonija Luzi, Manuela Russo, Alessandra Paparelli, et al. 2009. "High-potency cannabis and the risk of psychosis." *British Journal of Psychiatry* 195 (6): 488–491.

Dolan, Paul, Richard Layard, and Robert Metcalfe. 2011. *Measuring subjective wellbeing for public policy: Recommendations on measures.* Office for National Statistics.

Freisthler, Bridget, Nancy J. Kepple, Revel Sims, and Scott E. Martin. 2013. "Evaluating medical marijuana dispensary policies: Spatial methods for the study of environmentally-based interventions." *American Journal of Community Psychology* 51 (1-2): 278–288.

Friedman, Jerome H. 2002. "Stochastic gradient boosting." *Computational Statistics & Data Analysis* 38 (4): 367–378.

Gavrilova, Evelina, Takuma Kamada, and Floris Zoutman. 2018. "Is legal pot crippling Mexican drug trafficking organisations? The effect of medical marijuana laws on US crime." *The Economic Journal, forthcoming.*

Gruber, Jonathan H., and Sendhil Mullainathan. 2005. "Do Cigarette Taxes Make Smokers Happier?" *The B.E. Journal of Economic Analysis Policy* 5 (1).

Hall, Wayne, and Louisa Degenhardt. 2009. "Adverse health effects of non-medical cannabis use." *The Lancet* 374 (9698): 1383–1391.

Hammer, Michael. 2015. "Patients, police, and care providers: The cultural dimension of medical marijuana implementation." *Administration & Society* 47 (3): 282–297.

Han, Benjamin H., Scott Sherman, Pia M. Mauro, Silvia S. Martins, James Rotenberg, and Joseph J. Palamar. 2017. "Demographic trends among older cannabis users in the United States, 2006–13." *Addiction* 112 (3): 516–525.

Hansen, Benjamin, Keaton Miller, and Caroline Weber. 2017. "How Extensive is Inter-State Diversion of Recreational Marijuana?" NBER Working Paper No. w23762. Cambridge (MA): National Bureau of Economic Research.

Hao, Zhuang, and Benjamin Cowan. 2017. "The cross-border spillover effects of recreational marijuana legalization." NBER Working Paper No. w23426. Cambridge (MA): National Bureau of Economic Research.

Harris, Katherine M., and Mark J. Edlund. 2005. "Self-medication of mental health problems: New evidence from a national survey." *Health Services Research* 40 (1): 117–134.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as non-parametric preprocessing for reducing model dependence in parametric causal inference." *Political Analysis* 15 (3): 199–236.

Hunt, Justin, and Daniel Eisenberg. 2010. "Mental health problems and help-seeking behavior among college students." *Journal of Adolescent Health* 46 (1): 3–10.

Iskedjian, Michael, Basil Bereza, Allan Gordon, Charles Piwko, and Thomas R. Einarson. 2007. "Meta-analysis of cannabis based treatments for neuropathic and multiple sclerosis-related pain." *Current Medical Research and Opinion* 23 (1): 17–24.

Jacobi, Liana, and Michelle Sovinsky. 2016. "Marijuana on main street? Estimating demand in markets with limited access." *American Economic Review* 106 (8): 2009–2045.

Kepple, Nancy J., and Bridget Freisthler. 2012. "Exploring the ecological association between crime and medical marijuana dispensaries." *Journal of Studies on Alcohol and Drugs* 73 (4): 523–530.

Kilmer, Beau, and Robert J. MacCoun. 2017. "How medical marijuana smoothed the transition to marijuana legalization in the United States." *Annual Review of Law and Social Science* 13:181–202.

Lee, Yunhwan. 2000. "The predictive value of self assessed general, physical, and mental health on functional decline and mortality in older adults." *Journal of Epidemiology and Community Health* 54 (2): 123–129.

Mair, Christina, Bridget Freisthler, William R. Ponicki, and Andrew Gaidus. 2015. "The impacts of marijuana dispensary density and neighborhood ecology on marijuana abuse and dependence." *Drug and Alcohol Dependence* 154:111–116.

Marie, Olivier, and Ulf Zölitz. 2017. "«High» achievers? Cannabis access and academic performance." *Review of Economic Studies* 84 (3): 1210–1237.

Marijuana Policy Project. 2016. *Medical marijuana protections in the 50 States.* https://www.mpp.org/issues/medical-marijuana/medical-marijuana-protections-50-states/. Accessed: 2016-05-26.

Martin-Sanchez, Eva, Toshiaki A. Furukawa, Julian Taylor, and Jose L. R. Martin. 2009. "Systematic review and meta-analysis of cannabis treatment for chronic pain." *Pain Medicine* 10 (8): 1353–1368.

Mehmedic, Zlatko, Suman Chandra, Desmond Slade, Heather Denham, Susan Foster, Amit S. Patel, Samir A. Ross, Ikhlas A. Khan, and Mahmoud A. El Sohly. 2010. "Potency trends of Δ9-THC and other cannabinoids in confiscated cannabis preparations from 1993 to 2008." *Journal of Forensic Sciences* 55 (5): 1209–1217.

Meier, Madeline H., Avshalom Caspi, Antony Ambler, HonaLee Harrington, Renate Houts, Richard S. E. Keefe, Kay McDonald, Aimee Ward, Richie Poulton, and Terrie E. Moffitt. 2012. "Persistent cannabis users show neuropsychological decline from childhood to midlife." *Proceedings of the National Academy of Sciences* 109 (40): E2657–E2664.

Morrison, Chris, Paul J. Gruenewald, Bridget Freisthler, William R. Ponicki, and Lillian G. Remer. 2014. "The economic geography of medical cannabis dispensaries in California." *International Journal of Drug Policy* 25 (3): 508–515.

Nahin, Richard L. 2015. "Estimates of pain prevalence and severity in adults: United States, 2012." *Journal of Pain* 16 (8): 769–780.

Nixon, David C. 2013. "Budgetary Implications of Marijuana Decriminalization and Legalization for Hawai'i." Public Policy Center, University of Hawai'i.

Odermatt, Reto, and Alois Stutzer. 2015. "Smoking bans, cigarette prices and life satisfaction." *Journal of Health Economics* 44:176–194.

Odermatt, Reto, and Alois Stutzer. 2018. "Subjective Well-Being and Public Policy." In *Handbook of Well-Being,* edited by Ed Diener, Shigehiro Oishi, and Louis Tay. Salt Lake City, UT: DEF Publishers.

Okoro, Catherine A., Robert D. Brewer, Timothy S. Naimi, David G. Moriarty, Wayne H. Giles, and Ali H. Mokdad. 2004. "Binge drinking and health-related quality of life: do popular perceptions match reality?" *American Journal of Preventive Medicine* 26 (3): 230–233.

Ozluk, Pelin. 2017. "The effects of medical marijuana laws on utilization of prescribed opioids and other prescription drugs." Working paper. https://ssrn.com/abstract=3056791.

Pacula, Rosalie L., Anne E. Boustead, and Priscillia Hunt. 2014. "Words can be deceiving: a review of variation among legally effective medical marijuana laws in the United States." *Journal of Drug Policy Analysis* 7 (1): 1–19.

Pacula, Rosalie L., David Powell, Paul Heaton, and Eric L. Sevigny. 2015. "Assessing the effects of medical marijuana laws on marijuana use: The devil is in the details." *Journal of Policy Analysis and Management* 34 (1): 7–31.

Patrick, Stephen W., Carrie E. Fry, Timothy F. Jones, and Melinda B. Buntin. 2016. "Implementation of prescription drug monitoring programs associated with reductions in opioid-related death rates." *Health A airs* 35 (7): 1324–1332.

Plunk, Andrew D., Arpana Agrawal, Paul T. Harrell, William F. Tate, Kelli E. Will, Jennifer M. Mellor, and Richard A. Grucza. 2016. "The impact of adolescent exposure to medical marijuana laws on high school completion, college enrollment and college degree completion." *Drug and Alcohol Dependence* 168: 320–327.

Powell, David, Rosalie L. Pacula, and Mireille Jacobson. 2018. "Do medical marijuana laws reduce addictions and deaths related to pain killers?" *Journal of Health Economics* 58:29–42.

Reiman, Amanda. 2009. "Cannabis as a substitute for alcohol and other drugs." *Harm Reduction Journal* 6 (35): 1–5.

Reinarman, Craig, Helen Nunberg, Fran Lanthier, and Tom Heddleston. 2011. "Who are medical marijuana patients? Population characteristics from nine California assessment clinics." *Journal of Psychoactive Drugs* 43 (2): 128–135.

Richter, Kimber P., and Sharon Levy. 2014. "Big marijuana - lessons from big tobacco." *New England Journal of Medicine* 371 (5): 399–401.

Ridgeway, Greg, Stephanie A. Kovalchik, Beth A. Griffin, and Mohammed U. Kabeto. 2015. "Propensity score analysis with survey weighted data." *Journal of Causal Inference* 3 (2): 237–249.

Room, Robin. 2014. "Legalizing a market for cannabis for pleasure: Colorado, Washington, Uruguay and beyond." *Addiction* 109 (3): 345–351.

Sabia, Joseph J., and Thanh Tam Nguyen. 2018. "The effect of medical marijuana laws on labor market outcomes." Forthcoming. *Journal of Law and Economics.*

Sabia, Joseph J., Jeffrey Swigert, and Timothy Young. 2017. "The effect of medical marijuana laws on body weight." *Health Economics* 26 (1): 6–34.

Sarvet, Aaron L., Melanie M. Wall, David S. Fink, Emily Greene, Aline Le, Anne E. Boustead, Rosalie Liccardo Pacula, Katherine M. Keyes, Magdalena Cerdá, Sandro Galea, et al. 2018. "Medical marijuana laws and adolescent marijuana use in the United States: a systematic review and meta-analysis." *Addiction* 113 (6): 1003–1016.

Schauer, Gillian L., Brian A. King, Rebecca E. Bunnell, Gabbi Promoff, and Timothy A. McAfee. 2016. "Toking, vaping, and eating for health or fun: Marijuana use patterns in adults, US, 2014." *American Journal of Preventive Medicine* 50 (1): 1–8.

Schubart, Christian D., Iris E. C. Sommer, Willemijn A. van Gastel, Rogier L. Goetgebuer, René S. Kahn, and Marco P. M. Boks. 2011. "Cannabis with high cannabidiol content is associated with fewer psychotic experiences." *Schizophrenia Research* 130 (1): 216–221.

Semple, David M., Andrew M. McIntosh, and Stephen M. Lawrie. 2005. "Cannabis as a risk factor for psychosis: Systematic review." *Journal of Psychopharmacology* 19 (2): 187–194.

Shi, Yuyan, Kristin Meseck, and Marta M. Jankowska. 2016. "Availability of medical and recreational marijuana stores and neighborhood characteristics in Colorado." *Journal of Addiction,* no. 193740.

Słoczyński, Tymon. 2017. "A general weighted average representation of the ordinary and two-stage least squares estimands." Mimeo, Brandeis University.

Słoczyński, Tymon, and Jeffrey M. Wooldridge. 2018. "A general double robustness result for estimating average treatment effects." *Econometric Theory* 34 (1): 112–133.

Smart, Rosanna. 2015. "The kids aren't alright but older adults are just fine: effects of medical marijuana market growth on substance use and abuse." Mimeo, RAND Corporation.

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

Strobl, Carolin, Anne-Laure Boulesteix, and Thomas Augustin. 2007. "Unbiased split selection for classification trees based on the Gini index." *Computational Statistics & Data Analysis* 52 (1): 483–501.

Stuart, Elizabeth A. 2010. "Matching methods for causal inference: A review and a look forward." *Statistical Science* 25 (1): 1–21.

Ullman, Darin F. 2017. "The effect of medical marijuana on sickness absence." *Health Economics* 26 (10): 1322–1327.

Van Ours, Jan C., Jenny Williams, David Fergusson, and L. John Horwood. 2013. "Cannabis use and suicidal ideation." *Journal of Health Economics* 32 (3): 524–537.

Wall, Melanie M., Christine Mauro, Deborah S. Hasin, Katherine M. Keyes, Magdalena Cerda, Silvia S. Martins, and Tianshu Feng. 2016. "Prevalence of marijuana use does not differentially increase among youth after states pass medical marijuana laws: Commentary on and reanalysis of US National Survey on Drug Use in Households data 2002–2011." *International Journal of Drug Policy* 29: 9–13.

Wang, George S., Genie Roosevelt, and Kennon Heard. 2013. "Pediatric marijuana exposures in a medical marijuana state." *JAMA Pediatrics* 167 (7): 630–633.

Wen, Hefei, Jason M. Hockenberry, and Janet R. Cummings. 2015. "The effect of medical marijuana laws on adolescent and adult use of marijuana, alcohol, and other substances." *Journal of Health Economics* 42: 64–80.

Whiting, Penny F., Robert F. Wolff, Sohan Deshpande, Marcello Di Nisio, Steven Duffy, Adrian V. Hernandez, J. Christiaan Keurentjes, Shona Lang, Kate Misso, Steve Ryder, et al. 2015. "Cannabinoids for medical use: a systematic review and meta-analysis." *JAMA* 313 (24): 2456–2473.

Williams, Arthur R., Mark Olfson, June H. Kim, Silvia S. Martins, and Herbert D. Kleber. 2016. "Older, less regulated medical marijuana programs have much greater enrollment rates than newer medicalized programs." *Health A airs* 35 (3): 480–488.

Appendices for online publication

Appendix A Supplementary analyses

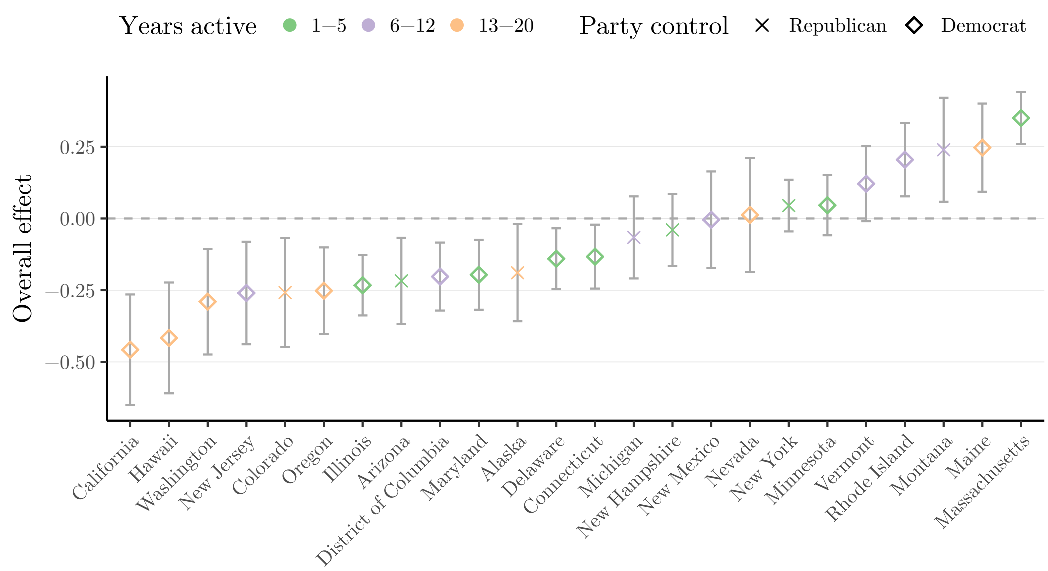
A.1 State heterogeneity

MMLs vary substantially across states and over time. For example, the most liberal law was introduced in California. The state's initial regulation did not impose clear marijuana possession limits, and in 2010, the California Supreme Court officially lifted such restrictions altogether (Cartier 2010). An example of a state that enforced a conservative regulation is New Jersey, where five years after the passage of the law, three operational dispensaries were the only legal sources for medical cannabis. Besides differences in the scope of the law and in juristic details, heterogeneity in treatment responses might also arise due to differences in effective law enforcement. Hawaii is an example. While a nominally restrictive medical marijuana law was passed in 2000, its lax enforcement made the net effect of its passage ambiguous from a theoretical point of view (Nixon 2013).

In a first attempt to gauge the effect of legal heterogeneity, Figure 14 reports estimates for the overall effect of an MML for every state separately, ordered in accordance with the size and direction of the effect. As previously, we employ the same control strategy except that we replace the overall MML dummy with a separate one for every treated state. Here, the effect of the introduction is strongest for California, being both highly systematic and sizeable in reducing the total number of bad mental health days by 0.45 per month. Given the liberal nature of California's implementation, the magnitude is plausibly driven not only by patients, but also by changes in recreational use. Massachusetts, on the other hand, exhibits the most pronounced detrimental effect, i.e., a precisely measured positive effect on bad mental health days of around the same magnitude as California's improvements in mental health. The 90% confidence interval for the average of all the coefficients is approximately $[-0.14, -0.04]$. Hence, giving all the states that introduced a MML equal weight (but allowing for differential estimates of the standard errors) results in a smaller yet more precisely measured overall improvement in mental health. This result reports the average effect of the ***average law*** rather than the treatment effect predicted to be experienced by a random person from the sample.

In order to understand the ranking across states, we include further characteristics of the states and their MMLs in Figure 14. First, we consider single variables which might drive the observed differences. For example, the length of time that a law was in place up until the end of our observation period in 2015 could influence estimates. We check this and indicate the number of years that a law was active by the coloring of the point estimates in the chart. It is revealed that the early adopters appear to be over-represented on the left, where six out of eight states exhibit negative net effects on bad mental health days. With regard to the number of years that an MML was implemented, no pattern emerges for years 1–5 and 6–12. Second, we consider whether states are predominantly led by Democrats or by Republicans, indicated by a diamond- or cross-shaped point estimate. Partisanship might well be associated with unobservables, such as more or less liberal attitudes or the restrictiveness in the implemented cannabis regulations. Figure 14, however, reveals no systematic pattern along this dimension. In an additional analysis (not shown), we also compared laws which were passed by popular ballot with those approved by state legislatures. Again, no clear pattern appears.

Figure (14) - Overall effect of medical marijuana laws by state.

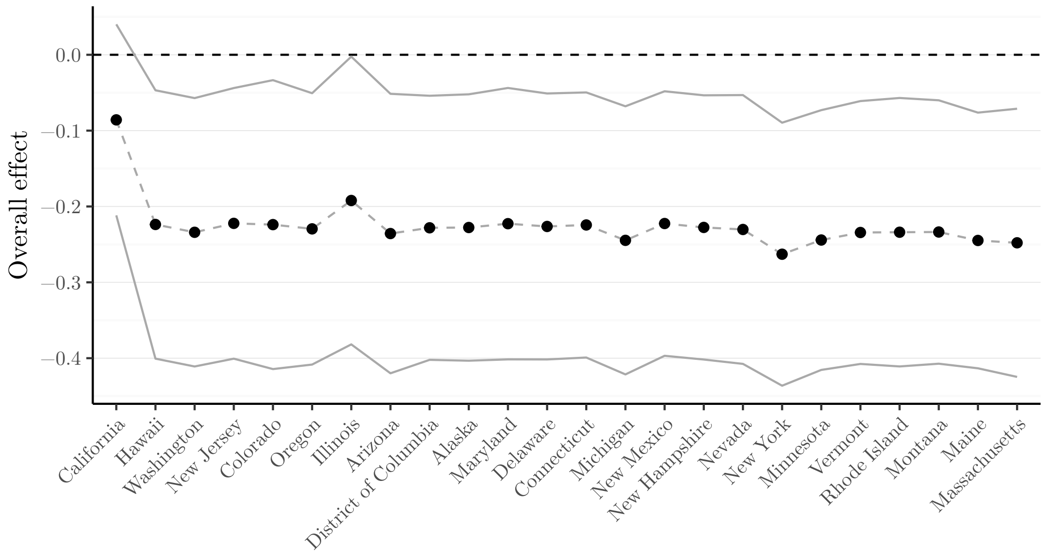


Notes: The symbol shapes and colors represent political control and the time the respective law has been in place, respectively. Categories for the number of years a law has been active have been set using the sample terciles. Party control is determined by the state time average of the partisanship score introduced in Caughey, Warshaw, and Xu (2015). Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

In an additional analysis, we investigate the extent to which the overall effect changes if one individual state that introduces an MML is left out. The results of this exercise are reported in Figure 15. As the plot shows, the effect of MMLs is considerably smaller and statistically not different from zero when California is excluded. This is in line with the large effect size in Figure 14 and the importance California has in our estimates due to its population size. The use of survey weights which respect differential state population sizes can also explain why the exclusion of Hawaii barely changes the estimate, although it exhibits a considerable effect size. In 2012, this state only accounted for approximately 0.4% of the United State's inhabitants.[19] The exclusion of all other states besides California

19. Alternative weighting schemes could be applied. For instance, one might require weights inside a state-year cell to sum to one. Alternatively, we could simply dispense with sample weights. We report a re-estimation of Table 1 using these possibilities in Appendix A.5.

Figure (15) - Sensitivity of overall effects regarding the omission of the specified individual states one at a time.



Notes: Estimations are based on the main specification (1). The states listed on the x-axis were individually omitted in each of the successive estimations. Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

does not alter the estimated effect size, however. Given the importance of observations from California, we also report the results for effect heterogeneity across consumer groups in Section 5 when California is excluded.

A.2 Law heterogeneity

In order to analyze the extent to which institutional variation might drive the differential effects previously discussed in Figure 14, we define further dummy variables characterizing important policy dimensions of the MMLs.

MML dimensions

- **Law only:** Possession of marijuana for the treatment of certain medical condition is legal. Sometimes, access is conditional on the patient being registered with an official office. Under this MML regime, access is eased in so far as doctors can prescribe cannabis for a specific set of ailments, *excluding* chronic pain.
- **Home cultivation:** In addition to the juristic protection offered by a "law only" regime, citizens who receive medical cards from a state office either as patients or caregivers can cultivate some amount of marijuana at home.
- **Unspecific pain:** "Pain due to an unspecified cause" is one of the conditions that allows a physician to legally issue prescriptions. This means that the experienced pain

does not need to be diagnosed as resulting from an acknowledged illness.

- **Dispensaries:** At least one operational state-approved dispensary exists for issuing.

Table (3) - Summary of major policy dimensions characterizing states' medical marijuana laws.

State	Law only	Home only	Pain only	Disp. only	Home & pain	Home & disp.	Pain & disp.	Home, pain & disp.
Maryland	✗	-	-	-	-	-	-	-
New Hampshire	✗	-	-	-	-	-	-	-
New York	✗	-	-	-	-	-	-	-
Connecticut	✗	-	-	✗	-	-	-	-
District of Columbia	✗	-	-	✗	-	-	-	-
Illinois	✗	-	-	✗	-	-	-	-
Montana	✗	-	-	✗	-	-	-	-
New Jersey	✗	-	-	✗	-	-	-	-
Massachusetts	-	✗	-	-	-	✗	-	-
Maine	-	✗	-	-	-	✗	-	-
New Mexico	-	✗	-	-	-	✗	-	-
Vermont	-	✗	-	-	-	✗	-	-
Delaware	-	-	✗	-	-	-	✗	-
Washington	-	-	✗	-	✗	-	-	-
Arkansaw	-	-	-	-	✗	-	-	-
Hawaii	-	-	-	-	✗	-	-	-
Montana	-	-	-	-	✗	-	-	-
Arizona	-	-	-	-	✗	-	-	✗
California	-	-	-	-	✗	-	-	✗
Colorado	-	-	-	-	✗	-	-	✗
Michigan	-	-	-	-	✗	-	-	✗
Nevada	-	-	-	-	✗	-	-	✗
Oregon	-	-	-	-	✗	-	-	✗
Rhode Island	-	-	-	-	✗	-	-	✗

Notes: This table represents the spatial variation in legal heterogeneity exploited for the estimation of the differential effects of policy dimensions in Table 4. States can be mentioned in two categories owing to regime changes over time.

Table 3 provides an overview of the heterogeneity in MMLs across states with regard to the different policy dimensions. This variation in the implemented laws is exploited for the estimation of the differential effects of the various policy dimensions. Table 4 reports the results for the inclusion of separate dummies one by one for the different dimensions in columns (2)–(4). The estimates corresponding to these dummies need to be interpreted as interactions with the presence of a basic MML which only provides legal defense for medical possession. In columns (5)–(7), two dimensions are included at a time, respectively, and they are allowed to interact with each other. Both columns (8) and (9) include all four policy dimensions (with "law only" being the reference category captured in "overall MML" in column (8)) and their interactions. The difference between the columns is that (8) introduces them all as interaction dummies (like in the columns before) whose coefficients report deviations from the effect of a basic MML, whereas column (9) presents a *satiated* formulation. This means that a separate dummy for every possible *combination* of policy dimension is introduced, of which there are eight. This facilitates the interpretation, as the coefficients directly indicate the net effect of these policies. Hence, for example, the coefficient “home & pain” in column (9) suggests that an MML which allows for both home cultivation and unspecific pain results in an average reduction of 0.294 bad mental health days per month. This estimate incorporates a possible interaction between the home cultivation and the unspecific pain dimension.[20]

The various dimensions seem to affect mental health to different degrees. Concentrating on column (9), a striking pattern emerges, highlighting “unspecific pain” (i.e., pain where a clear diagnosis is not required for the prescription) as a strong predictor of positive effects on mental health. The large effects of these regimes might be driven by the patients’ easier access to the drug. However, they could equally result from recreational use as this policy might come along with more opportunities for diversion. In Section 5.3, we provide some evidence that the first driver presumably dominates, which is based on an analysis of the population of likely pain sufferers.

The consequences of operating dispensaries for mental health remain ambiguous in our results. If anything, they seem to reduce eventual benefits of MMLs somewhat, which is most pronounced in the presence of all three policy dimensions. A possible explanation for this is that dispensaries need not necessarily work as *additional* sources of supply, but rather as *replacements* for existing outlets. For instance, in March 2011, Arizona declared that only people living further than 25 miles away from a state dispensary were allowed to be served by a grower. Another difficulty is finding an appropriate measure for dispensaries. For example, it is debatable whether measuring the effect of dispensaries by a simple binary variable (rather than, say, a population coverage measure) is appropriate. Furthermore, it remains ambiguous whether so-called “collective gardens”, which are unions of licensed growers and/or patients, should not be treated as de facto dispensaries too.[21]

A.3 Dynamic effects

Figure 16 illustrates the dynamic effects of MMLs in a flexible way, allowing for differential effects four years before, and up to five years after the introduction. We observe anticipatory effects before the actual implementation. This would be consistent with a de facto liberalization in the respective states before the de jure implementation. Improvements in terms of fewer bad mental health days seem to accrue primarily in the early years after adoption and start to vanish four years after introduction. An interpretation

20. Note that there is not enough variation to credibly estimate the effect of the “pain only” and “pain & dispensary” law configurations. The coefficients reflect more or less state rather than dimension specific law effects which might well be driven by other aspects than just juristic differences.

21. Some recent articles study the spatial effects of cannabis dispensaries. While Mair et al. (2015), Kepple and Freisthler (2012), Morrison et al. (2014) and Freisthler et al. (2013) study neighbourhood effects regarding metrics such as cannabis dependence or crime rates in California, Shi, Meseck, and Jankowska (2016) show that the placement of dispensaries in Colorado is highly non-random and seemingly tied to household incomes and ethnic separation.

Table (4) - Effects of the different policy dimensions characterizing medical marijuana laws (MML) across US states on bad mental health days.

	No. of days (1)	No. of days (2)	No. of days (3)	No. of days (4)	No. of days (5)	No. of days (6)	No. of days (7)	No. of days (8)
Two years before MML	−0.108 (0.077)	−0.108 (0.077)	−0.107 (0.077)	−0.104 (0.077)	−0.104 (0.078)	−0.104 (0.077)	−0.104 (0.077)	−0.101 (0.078)
One year before MML	−0.087 (0.110)	−0.086 (0.112)	−0.088 (0.112)	−0.082 (0.111)	−0.083 (0.113)	−0.082 (0.112)	−0.086 (0.113)	−0.081 (0.115)
Overall MML	−0.227** (0.106)	−0.245*** (0.095)	−0.117 (0.109)	−0.225** (0.102)	−0.182* (0.098)	−0.234*** (0.090)	−0.103 (0.107)	−0.174* (0.099)
Home cultivation	–	0.028 (0.098)	–	–	0.373*** (0.090)	0.017 (0.098)	–	0.385*** (0.086)
Unspecific pain	–	–	−0.181 (0.112)	–	−0.143 (0.101)	–	−0.202* (0.106)	−0.148 (0.099)
Operating dispensaries	–	–	–	0.113* (0.064)	–	0.046 (0.062)	−0.002 (0.052)	0.041 (0.060)
Home × pain	–	–	–	–	−0.336*** (0.103)	–	–	−0.359*** (0.101)
Home × dispensary	–	–	–	–	–	0.082 (0.108)	–	−0.080 (0.104)
Pain × dispensary	–	–	–	–	–	–	0.156* (0.092)	0.006 (0.050)
Home × pain (x) dispensary	–	–	–	–	–	–	–	0.191* (0.114)
Border × MML	0.082 (0.061)	0.086 (0.063)	0.055 (0.064)	0.062 (0.060)	0.044 (0.060)	0.065 (0.065)	0.038 (0.063)	0.025 (0.059)
Border MML	−0.074 (0.068)	−0.074 (0.067)	−0.063 (0.071)	−0.074 (0.066)	−0.051 (0.072)	−0.072 (0.066)	−0.060 (0.070)	−0.049 (0.070)
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
State trends	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Clusters	51	51	51	51	51	51	51	51
Observations	6,349,173	6,349,173	6,349,173	6,349,173	6,349,173	6,349,173	6,349,173	6,349,173
Adjusted R^2	0.088	0.088	0.088	0.088	0.088	0.088	0.088	0.088

* p < 0.1 ; ** p < 0.05 ; *** p < 0.01

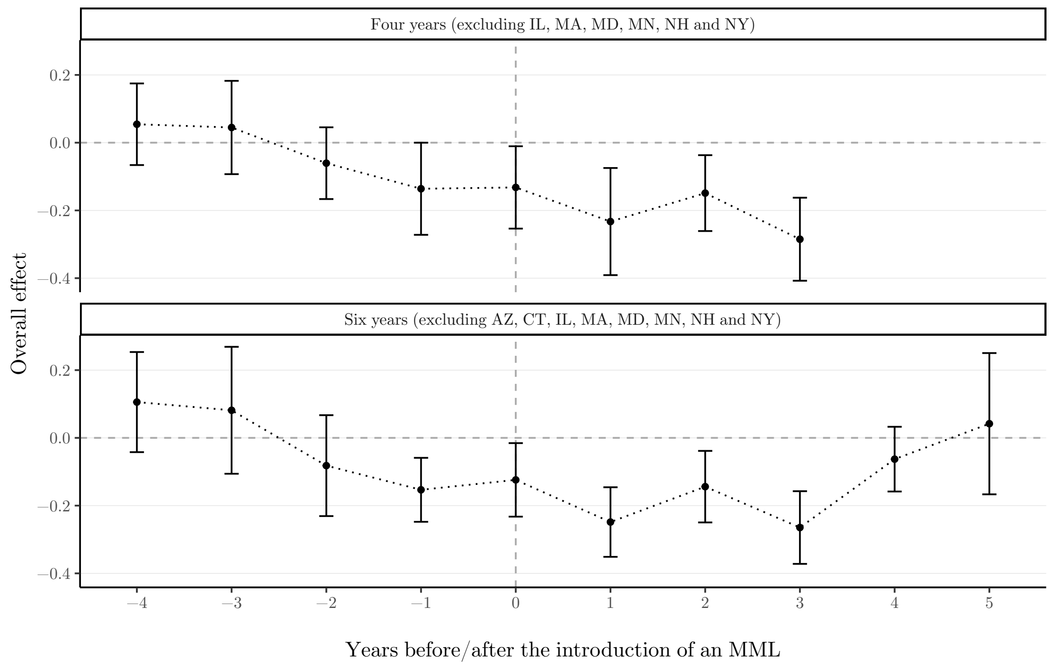
(Satiated specification)	No. of days (9)
Two years before MML	−0.101 (0.078)
One year before MML	−0.081 (0.115)
Law only	−0.174* (0.099)
Home cultivation	0.211* (0.109)
Unspecific pain	−0.322*** (0.101)
Operating dispensaries	−0.133 (0.124)
Home and pain	−0.296*** (0.108)
Home and dispensary	0.172 (0.192)
Pain and dispensary	−0.275** (0.110)
Home and pain and dispensary	−0.139* (0.074)
Border × MML	0.025 (0.059)
Border MML	−0.049 (0.070)
Time FE	✓
State FE	✓
State trends	✓
Controls	✓
Clusters	51
Observations	6,349,173
Adjusted R^2	0.088

* p < 0.1 ; ** p < 0.05 ; *** p < 0.01

Notes: The average number of bad mental health days (dependent variable) amounts to 3.36. Clustered standard errors are reported in parentheses. "Satiated specification" (column (9)) means that coefficients for different treatments can be interpreted autonomously; i.e., they are not interactions but an estimate of the combined effect of all the policy dimensions involved. An overview of which states contribute to the respective policy dimension estimates can be found in Table 3.
Data source: BRFSS. Calculated using survey weights.

is difficult, not least because a limited set of states contributes to the estimation of the dynamic effects (as in some states laws are not in place for long enough). In this context, Williams et al. (2016) highlight that early adopter states have vastly higher enrollment rates for their programs than late adopter states. This renders the assumption of homogeneous treatments questionable, which we implicitly endorsed in our main specification (1). We consider this objection in Section A.2, where we consider the heterogeneity in the law.

Figure (16) - Dynamic overall effects of medical marijuana laws (MML).



Notes: States which do not contribute observations throughout the considered time frame are excluded to avoid spurious effects due to changes in the sample. Confidence intervals are set at 90%. *Data source:* BRFSS. Calculated using survey weights.

A.4 Differential effects on demographic subgroups

Public debates on MML policy often primarily address the possible effects of cannabis regulations on particular demographic groups. Previous research has found differential effects for different segments of the population (see Section 2 for a discussion). Adolescents and young adults, especially males, are a prime example (Hammer 2015). We therefore estimate separate MML effects for different groups conditional on certain age ranges and gender. The results are reported in Table 5. Comparing columns (1) and (2), which distinguish the results between men and women, we observe that, overall, men and women do not seem to be affected differently by an MML regarding the main effect (i.e., the effect in the absence of spill-overs from neighboring MMLs). However, compared to women, the treatment effect for men is smaller if a bordering state's MML is in place at the time of the law's introduction. Furthermore, the positive effect of the ***no jail*** policy on bad mental health is disproportionately stronger for women, while it is unsystematic for men. In columns (3) and (4), samples are further restricted to the age category 18 to 24, which is the youngest category available in the BRFSS. This sample restriction, combined with the gender separation, obviously leads to a substantial loss of precision, as each subsample only comprises approximately 2% of all observations. Still, the estimates suggest that MMLs have no systematic harmful effect on either group. The sample selection in column (5) is inspired by Reinarman et al. (2011) who identify young to middle-aged white men as those most likely to apply for marijuana cards in California. We find no hints of systematic negative effects on mental health for this supposed "risk group". While the negative treatment point estimate is somewhat larger than in the general population, it is imprecisely measured and not statistically significantly different from zero. This result is in line with, for example, the decreased sickness absence at work in middle aged men reported by Ullman (2017).

Column (6) focuses on people older than 64 years of age. This separate analysis is informative for two reasons. First, as shown by Han et al. (2017), this group has become increasingly responsive to MMLs over the recent years. This fact has only recently gained recognition in public debates. One reason for the upward trend in consumption might be the availability of new forms of administering the drug, such as vaping or marijuana smoothies (Schauer et al. 2016). Second, as the hazard of conditions such as chronic pain or neuropathy rises with age, cannabis use in this group is more likely to be due to genuine medical needs. Interestingly, while the treatment effect decreases in size, its precision increases. Furthermore, we observe a significant "fade-in" one year prior to the effective implementation. Including the beneficial border effect (which is completely offset by the border interaction), the results are consistent with the hypothesis that elderly patients experience relevant improvements owing to the new therapeutic options.

A.5 Alternative specifications and sample restrictions

Given the considerable response differences across states reported in Figure 14, the question about the "appropriate" weighting of observations is an important one. Throughout the main article, we employ the survey weights constructed by the BRFSS. In this way, we respect both the differential precisions of the states' contributions to the ATE and come closer to an appropriate weighting of average partial effects (at least for matching estimates, see Solon, Haider, and Wooldridge 2015). Yet, one might hold the position that an unweighted average of every state intervention is more informative for policy makers. While we report such an estimate in Section A.1, another approach would be to normalize weights in state-year cells such that they sum to one.[22] We report such estimates in the columns (1) to (4) in Table 6 where the specification is taken from the global regression in Table 1.

22. Note that this approach has the drawback of biasing standard errors. A weight normalization implies that we ignore the differential precision of states' contributions to the regression estimates due to varying subsample sizes.

Table (5) - Overall effect of medical marijuana laws (MML) on bad mental health of selected demographic groups.

	All ages		Age 18-24			
	Men	Women	Men	Women	White men 24-64	Age 64-100
	No. of days	No. of days	No. of days	No. of days	No. of days	No. of days
	(1)	(2)	(3)	(4)	(5)	(6)
Two years before MML	−0.053	−0.154	0.012	−0.490*	−0.073	0.034
	(0.057)	(0.111)	(0.218)	(0.293)	(0.081)	(0.085)
One year before MML	−0.045	−0.119	0.103	−0.307	−0.220	−0.135**
	(0.089)	(0.142)	(0.184)	(0.281)	(0.151)	(0.067)
Overall MML	−0.233**	−0.217*	−0.089	−0.328	−0.264	−0.146**
	(0.101)	(0.116)	(0.207)	(0.379)	(0.167)	(0.073)
Border × MML	0.161**	0.002	0.204	−0.136	0.052	0.094
	(0.077)	(0.063)	(0.197)	(0.261)	(0.083)	(0.080)
Border MML	−0.094	−0.060	−0.060	−0.078	−0.112	−0.094**
	(0.072)	(0.070)	(0.111)	(0.169)	(0.085)	(0.043)
Sample means	2.68	3.78	3.38	4.85	2.86	2.21
Time FE	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓
State trends	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Clusters	51	51	51	51	51	51
Observations	2,516,826	3,832,347	136,549	159,982	1,375,729	1,844,252
Adjusted R^2	0.084	0.084	0.022	0.025	0.031	0.093

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: The row *sample means* reports the average number of bad mental health days (dependent variable) in the respective sample. Clustered standard errors are reported in parentheses. *Data source:* BRFSS. Calculated using survey weights.

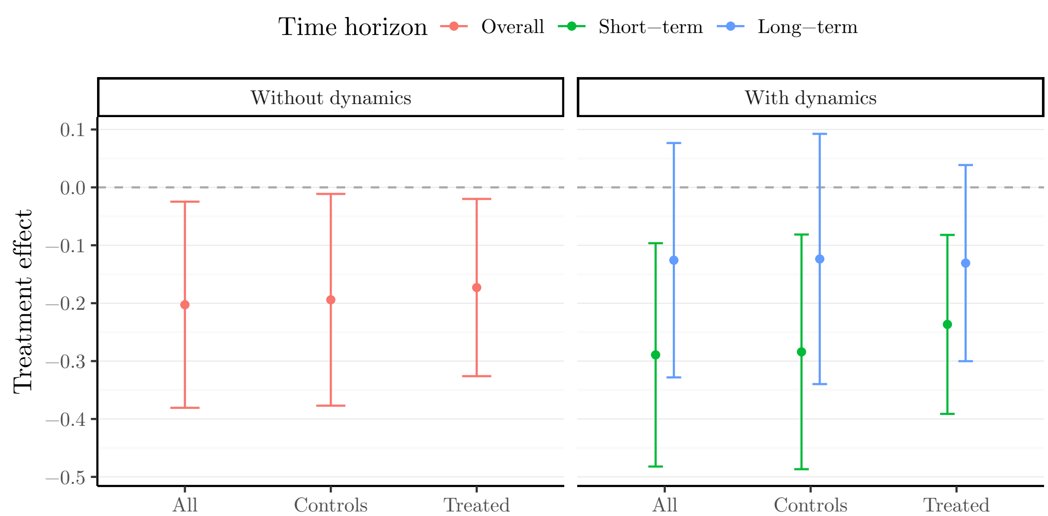
Besides weight normalization, there are arguments which cast doubt on whether survey weights should be employed at all (again, see Solon, Haider, and Wooldridge 2015). We do not follow this reasoning in the main analysis because the responses to MMLs are surely quite different across the population. Since we are interested in net policy effects, these differing average partial effects should be aggregated correctly. Otherwise, it is not clear which population is the subject of the inferences that we draw. Still, for the interested reader, we report estimates without survey weights in columns (5) to (8) in Table 6.

A.6 Matching estimates for global effects

In a supplementary analysis, we report matching estimates of the ATE under the global specification in Table 1. We include this as a robustness check to see whether an alternative weighting scheme of average partial effects over observables has a substantive impact. With regard to the left-hand side of Figure 17, differences to the fixed effect estimates are minor. Furthermore, we see no substantive differences between estimates for the effects on the treated and the controls. In addition, the right-hand side of the plot reports estimates when we allow for a two-year transitory period after the implementation of an MML (green) and a long-term effect thereafter (blue). In agreement with Figure 16, benefits accrue in earlier periods and become unsystematic later on.

Figure (17) - Regression-adjusted matching estimates of (dynamic) overall MML effects on bad mental health.

Confidence intervals are set at 90%.



Notes: Besides all individual characteristics, heterogeneous effects due to differential state Medicaid expenditures and unemployment rates are reflected in the estimates. *Data source:* BRFSS. Calculated using survey weights.

Table (6) - Global specifications for the effects of MMLs on bad mental health applying different weights.

	Normalized survey weights				No survey weights			
	No. of days	No. of days	No. of days	No. of days	No. of days	No. of days	No. of days	No. of days
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Two years before MML	−0.054	−0.058	−0.061	−0.064	−0.039	−0.040	−0.044	−0.048
	(0.051)	(0.049)	(0.050)	(0.049)	(0.039)	(0.037)	(0.037)	(0.034)
One year before MML	−0.009	−0.017	−0.024	−0.030	−0.020	−0.028	−0.034	−0.041
	(0.056)	(0.054)	(0.054)	(0.047)	(0.046)	(0.045)	(0.045)	(0.040)
Overall MML	−0.027	−0.024	−0.031	−0.040	−0.028	−0.043	−0.045	−0.052
	(0.065)	(0.060)	(0.062)	(0.058)	(0.057)	(0.055)	(0.056)	(0.054)
Border × MML	–	−0.018	−0.022	0.009	–	0.015	0.003	0.018
		(0.069)	(0.071)	(0.069)		(0.053)	(0.054)	(0.055)
Border MML	–	−0.075	−0.079	−0.172***	–	−0.077*	−0.066	−0.125***
		(0.049)	(0.057)	(0.045)		(0.044)	(0.049)	(0.041)
Sample means	3.365	3.365	3.368	3.363	3.343	3.343	3.385	3.31
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
State trends	✓	✓	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓	✓	✓
Clusters	51	51	45	29	51	51	45	29
Observations	6,349,173	6,349,173	5,540,433	3,631,697	6,349,173	6,349,173	5,540,433	3,631,697
Adjusted R^2	0.099	0.099	0.099	0.096	0.105	0.105	0.105	0.102

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: In columns (3) and (7), only states are included which introduce or have in place some kind of liberalising cannabis regulation during our sample period. The samples in columns (4) and (8) are further restricted to states which eventually introduce some kind of effective marijuana measure as categorized in the Marijuana Policy Project (2016). Columns (1) to (4) normalize sample weights, such that they sum to one inside state-year cells. In columns (5) to (8), no survey weights are used. Standard errors are reported in parentheses. The row *sample means* reports the average bad mental health days (dependent variable) of the respective sample.
Data source: BRFSS.

Appendix B Descriptive statistics and regression outputs

Table (7) - Summary descriptives for the BRFSS and complementary state-level data.

Variable	Mean/SD
Alcohol days	4.40 (7.71)
Smoking:	
No	81.6%
Yes	18.4%
Children:	
None	68.9%
One	12.5%
Two	11.6%
Three	7.06%
Sex:	
Female	60.4%
Male	39.6%
MML	0.24 (0.43)
BMI	27.3 (5.72)
No jail	0.26 (0.44)
Recreational	0.01 (0.11)
Healthplan:	
No	11.0%
Yes	89.0%
Medicaid	720 (234)
TANF	34.6 (33.6)
Race:	
White	81.1%
Black	7.95%
Asian	1.79%
Nativ	1.47%
Other	7.72%
Marital status:	
Married	54.9%
Divorced	13.8%
Widow	12.5%
Separated	2.22%
Never married	14.1%
Couple	2.43%
Education:	
None	0.18%
Elementary	3.09%
Some highschool	6.34%
Highschool	30.2%
Some college	27.1%
College	33.0%
Employment:	
Employed	46.1%
Self-employed	8.59%
Long-term unemployed	2.13%
Short-term unemployed	2.43%
Homemaker	7.24%
Student	2.33%
Retired	25.2%
Unable	6.02%

Variable	Mean/SD
Income:	
Less 10k	4.99%
Less 15k	5.36%
Less 20k	7.02%
Less 25k	8.72%
Less 35k	11.6%
Less 50k	14.3%
More 50k	35.0%
NA	13.1%
Age:	
[18, 24)	4.67%
[24, 34)	11.8%
[34, 44)	16.0%
[44, 54)	18.8%
[54, 64)	19.7%
[64, 100)	29.0%
Unemployment rate	6.03 (2.06)
Mental health	3.34 (7.61)
Home cultivation	0.19 (0.39)
Operating dispensaries	0.08 (0.28)
Mental health expenditures	75.9 (45.3)
Beer tax	0.19 (0.16)
Unspecific pain	0.17 (0.37)
Cigarette tax	0.76 (0.56)
Border MML	0.43 (0.50)
Border recreational	0.04 (0.20)
Dynamic MML:	
None	75.6%
Transitory	4.67%
Long-term	19.7%
Minimum wage	4.57 (0.48)
City taxonomy:	
Rural	26.7%
Urban	30.1%
Metropolitan	28.9%
NA	14.3%
Party control:	
Totally republican	29.0%
Mostly republican	18.7%
Bipartisan	1.43%
Mostly democratic	22.6%
Totally democratic	28.3%

Table (8) - Complete output for the global regression underlying Table 1, column (2) based on data from the BRFSS and calculated using survey weights.

Cannabis regulations	
Two years before MML	−0.11 (0.08)
One year before MML	−0.08 (0.11)
Effective MML	−0.22 (0.11)**
Border MML	−0.08 (0.07)
Border × MML	0.07 (0.06)
No jail	0.14 (0.07)**
Recreational	0.24 (0.07)***
Border recreational	−0.15 (0.09)*
State and county controls	
Medicaid	0.00 (0.00)***
TANF	−0.00 (0.00)**
Unemployment rate	−0.00 (0.02)
Beer tax	−0.28 (0.35)
Cigarette tax	0.21 (0.07)***
Minimum wage	0.02 (0.04)
`Reference:` Rural	
Urban	0.15 (0.03)***
Metropoly	0.25 (0.03)***
Rural or urban: NA	0.04 (0.03)
`Reference:` Totally republican	
Mostly republican	−0.01 (0.05)
Bipartisan	−0.04 (0.10)
Mostly democratic	0.04 (0.06)
Totally democratic	−0.00 (0.07)
Individual controls	
`Reference:` No children	
One child	0.05 (0.02)**
Two children	−0.01 (0.03)
Three children	0.07 (0.04)*
`Reference:` Single	
Divorced	1.10 (0.02)***
Widow	0.45 (0.04)***
Separated	2.27 (0.14)***
Never married	0.45 (0.04)***
Couple	0.85 (0.05)***
Healthplan	−0.49 (0.12)***
Healthplan × Medicaid	−0.00 (0.00)
Healthplan × TANF	0.00 (0.00)***
`Reference:` Employed	
Self-employed	0.10 (0.02)***
Long unemployed	2.72 (0.08)***
Short unemployed	1.94 (0.05)***
Homeworker	0.12 (0.03)***
Student	0.30 (0.05)***
Retired	0.47 (0.04)***
Unable to work	7.05 (0.10)***
`Reference:` Income: less 10k	
Income: less 15k	−0.40 (0.04)***
Income: less_20k	−0.79 (0.09)***
Income: less_25k	−1.00 (0.07)***
Income: less_35k	−1.29 (0.10)***
Income: less_50k	−1.49 (0.10)***
Income: more_50k	−1.87 (0.11)***
Income: NA	−1.65 (0.08)***
`Reference:` No school	
Elementary school	0.26 (0.08)***
Some highschool	0.78 (0.09)***
Highschool	0.22 (0.12)*
Some college	0.35 (0.11)***
College	−0.28 (0.11)***
`Reference:` Age [18, 24)	
Age [24, 34)	−0.15 (0.06)**
Age [34, 44)	−0.11 (0.07)
Age [44, 54)	−0.35 (0.07)***
Age [54, 64)	−1.23 (0.07)***
Age [64, 100)	−2.67 (0.07)***
Male	−1.54 (0.06)***
`Reference:` White	
Black	−1.01 (0.19)***
Asian	−1.27 (0.24)***
Native	0.09 (0.33)
Other race	−1.15 (0.15)***
`Reference:` Age [18, 24) and Male	
Age [24, 34) × Male	0.38 (0.08)***
Age [34, 44) × Male	0.25 (0.07)***
Age [44, 54) × Male	0.23 (0.06)***
Age [54, 64) × Male	0.51 (0.07)***
Age [64, 100) × Male	1.22 (0.07)***
`Reference:` Age [18, 24) and White	
Age [24, 34) × Black	0.29 (0.13)**
Age [34, 44) × Black	0.27 (0.14)*
Age [44, 54) × Black	0.22 (0.13)*
Age [54, 64) × Black	0.03 (0.15)
Age [64, 100) × Black	0.50 (0.14)***
Age [24, 34) × Asian	−0.03 (0.15)
Age [34, 44) × Asian	−0.41 (0.13)***
Age [44, 54) × Asian	−0.16 (0.23)
Age [54, 64) × Asian	0.16 (0.19)
Age [64, 100) × Asian	0.71 (0.29)**
Age [24, 34) × Native	0.20 (0.23)
Age [34, 44) × Native	0.34 (0.33)
Age [44, 54) × Native	1.19 (0.38)***
Age [54, 64) × Native	0.93 (0.32)***
Age [64, 100) × Native	−0.01 (0.33)
Age [24, 34) × Other race	−0.11 (0.13)
Age [34, 44) × Other race	0.11 (0.14)
Age [44, 54) × Other race	0.73 (0.12)***
Age [54, 64) × Other race	1.23 (0.17)***
Age [64, 100) × Other race	1.43 (0.30)***
`Reference:` Male and White	
Male × Black	0.33 (0.14)**
Male × Asian	0.70 (0.27)**
Male × Native	−0.02 (0.57)
Male × Other	0.69 (0.16)***
`Reference:` Age [18, 24) and Male and White	
Age [24, 34) × Male × Black	−0.07 (0.11)
Age [34, 44) × Male × Black	−0.29 (0.19)
Age [44, 54) × Male × Black	−0.11 (0.18)
Age [54, 64) × Male × Black	0.13 (0.14)
Age [64, 100) × Male × Black	−0.02 (0.13)
Age [24, 34) × Male × Asian	−0.26 (0.28)
Age [34, 44) × Male × Asian	0.32 (0.24)
Age [44, 54) × Male × Asian	0.17 (0.45)
Age [54, 64) × Male × Asian	0.02 (0.22)
Age [64, 100) × Male × Asian	−0.40 (0.38)
Age [24, 34) × Male × Native	0.34 (0.47)
Age [34, 44) × Male × Native	0.83 (0.68)
Age [44, 54) × Male × Native	−0.64 (0.67)
Age [54, 64) × Male × Native	−0.54 (0.60)
Age [64, 100) × Male × Native	0.42 (0.53)
Age [24, 34) × Male × Other race	−0.14 (0.12)
Age [34, 44) × Male × Other race	−0.32 (0.19)*
Age [44, 54) × Male × Other race	−0.47 (0.13)***
Age [54, 64) × Male × Other race	−0.57 (0.20)***
Age [64, 100) × Male × Other race	−0.93 (0.22)***

Table (9) - Effects on those who are unlikely or likely to suffer from recent pain.

	No. of days	
	Unlikely to suffer pain	Likely to suffer pain
Two years before MML	-0.078	-0.292
	(0.087)	(0.231)
One year before MML	0.011	-0.446
	(0.086)	(0.320)
Overall MML	-0.188^{*}	-0.321^{**}
	(0.100)	(0.157)
Border × MML	0.045	0.198
	(0.063)	(0.183)
Border MML	-0.073	-0.033
	(0.074)	(0.121)
Sample means	2.848	7.329
Time FE	✓	–
State FE	✓	–
State trends	✓	–
Controls	✓	–
Clusters	51	–
Observations	5,333,201	1,015,797
Adj. R^2	0.09	–

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Columns (1) and (2) belong to the same regression. Treatment and neighbourhood effects are satiated regarding the propensity group; i.e., coefficients per column can be interpreted independently of the other columns belonging to the same regression. Propensity thresholds have been chosen to reproduce the US national prevalence of chronic pain in 2012 (Nahin 2015). The row ***sample means*** reports the average number of bad mental health days (dependent variable) in the respective sample. Standard errors have been bootstrapped using 500 replications. ***Data source:*** BRFSS. Calculated using survey weights.

Table (10) - Overview of effective introduction dates of cannabis control policies by state.

State	Code	No jail	MML	Home cultivation	Operating dispensaries	Unspecific pain	Recreational
Alaska	AK	-	04.03.1999	04.03.1999	-	04.03.1999	24.02.2015
Arizona	AZ	-	01.04.2011	01.04.2011	06.12.2012	01.04.2011	-
California	CA	01.01.1993	06.11.1996	06.11.1996	01.01.2003	06.11.1996	-
Colorado	CO	01.01.1993	01.06.2001	01.06.2001	01.01.2005	01.06.2001	10.12.2012
Connecticut	CT	30.06.2011	31.05.2012	-	22.09.2014	-	-
Delaware	DE	01.06.2015	01.07.2011	-	25.06.2015	13.05.2011	-
Hawaii	HI	-	28.12.2000	28.12.2000	-	28.12.2000	-
Illinois	IL	-	01.01.2014	-	09.11.2015	-	-
Maine	ME	01.01.1993	22.12.1999	22.12.1999	09.03.2011	-	-
Maryland	MD	01.10.2014	01.06.2014	-	-	-	-
Massachusetts	MA	01.01.2009	01.01.2013	01.01.2013	24.06.2015	-	-
Michigan	MI	-	04.12.2008	04.12.2008	*	04.12.2008	-
Minnesota	MN	-	30.05.2014	-	01.07.2015	-	-
Montana	MT	01.01.1993	02.11.2004	02.11.2004	-	02.11.2004	-
Nevada	NV	01.10.2001	01.10.2001	01.10.2001	31.07.2015	01.10.2001	-
New Hampshire	NH	-	23.07.2013	-	-	-	-
New Jersey	NJ	-	01.10.2010	-	06.12.2012	-	-
New Mexico	NM	-	01.07.2007	01.07.2007	01.07.2009	-	-
New York	NY	01.01.1993	05.07.2014	-	-	-	-
Ohio	OH	01.01.1993	-	-	-	-	-
Oregon	OR	01.01.1993	03.12.1998	03.12.1998	15.08.2013	03.12.1998	01.07.2015
Rhode Island	RI	01.04.2013	03.01.2006	03.01.2006	19.04.2013	03.01.2006	-
Vermont	VT	01.07.2013	01.07.2004	01.07.2004	01.06.2013	-	-
Washington	WA	-	03.11.1998	01.07.2008	-	03.11.1998	09.12.2012
District of Columbia	DC	-	27.07.2010	-	30.07.2013	-	26.02.2015

Notes: Dates earlier than 01.01.1993 are forced onto that date (i.e., the starting point of our survey data) and introductions later than 31.12.2015 are shown as '-'. * We code Michigan as the only state that introduced operating dispensaries and abolished them later on. We set the respective dates at 01.09.2009 and 01.08.2012.

Appendix C Prediction diagnostics

In Section 5, we study the differential effects of MMLs on the mental well-being of likely medical cannabis users, likely recreational cannabis users as well those likely to suffer from pain. In the following, we explain our design choices when calculating the respective propensities.

C.1 Predictors

For the calculation of the likely consumer status, we include as many controls as possible from the second step (mental health regression) in the first step (consumption regression) in order to minimize dependence between controls and propensity scores. The only restriction is that predictors need to be elements of the variable intersection between the BRFSS and the NSDUH. Table 11 reports basic statistics for the used variables.

As we were not granted access to the scientific-use file of the NSDUH, we cannot use state-level variables for predictions. In addition to socio-economic characteristics, we include the frequency of alcohol consumption during the past 30 days, the smoking status as well as the Body Mass Index as predictors in our estimation of the propensities. While these latter variables would be questionable controls in the equation applied in the second step due to endogeneity concerns, we deem the gain in predictive power in the first step sufficient to compensate for an eventual "pollution" of the second step (as some endogenous parts in the propensity scores might be captured).

C.2 Performance

We report performance diagnostics for our boosting predictions along two dimensions. First, we check the accuracy of the probabilities generated by our model. We assess this using a calibration plot as shown in Figure 18. Observations are sorted into bins according to their predicted probability of being a positive case. We choose ten-percentage-points bins where the numbers in the boxes report the number of observations falling into the respective bin. For each bin, the true share of positive cases is then compared to the predicted one. A deviation from the diagonal line indicates bias on the side of the classifier. As can be seen, our model is excellently calibrated with a slight underprediction of positive cases in the lowest range.

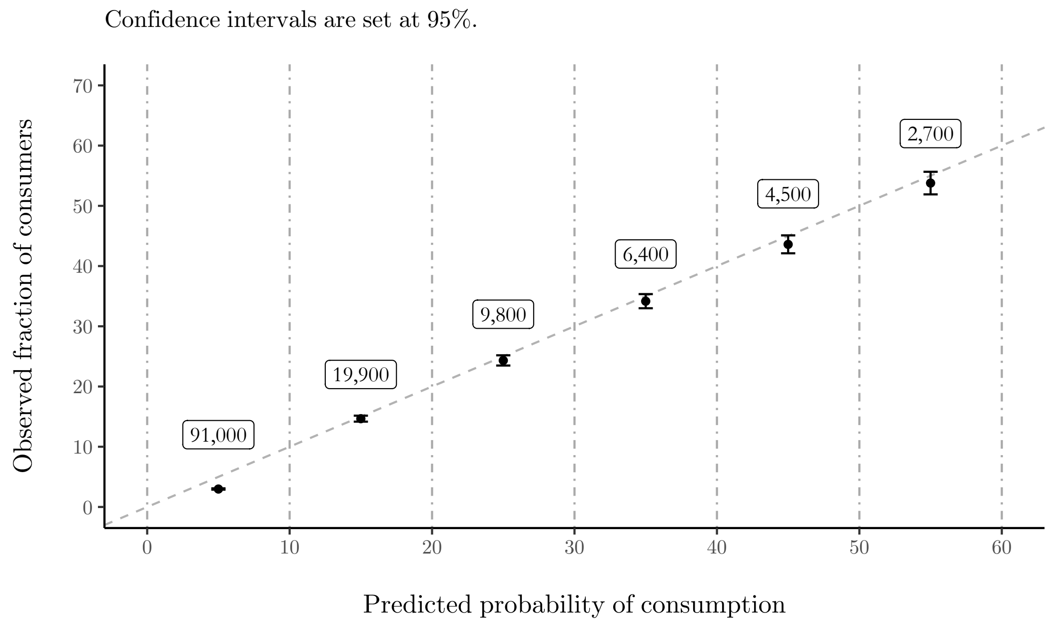
In addition to the calibration plot, we report a confusion matrix and standard metrics of separation quality in Table 12. The right-hand column of the matrix reveals that our ability to separate abstainers from consumers using a propensity threshold is rather modest. According to the metrics, we can expect around 38% of those classified as probable positive cases to be actual consumers. In contrast, non-consumers are easy to predict. Although we use AUC to hypertune `xgboost` due to imbalance, predictions are still conservative regarding negative cases.

Table (11) - Summary descriptives for the NSDUH variables used to predict the cannabis consumption status.

Variable	Mean/SD
Alcohol days:	4.36 (6.81)
Smoking:	
no	68.8%
yes	31.2%
Children:	
none	66.6%
one	14.9%
two	12.1%
three+	6.45%
Sex:	
female	54.1%
male	45.9%
Income:	
less 10k	35.0%
less 20k	21.2%
less 50k	28.6%
more 50k	11.2%
NA	4.05%
Age:	
[18, 24)	34.2%
[24, 34)	28.6%
[34, 64)	31.0%
[64, 100)	6.21%

Variable	Mean/SD
Education:	
some highschool	16.9%
highschool	32.7%
some college	28.8%
college	21.6%
Employment:	
employed	70.4%
unemployed	6.37%
other	23.2%
Race:	
white	64.7%
black	12.5%
other	22.8%
Cannabis days	1.46 (5.61)
Marital status:	
married	39.5%
divorced	9.47%
widowed	2.73%
never married	48.3%
BMI	27.1 (6.12)
Medical marijuana	0.02 (0.20)

Figure (18) - Calibration plot of boosting predictions regarding the extensive margin of cannabis consumption.



Notes: Performance is based on a 20% evaluation subsample which has not been used during training. The x-axis is capped such that 99,9% of the classified observations are represented in the graph. The numbers in the white boxes are the counts of observations inside the respective bins.

Table (12) - Confusion matrix and performance indicators of boosting predictions for the extensive margin of cannabis consumption.

		Prediction		
		No use	*User*	*Sum*
Actual value	*No use*	TN 122,100	FP 5,100	127,200
	User	FN 5,000	TP 3,100	8,100
	Sum	127,100	8,200	

Performance metrics	
AUC	85 %
Accuracy	93 %
Sensitivity	29 %
Specificity	97 %
Balanced acc.	63 %

Notes: The threshold is set to reproduce the estimated national prevalence in the NSDUH sample. Frequencies are rounded to the hundreds digit, and metrics are rounded to the second decimal for readability. *Data source:* NSDUH. Calculated using survey weights.

C.3 Threshold selection for propensity groups

In general, we choose thresholds in propensity regressions such that hypothetical national prevalences for cannabis consumption in the NSDUH are reproduced. When we analyze three categories in the regression underlying Table 2, we face the problem that some observations are both likely to be recreational *and* medical consumers. Since these categories are mutually exclusive, we see no means of interpreting this "cross-over" category in a natural way. Let "clear" predictions be those who are either likely to be recreational *or* medical users. To the best of our knowledge, no scientific consensus has been reached so far in how to arrive at clean categorisations for all observations in the presence of overlap. Hence, we propose an idiosyncratic procedure which works in four steps:

(1) Set the propensity thresholds for recreational and medical consumers such that, in the absence of overlap, national hypothetical prevalences would be enforced. In our case, some predictions will have a high propensity for both classes, making the clear predictions fall short of national rates.

(2) Decrease the threshold for both classes proportionally to their distance to 100% until the sum of clear positive and cross-over predictions equals the national prevalence of cannabis consumption.

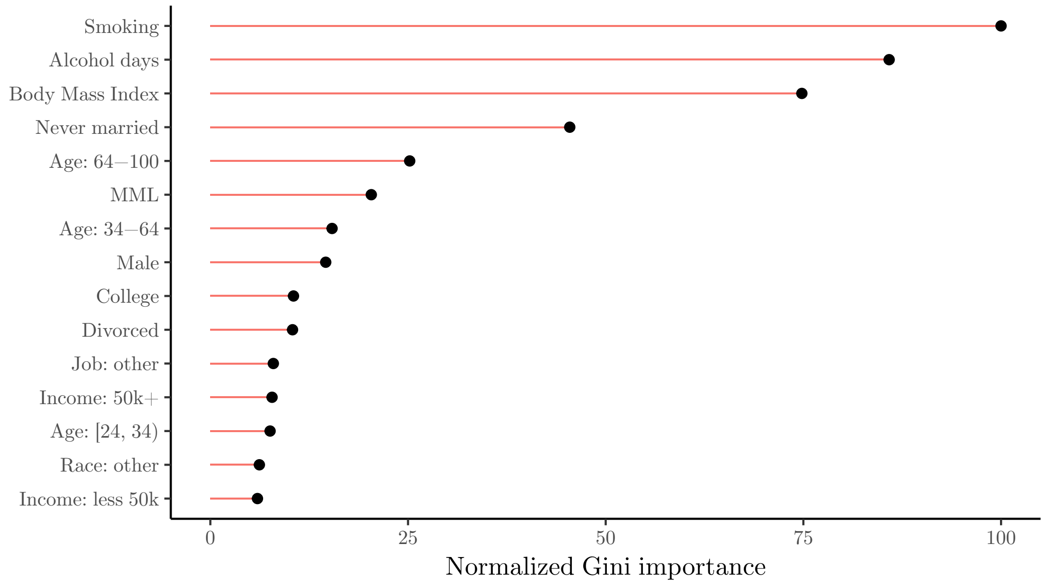
(3) Inside the cross-over category, standardize the propensity scores for recreational and medical use separately. Then take the difference between standardized recreational and medical scores. If the value of one such difference is k, the interpretation is that the recreational score lies k standard deviations further above the recreational cross-over mean than the medical score lies above the medicinal mean.

(4) Taking the standardized score differences from the top downwards, classify so many observations as *recreational* such that clear predictions plus the newly assigned observations equal the national prevalence for recreational use. The remaining observations are then classified as *medical*, enforcing the medical prevalence as well by construction.

C.4 Predictor rankings for probable recreational/medical consumers and probable pain sufferers

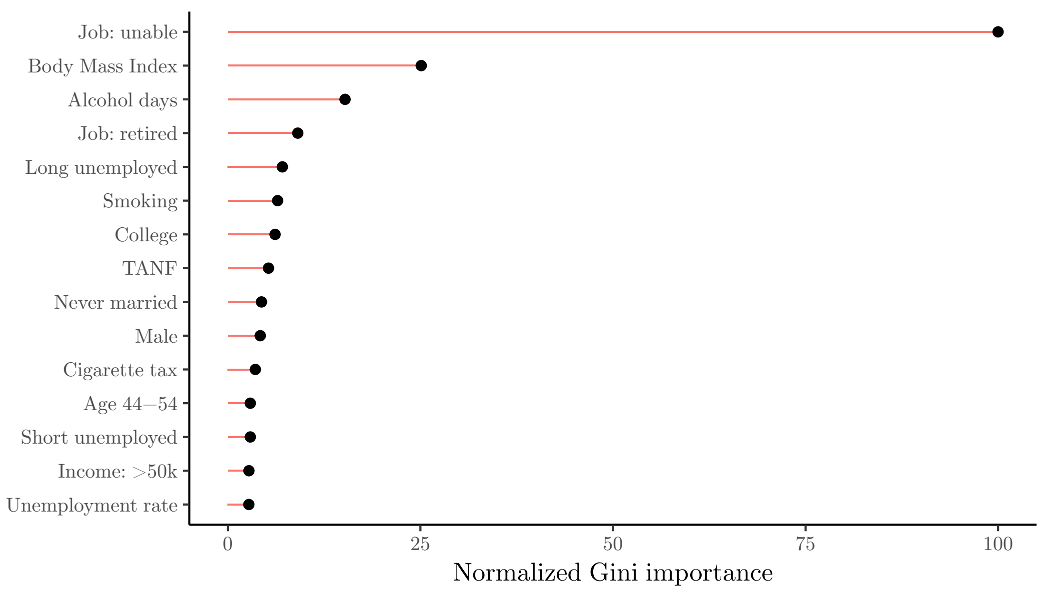
Figures 19 and 20 show which individual characteristics contribute the most to predicting the status of a likely recreational/medical consumer or a likely pain sufferer.

Figure (19) - Ranking of top 15 predictors of recreational and medical cannabis consumption across all decision trees fitted using stochastic gradient boosting.



Notes: The underlying importance metric is ***average Gini impurity reduction*** (Strobl, Boulesteix, and Augustin 2007). We do not report separate importance rankings for medical and recreational users as the R package `xgboost` does not offer such a feature. *Data source:* NSDUH. Calculated using survey weights.

Figure (20) - Ranking of top 15 predictors of frequent pain across all decision trees fitted using stochastic gradient boosting.



Notes: The underlying importance metric is ***average Gini impurity reduction*** (Strobl, Boulesteix, and Augustin 2007). *Data source:* BRFSS. Calculated using survey weights.

Appendix D Matching diagnostics

In order to assess the inverse probability weighting (IPW) in our matching in Section 5, we report two graphical diagnostics (Austin and Stuart 2015). Note that we employ stabilized weights which take into account the unconditional likelihood of receiving the factual treatment (Cole and Hernán 2008). Figure 21 shows a so-called "Love plot" of the ten most unbalanced covariates in the survey weighted sample after adjustment as measured by the standardized mean difference (SMD).[23] Let $\overline{x}_t$ and s_t be the mean and standard deviation of some variable in the treated sample, respectively. If c instead of t is used in the subscript, the variables belong to the control sample. The standardized mean difference is then given as

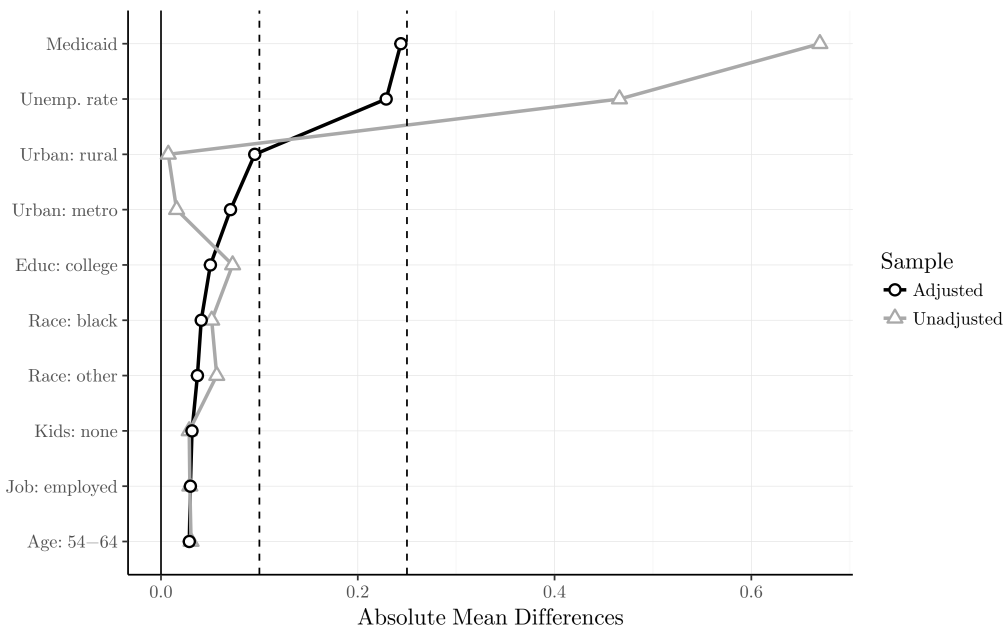
$$d = (\overline{x}_t - \overline{x}_c)/\sqrt{(s_t^2 + s_c^2)/2}.$$

As an empirical rule of thumb, a SMD below 0.1 is deemed excellent while .25 has been proposed as well as a less conservative threshold. In the plot, we see that reweighting decreases the SMD for all the shown variables below .25 (with slight yet insubstantial decline in balance for some variables). This holds for all other variables not shown in the plot as well, including first-order interactions. The relevant output is available upon request. In summary, the achieved balance after the inverse probability weighting is acceptable.

We only balance for state Medicaid expenditures and unemployment rates in the IPW step instead of all institutional variation. The reason is that ***overlap***, one of the components of the ***strong ignorability*** assumption which is necessary for causal inference, requires that $0 < \mathbb{P}(\text{MML}_{ist} = 1 \mid X_{ist}, Z_{st}) < 1$. Our non-parametric boosting approach , however, quickly identifies perfect separation between the treated and controls if we allow for "too many" state variables in the propensity stage. For this reason we restrict the selection to those two institutional variables for which we expect the highest moderating effect on the treatment response. Furthermore, in order to enforce overlap, we dropped 78,000 observations, which is approximately 1% of the sample. The minimum, mean, and maximum of the estimated ATE IPW weights are .21, 1.02 and 28.32, respectively.

23. We employ survey weights both in the propensity score estimation and the second-step outcome regression. As shown by Ridgeway et al. (2015), this is necessary to conduct inference over the right population (i.e., non-institutionalised US adults).

Figure (21) - Overview of pre-post matching covariate balance as quantified by the variable-wise standardized absolute difference.



Notes: The selection is restricted to the ten most unbalanced variables in the adjusted sample in descending order. The dashed lines at .1 and .25 represent rules of thumb proposed in the literature to distinguish between excellent, acceptable and bad balance after matching (Stuart 2010). *Data source:* BRFSS. Calculated using survey weights.