Flying High?
Legalization and the Black Market for Cannabis *

Tiffanie Perrault†

November 12, 2023
[click here for the latest version]

Abstract

How does legalization affect the black market for cannabis? I assemble a novel dataset on US city-level prices and THC potencies, used as proxies for quality, in both prohibition and legalization environments. Difference-in-difference analyses show that legalization and the introduction of legal retailers yield an immediate and large drop in black-market prices, as well as a limited increase in equilibrium quality. This effect on price is driven by medium potency products being subject to important decreases in price, whereas the price of the most potent products remains unchanged *ex-post*. This heterogeneity suggests legalization selecting high potency products on the black market. While the empirical literature has overlooked consumer preferences for cannabis quality, policy design cannot ignore this dimension. To better understand how quality affects the demand and supply of cannabis, I complement the analysis by evaluating a structural model accounting for quality, combining administrative data on legal prices and consumption microdata for the state of Washington. Cross-price elasticities of consumption between legal and illegal cannabis are relatively small. However, changes in THC potency yield sensible substitution between the two products. Counterfactual analysis presents high quality provision as a creditable tool to drive illegal retailers out of the market.

Keywords: cannabis, legalization, policy, demand estimation
JEL Classifications: D12, H80, I18, K49, L19, L51

*I am indebted to Joshua Lewis and Mathieu Marcoux for their outstanding supervision and numerous insightful discussions in every step of the project. I am particularly grateful to Brett Hollenbeck for providing the data from the Washington State Liquor and Cannabis Board, on retail sales for legal cannabis. I also would like to thank Laura Lasio, Erin Strumpf, Simon Franklin, Shweta Mital, as well as seminar participants at the Université de Montréal, McGill University, Queen Mary University of London, the 2022 annual congresses of the Canadian Economics Association and the Société Canadienne de Science Economique, the 21st Annual Canadian Health Economists' Study Group, and the 15th IHEA World Congress on Health Economics for their valuable feedback and constructive comments. All remaining errors are mine.

†Department of Economics, McGill University

1 Introduction

In response to the growing concern about the effectiveness of the *War on Drugs*, the past decade has witnessed a rapid global movement towards the liberalization of cannabis. This momentum was catalyzed in 2012 when the states of Colorado and Washington voted in favor of legalization. As of fall 2023, these policy shifts have extended to twenty-one other states and the District of Columbia.[1] Although these policies align with various government priorities, such as improving consumer welfare, stimulating the economy, and generating fiscal revenue, they all share a common objective: mitigating the influence of the black market and its associated negative externalities.

While the legal market theoretically has the potential to displace illegal retailers, there is limited empirical evidence to evaluate the extent to which current policies have accomplished this objective. Analyzing shifts in cash circulation following the Canadian legalization, Goodhart and Ashworth (2019) suggest significant disruption to the black market. On the other side of the Atlantic, an unintended experience of cannabis liberalization in Italy decreased revenues from cannabis sales on the illegal market by 90-170 million euros (Carrieri et al., 2019).

A number of key empirical issues hinder estimating the effects of legalization on the demand for black market cannabis. These indeed rely on consumers substitution patterns between the legal and the illegal sectors; the estimation of which requires information on products from both sectors within the same market. Due to its illegal – and thereby hidden – nature, seeking data on the black market is particularly challenging. Most data sources on illegal cannabis used by governments and researchers are surveyed or crowd-sourced. Most of them focus on prices and either ignore quality (e.g. the National Survey on Drug Use and Health) or rely on self-assessed discrete categories for quality (e.g. crowd-sourced data from `www.priceofweed.com`). Yet, because the black market for cannabis features high vertical differentiation (Červenỳ and van Ours, 2019), studying the market for cannabis requires objective information on quality. Finally, such an analysis calls for modeling the simultaneous equilibrium interactions between the two markets. This involves obtaining information on cannabis consumption; which remains sensitive, even though social norms have been evolving, and constitutes then another data requirement challenge.

In this paper, I investigate the ability of legalization policies to eradicate the black market. I assemble a novel dataset on city-level crowd-sourced cannabis prices and quality in the US. I digitize 20 years of prices and strains from the Trans-High Market Quotation (THMQ) and match them with their expected THC potency levels, which I webscrape from Leafly's online cannabis consumer guide. These provide an objective measure of quality.[2] To

[1] Internationally, the recreational use of cannabis is now legal in Uruguay, Canada, Georgia, and Mexico.

[2] Using purity or potency as a measure of quality is relatively standard in the literature on drugs (see

analyze the interactions between the illegal and the legal market, I complement this dataset with two additional data sources: legal retail prices from the Washington State Liquor and Cannabis Board (WSLCB) and the Behavioral Risk Factor Surveillance System (BRFSS).[3] This provides me with local prices and quality for both sectors, as well as local cannabis consumption across the state of Washington. Exploiting these data, I model equilibrium responses to legalization using reduced-form and structural methods.

The first part of this work quantifies average black-market price and quality responses to legalization and the implementation of retail sales for legal cannabis. It relies on the THMQ data. Difference-in-difference and event-study estimations show legalization reforms are responsible for the equilibrium black-market prices dropping by up to 20% and THC potency rising by almost 1.4%.[4] Legalization mechanically enhances competition, bringing down the price-cost margin of black-market cannabis. However, this result is driven by medium potency black market products, which are subject to important drops in prices *post-legalization*. This is not necessarily the case for higher potency products, for which reactions are more difficult to predict and which may display zero to positive change in price. This reduced-form analysis confirms the ability of the black market to respond to the legal retail market by combining price and quality adjustments. However, it does not allow to confirm whether the illegal market thrives or shrinks. In a scenario where the price for legal cannabis is "too high" for the legal market to compete efficiently with illegal market, the black market could still respond to the legal market by reducing its price and flourish (see Auriol et al., 2020).

Based on the reduced-form evidence, I propose a structural model of cannabis supply and demand to study the role of price and quality changes induced by legalization. Consumers value price and quality, both on which retailers compete. The core of the analysis relies on a random utility discrete choice model evaluating the choices of consumers in the state of Washington. I estimate the price-elasticity of participation[5] to the black market to lie between -0.2 and -0.3.[6] The elasticity of participation to the legal market is around -0.5. While I find low substitution between the legal and the illegal products with respect to price, consumers are more likely to switch between products upon changes in THC potency. Counterfactuals enable to characterize eviction price and quality strategies for legal cannabis, such that the black market does not survive.

for example Galenianos and Gavazza, 2017).

[3]This annual health survey is conducted by the Centers for Disease Control and Prevention (CDC) and collects state data about US residents.

[4]These average results come from TWFE estimates and are subject to heterogeneity across states and time.

[5]The price-elasticity (respectively quality-elasticity) of participation is defined by the variation in the extensive margin of consumption following a 1% change in price (respectively quality).

[6]This is in line with the results of Jacobi and Sovinsky (2016).

The contribution of this paper to the literature on cannabis legalization is twofold. While the social effects of policy changes have been largely investigated, few projects have quantified the responses of consumption to combined changes in policy and product characteristics. This work further contributes to the literature by being the first to provide estimates for consumer sensitivity with regards to changes in quality (here measured by THC potency).[7] This dimension in consumer preferences has been overlooked in the literature, which has focused on sensitivity to price, availibility and risk.

Following the 2010s wave of legalization, a new strand of literature has studied the reactions to policy changes in terms of crime and consumption. Liberalization policies have resulted in local (Dills et al., 2017; Dragone et al., 2019; Brinkman and Mok-Lamme, 2019) and trans-border decreases in drug trafficking crime (see Morris et al., 2014; Gavrilova et al., 2019; Chang and Jacobson, 2017, for the example of the US-Mexico border). While cannabis legalization shows the intended effects of reducing the negative externalities associated with prohibition, it also increases overall use, as highlighted by Miller et al. (2017) using survey data on undergraduate students at Washington State University.

Three channels drive this effect: price, risk and availability. Most saliently, legalization creates a *riskless* alternative for cannabis consumption and causes the risk of getting caught for illegal consumption to practically disappear.[8] Therefore, since cannabis consumers respond to risk (Jacobson, 2004), they naturally tend to consume more. Retail sales make cannabis more available, granting easier access to the substance. Using a structural model of demand, Jacobi and Sovinsky (2016) extrapolate that stigma and availability effects of legalization would cause cannabis use to increase by 48%. While responses to risk and availibility are well documented, analyses led under prohibitive frameworks miss part of the information necessary to assess retailers' strategic responses. Retail sales of legal cannabis introduce competition with the illegal market, which reacts by setting lower prices. Since both the intensive (Davis et al., 2016; van Ours and Williams, 2007) and the extensive (Jacobi and Sovinsky, 2016) margins of consumption for black-market cannabis are sensitive to price, this strategic response drives up consumption, Consumers are also sensitive to the price of legal cannabis (Hansen et al., 2017; Hollenbeck and Uetake, 2021). While price reveals to be a potential tool for regulating the market for licit cannabis, the literature has focused on either the black market under prohibition or the legal market. This paper is the first to combine information on both illegal and legal products simultaneously to directly evaluate the impact of legalization on the demand for illegal cannabis.

[7] Data have limited other work to discrete measures of quality. Davis et al. (2016) include an indicator for self-assessed high quality in their analysis, while Jacobi and Sovinsky (2016) differentiate "leaf", "head" and "hydro" product types. In my data, quality is objective and continous; which enables me to evaluate elasticities of demand with respect to this dimension.

[8] Under prohibition, simply possessing cannabis is illegal and, hence, liable to sanctions. The legal status of cannabis decreases this risk, making illegal transactions more difficult and more costly to detect.

The sensitivity of consumers to prices provides governments with pricing tools able to reduce increases of consumption induced by legalization (like suggested by the First Article). Taxing legal cannabis not only provides governments with fiscal revenues, it also enables to adjust the price of legal cannabis – and thereby curb use. Hollenbeck and Uetake (2021) show that the retail market for cannabis in the state of Washington, where taxes reach 37%, is still on the upward sloping portion of the Laffer curve. In addition, targetting a given level of consumption through price regulation yields higher social welfare than when employing supply quotas (Thomas, 2019). However, heterogenous effects of legalization on the price of black-market cannabis suggest the equilibrium response of the black market potentially involves the selection of higher potency products, which are more harmful (Di Forti et al., 2019). Quantifying the preferences for potency is therefore key to design legalization policies. The structural results of this paper on both price and quality preferences allow to explore alternative counterfactual policies aimed at eliminating the black market. I show that when the legal sector only competes in price, it has to sacrifice traceability requirements and controls to be able to eradicate the illegal retailers. Enhancing the quality on the legal market overcomes this trade-off.

This article is organized as follows. Section 2 describes the data used in this project. I describe the relationship between the market equilibrium dynamics in terms of prices and THC potencies and the legal status of cannabis in section 3, employing reduced-form techniques. The structural demand model appears in section 4. Finally, section 5 discusses the results, the possible extensions of this work and concludes.

2 Data

This section presents the data used throughout the project. I use a combination of three data sources. Black-market prices, on which I focus in the first part of this work, were retrieved from *High Times'* Trans-High Market Quotation (THMQ). In the second part, I link these data to detailed administrative data on the retail market transactions for legal recreational cannabis in the state of Washington from the Washington State Liquor and Cannabis Board (WSLCB), along with consumption and health data from the Washington State Behavioral Risk Factor Surveillance System (BRFSS) survey. Combining both prices on the black market and retail prices for licit cannabis with consumption data enables me to estimate substitution patterns between legal and illegal cannabis after legalization. These data sources are described below, with more detailed information provided on the THMQ data, as its use in the literature has been relatively sporadic[9] – while the recent IO

[9] Although different extracts of the THMQ data have been used in the economic literature (see Jacobson, 2004; Anderson et al., 2013).

literature (Hollenbeck and Uetake, 2021; Hansen et al., 2017; Thomas, 2019) has featured the data on retail transactions from the WSLCB, and the BRFSS has been well established as a data source in the Health Economics literature.

Consumption data from the WA BRFSS

The BRFSS is a state-based yearly survey, conducted throughout the United States and their territories. The survey is partnered with the Centers for Disease Control and Prevention (CDC) to ensure federal and state public health surveillance. In particular, it aims at monitoring individual health behaviors and conditions, as well as preventive health services.

I use the Washington State BRFSS data, from 2011 to 2017. This micro data includes core questions on individual demographics, socio-economic background and general health. It also includes indicators of extensive margins of cannabis consumption: these consist in two binary variables indicating whether an individual has used cannabis in the past month or year. Since the BRFSS does not provide information on cannabis prices, I combine this data with the price data for the legal and the illegal markets described in the following paragraphs.

Legal prices from the WSLCB *seed-to-sale* tracking system

The data on legal prices is obtained from the Washington State Liquor and Cannabis Board (WSLCB), which oversees the retail cannabis market in Washington. All transactions within the legal retail market until 2017 were recorded in the *seed-to-sale tracking system* known as BiotrackTHC. This measure was implemented to improve traceability, ensure consumer protection, and combat the *grey economy*.

Each plant or clone is given a unique 16-digit identifier at the cultivation stage. This identifier records all relevant information related to the growing and plant maturation process. After harvest, all cannabis components and derivatives are organized in batches. These batches are then assigned another 16-digit identifier, which is linked to the plant identifier – and hence the information it contains. Once at the dispensary, each individual product is given a new code, which is itself linked to the batch.

The data I use account for all retail transactions for legal cannabis in the state of Washington from 2014 to 2017. Each observation contains the retailer license code, the date of transaction, the product type, its strain, upstream and downstream prices, as well as quantities sold. [10] I aggregated this data into local price indices, at the Metropolitan Statistical Area level (see appendix B.1 for detail).

[10] To allow for comparison between quantities of dried cannabis and concentrates, the data includes information of *usable weight*, which refers to the amount of dried cannabis that can be smoked directly, in addition to the variable *weight*.

Black market prices from the THMQ data

The Trans-High Market Quotation (THMQ) data are collected by the *High Times* magazine. First published in 1974, the magazine targets cannabis enthusiasts and advocates for a safe cannabis industry. It encourages readers to share market information, compiled into a monthly price index.

Black-market transaction data used in the literature has primarily relied on relatively short-term data from questionnaire surveys or online crowd-sourced platforms such as `priceofweed.com`(as in Davis et al., 2016).[11] While the website `priceofweed.com` was launched in 2010, i.e. at the verge of the first legalization wave, the *High Times* magazine has been monthly publishing the THMQ for nearly fifty years. This index for black-market prices has become well established in the pool of cannabis consumers, as well as an advantageous data source for studies covering long periods of time.

The THMQ is an unbalanced panel of prices, classified by state. To each state is associated one or several locations – usually a city – to which is associated in turn at least one cannabis strain and its corresponding price. Recent versions of the THMQ usually display prices per ounce, as in figure 1. Older versions, as in figure 2, provide more detail and quantity-price couples and thereby possible quantity discounts.

I collected the THMQ data covered in the *High Times* issues from January 1999 to February 2019. The prices listed are usually collected 3 months before the magazine is issued. Dropping the observations relating to other drugs than cannabis and outside of the United-States, this data set includes 10, 379 prices covering all the US states as well as the District of Columbia. Computing the average price per ounce at which each product (strain) is sold in each location at a given point of time yields a dataset of 8,918 observations.

Information on strain is relatively specific to the THMQ data – compared to other data sources on *illegal* cannabis prices. Strains do not only represent different kinds of plants and tastes, they also reflect diverse THC potencies. In the literature on markets for illicit drugs, measuring quality by using potency or purity is relatively conventional (see for instance Galenianos and Gavazza, 2017). For this reason, I paired the observed cannabis strains with THC potencies scrapped from the website `leafly.com`.[12] Appendix B.2 provides detail on how the data were cleaned and matched.

[11] Another source of data for cannabis prices is the System to Retrieve Information from Drug Evidence (STRIDE), managed by the Drug Enforcement Administration (DEA). However, this data is obtained from undercover buys made by DEA agents. It reflects interactions between law enforcement and targetted suppliers, whereas self-report sources provide information on prices paid by users. Since most transactions occur between people who are already acquainted (Caulkins and Pacula, 2006), the choice of crowd-sourced data, such as the THMQ, could better represent the prices paid by consumers.

[12] This website is one of the largest online cannabis consumer guides. Among other things, it produces a cannabis strain explorer, which, along with crowd-sourced information on effects and reviews, provides the average expected THC potency for each strain.

TRANS HIGH MARKET QUOTATIONS

STATE	CITY	STRAIN	PRICE
ALABAMA	Prattville	Northern Lights #5	$285
ARKANSAS	Little Rock	Tahoe OG	250
ARIZONA	Phoenix	Grape Ape	350
CALIFORNIA	Los Angeles San Francisco	Kosher Kush Guava Chem	300 320
COLORADO	Denver Pueblo	Blue Dream Ghost Train Haze	300 250
CONNECTICUT	Hartford	Trainwreck	360
FLORIDA	North Port	Lamb's Bread	250
GEORGIA	Atlanta	Juicy Fruit	380
HAWAII	Maui	Northern Lights	360
ILLINOIS	Chicago	Gorilla Glue	380
INDIANA	Indianapolis	Critical+	380
IOWA	Des Moines	Death Star	350
KENTUCKY	Albany	Lithium OG Kush	300
LOUISIANA	New Orleans	Skywalker OG	400
MAINE	Portland	Sour Diesel	260
MARYLAND	Baltimore	Blue Dream	380
MASSACHUSETTS	Provincetown	Dakini Kush Girl Scout Cookies	240 240
MICHIGAN	Ann Arbor	Deadhead OG	350
MINNESOTA	Minneapolis	Purple Haze	375
MISSISSIPPI	Oxford	Master Kush	380
MONTANA	Helena	Blue Dream	330
NEVADA	Las Vegas	Three Kings	380
NEW JERSEY	Trenton	Tahoe OG	380
NEW YORK	New York Brooklyn	Gorilla Glue #4 Strawberry Cough	375 360
NORTH DAKOTA	Fargo	Funky Monkey	300
OHIO	Columbus	G-13	360
OREGON	Portland	Goji OG	250
PENNSYLVANIA	Philadelphia	Grand Daddy Purp	400
TENNESSEE	Nashville	Mids Hydro	130 300
TEXAS	Austin	East Coast Sour Diesel	380
UTAH	Salt Lake City	Jedi Kush	360
VERMONT	Bennington	Tangerine Dream	320
VIRGINIA	Richmond	Super Silver Haze	380
WASHINGTON	Seattle	Godfather OG	260
WISCONSIN	Madison	Banana Kush	375
INTERNATIONAL			
CANADA	Montreal Toronto	Bruce Banner Girl Scout Cookies	$C180 150
BELGIUM	Brussels	Jack Herer Casey Jones	€227 227
UNITED KINGDOM	Birmingham	Cheese	£300

Figure 1: THMQ for the September 2017 issue of *High Times*

OHIO
Youngstown: Commercial Brown Brick, "Smells bad, looks bad, tastes bad, but man, did it ever fuck me up, nice quick high, shared four bowls with my four buds, and we were hitting on the mailbox": **$20** 1/4-oz; **$40** 1/2-oz; **$80** oz.
Avon: Kind Buds, "Said to be AK-47 but I am uncertain, bright-green with a great piney taste, fat, dense nuggets with bright-orange hairs and many crystals, uplifting four-hour high off two or three hits": **$50** 1/8-oz; **$100** 1/4-oz; **$400** oz.
North Canton: The Grape, "Light-green and completely covered in crystals, smells like somebody opened up a jar of Smucker's": **$100** 1/4-oz.
Afghani, "A friend in town grew this bud, the buds from this plant are absolutely gorgeous, he got the seeds from a friend in Athens and definitely did well with them, this is the best Afghani around, unfortunately, it's grown only for personal use": **FREE!**
B.C. Buds, "Seeds from this herb came straight from British Columbia, buds are dense, covered in crystals, and large, killer on the head": **$185** 1/2-oz; **$350** oz.

NEW JERSEY
Brick: Blueberry Hydro, "Green nugs with a hint of blue, hairy as hell, loaded with crystals, packs a delightful fruity taste, two hits is more than enough to do the trick": **$65** 1/8-oz.

PENNSYLVANIA
Sayre: RuPaul Bud, "Dark-purple buds with yellow hairs, good high that lasts about five hours with an intense 30-minute plateau": **$75** 1/4-oz; **$125** 1/2-oz; **$210** oz.
Pittsburgh: Kind Nugs, "The buds are dark with a strong musty odor, plenty of hairs and no seeds": **$60** 1/8-oz; **$300** oz.

TENNESSEE
Oak Ridge: Commercial Bud, "Killer shit, great high, one joint will do it, this stuff is really cheap": **$105** oz.
LSD, gel-tabs, "Dark-purple gels, good trip, lasts about 6-8 hours": **$6**/hit.
Oak Ridge Killa, "Light-green, not much smell to it, but it comes through the back door on ya', roll a phatty and get gone!": **$60** 1/2-oz; **$110** oz.
Murfreesboro: Schwag, "Compressed, short-lived buzz, get ya' high but not stoned, this one's good for everyday smoking, going to work and catching a daytime buzz": **$25** 1/4-oz; **$100** oz.

ILLINOIS
Clarendon Hills: White Widow, "Don't buy into this bullshit, nice popcorn puffy buds and a good smoke but not worth the price, better-than-average high with an easy comedown period": **$30**/gm.
Wonka, "This is some quality greenery, one of the longest-lasting highs I have ever experienced, after eight hours you feel so clean that you are in love with everything": **$60** 1/4-oz.
Oak Park: Decent Schwag, "This is OK shit, a few bong rips will get you fucked up, cheap": **$70** oz.
Ecstasy, "Stickman": **$25-$30**/tab

NEW YORK
Buffalo: Killer Green, "*Sativa*, decent-size buds, not the best shit, but it works, a joint does the trick, this one's a C-r-e-e-p-e-r, these are decent buds at a decent price": **$90** oz.
Nasty-Ass Schwag, "Brown, stems, seeds and it smells like dogshit! Sorely lacking hairs and crystals, bricked like a house, can't tell you about potency because I won't touch the shit, get the green, it's a better deal": **$35** 1/4-oz; **$110** oz.
Skunk, "A joint of this and you will stink for a week, knocks you off your chair and right on your ass, the real two-hit shit here, no deals on quantity": **$100** 1/2-oz; **$200** oz.
Schwag #2, "It looks good, it smells good, it tastes good, but it's just not kind bud, got me buzzed for a half hour and then sent me straight to bed, the price on this shit is outrageous, it's more than the Skunk!": **$250** oz.
'Shrooms, "No review, but available": **$100** oz.

KENTUCKY
Louisville: Commercial, "Green buds with plenty of seeds, decent buzz and worth the price": **$35** 1/4-oz; **$65** 1/2-oz.

NORTH CAROLINA
Charlotte: Nice Green Bud, "A few seeds, some orange hairs, excellent taste and aroma, about 3-4 hits for a nice 2-3 hour buzz": **$40** 1/4-oz; **$70** 1/2-oz; **$120** oz.

HAWAII
Maui: Puna Butter, "The real deal, avocado in color, nice smell, one bowl from the pinch-hitter and you're very blind, only one seed in the whole, I germinated it, then it died (sniff, sniff)": **$100** 1/4-oz.
Backyard Greens, "My third crop, it's coming along and it gets you pleasantly stoned, unavailable on the market, I grow it for myself, you sure can't beat the price!": **FREE!**

WASHINGTON
Tacoma: Bubblegum, "Tasty killer green bud, gets you so loaded you forget to REBAKE until hours later, it has a light green that almost belongs on *The X-Files*, a nuclear-green glow as well as a nuclear high, hard to wipe the grin": **$40** 1/8-oz; **$350** oz.
Vancouver: Purple Buds, "Fields of crystals coating dark gray nugs, top-quality buds with a top-quality stone, two tokes and you know why they call it the Evergreen State!": **$50** 1/8-oz; **$345** oz.
Chemo, "Giant colas, fuzzy, light-green with large calyxes and dark-green leaves, all covered with a thick layer of crystals, tastes like no other and burns slow and long": **$45** 1/8-oz; **$310** oz.
Beaster, "B.C. mid-grade commercial hydro, mostly popcorn buds with an occasional big bud, nice smell and a decent taste, when you can't get any Washington ganja this is what to get, always around, always gets you stoned and fairly priced": **$80** 1/4-oz; **$220-$250** oz.
'Shrooms, "Nice Northwest closed caps that send you to the other side, lots of blue stems and phat caps": **$20** 1/8-oz; **$100** oz; **$300** 1/4-lb.

GUYANA
Georgetown: Amazonian Heritage Weed, "Nice smooth high, sometimes it can be a bit harsh due to inadequate handling and drying, but this is no dirtweed, though a bit seedy, these prices are for real!" **$5-7** oz; **$20** 1/4-lb; **$55-60** lb.
Jamaican, "From Mt. Roramina, has a heritage from the Blue Mountains of Jamaica, crossed with some native Amazon, nice high": **$50** lb.

Figure 2: THMQ for the December 1999 issue of *High Times*

3 Reduced-form evidence

This section provides reduced-form results on the black market equilibrium response to legalization reforms. Two strategic outcomes are observed: price and quality.

As one would expect, legalization causes the price for black market cannabis to drop. The newly retail market for legal cannabis introduces competition with the illegal market. Further, legalization introduces licit products which could be diverted to the illegal market, while making illegal behavior more difficult to detect. It could thereby lower barriers to enter the black market and atomize its supply.

On quality, estimation of a two-way fixed effects model on prices and THC potency show that operating legal retail sales of cannabis seems to yield higher quality on the black market. This supports the hypothesis of the black market becoming more competitive and responding to legalization by price and quality differentiation.

One should keep in mind that these are equilibrium results; in particular the effects of legalization on supply could be outweighed by a boom in demand following the reform.

3.1 Average effects of legalization on black-market prices and quality

In this paragraph, I attempt to quantify the average changes in equilibrium on the black market for cannabis, *post-legalization*, in terms of price and quality.

In the US, unlike in other jurisdictions such as Canada, legalization policies are usually implemented in two steps: first the recreational use of cannabis is legalized, then on average two years later, the first legal retail sales of cannabis are implemented (see Appendix A for more detail). I therefore consider two treatments: the legalization cannabis use – hereafter called "legalization" – and the operation of legal retail sales for recreational cannabis. The related twoway fixed effects (TWFE) model is given as follows:

$$y_{ist} = \theta_s + \psi_t + \beta_L \mathbf{L}_{st} + \epsilon_{ist} \tag{1}$$

where y_{ist} is the outcome of interest for observation i collected in state s during month t, θ_s is a state fixed effect,[13] ψ_t is a time fixed effect, $\mathbf{L}_{st}$ is a vector indicating the legalization status in state s at time t, and ϵ_{ist} is a state-level error term that may exhibit within group correlation but is independent from the other regressors. The vector $\mathbf{L}_{st}$ indicates whether recreational use of cannabis is legal, which will be denoted as *legal*, and whether legal retail

[13] In this model, state fixed effects correct for systematic variations in prices across states. States featuring easier access to cannabis *ex ante* could be more likely to liberalize cannabis use. In these states, the *pre-legalization* price for cannabis would be relatively low, which would bias estimates downwards. Besides, locations where cannabis is prohibited may be geographically close to areas in which cannabis is either legal, prohibited but more accessible, or largely exported– e.g. British Columbia or Mexico. Controlling for geographical fixed effects enables to rule out this kind spillover effect.

sales for cannabis are operational, denoted as *retail*.

Two issues here affect the unbiasedness of the TWFE estimator. In the presence of differential timing, the TWFE estimator $\hat{\beta_L}^{fe}$ measures a weighted composite of average treatment effets on the treated (ATT). For instance, Goodman-Bacon (2021) proposes a decomposition of the TWFE estimator into a weighted average of the difference-in-difference estimates resulting from the two-by-two comparisons between all the groups, with weights depending on group sizes and variance in treatment. de Chaisemartin and d'Haultfœuille (2020) show the TWFE estimator can be written as a weighted sum of ATT in each group and period. This implies that unless the treatment effect is homogenous across states and time, the TWFE estimator is biased. The other issue comes with retail sales of recreational cannabis being legalized after recreational use. These two policy changes are considered as two treatments dependent on one another, on top on their staggered adoption. By this design, the estimated effect of the implementation of retail sales is likely contaminated by the effect of the legalization of recreational use (Hull, 2018; Goldsmith-Pinkham et al., 2021; de Chaisemartin and d'Haultfœuille, 2022).

To address these issues, I follow de Chaisemartin and d'Haultfœuille (2020) to check the robustness of the TWFE estimator to treatment heterogeneity in my data.[14] The diagnostic tests proposed mainly consist in computing the weights of ATT in each group and period. In this decomposition, weights sum to 1 but some may be negative. In case the TWFE estimator is not a convex combination of the ATTs, then its sign can be opposite to the sign of the actual treatment effect. The tests also provide two statistics: the minimum variance in treatment such that the treatment effect is zero, as well as (should some of the weights be negative) the minimum variance in treatment such that the treatment effect is of opposite sign as the treatment effect.

I consider the three following specifications:[15]

(i) the effect of the legalization of recreational use was computed without controlling for the implementation of retail sales;

(ii) the effect of the implementation of retail sales was computed without controlling for the legalization of recreational cannabis use;

[14]This diagnostic test method is more suitable in this context than the decomposition proposed by Goodman-Bacon (2021). The latter requires *strictly balanced* panel data. To satisfy this assumption, I would have to aggregate the data at the year level and drop some states, losing a notable amount of information. Further, the decomposition in de Chaisemartin and d'Haultfœuille (2020) can be applied to settings such as this one, in which there are multiple treatments.

[15]In Appendix C, I conduct similar diagnostic tests for two other specifications in which I restrict the sample to observations such that no more than one treatment – *legal* or *retail* – has been applied. The results from these diagnostic tests indicate a likely problem of sign with the TWFE estimate when the effect of *retail* is analyzed on the subsample that has legalized cannabis, which is why I discard these two additional specifications from the analysis.

(iii) the effect of the implementation of retail sales was computed controlling for the legalization of recreational cannabis use.

Results of diagnostic tests for these specifications are provided in Appendix C. They indicate two things. First, when treatments *legal* and *retail* are taken separately the OLS estimates of TWFE model are a convex combination of the ATTs. However, should the treatment effect be highly heterogenous, either across time or units, its average could be zero. Second, contamination between the two treatments *legal* and *retail* is very likely.

Tables 1 through 3 describe the average impact of legalization on the price, quality and quality ajusted price (that is the price normalized by the THC potency) on the black market for cannabis. Columns (1) and (2) of each table provide the results for specification (i) respectively using OLS on the TWFE model specified in (1) and the DiD-M estimator introduced by de Chaisemartin and d'Haultfœuille (2020). Column (3) and (4) describes the OLS and DiD-M results for specification (ii). Column (5) gives the result of the DiD-M estimation of the impact of *retail* when controlling for legalization.

Main takeaways comparing the TWFE and the DiD-M estimators

The OLS estimation results of the TWFE model indicate that legalizing recreational cannabis would result in the black-market price to drop by 19.5% overall. This effect seems strenghened by the implementation of regulated retail sales for recreational cannabis, which result in a similar drop in the black-market price. On quality, THC potency on the black market is not affected by legalization, before retail sales are implemented. However, *retail* results in the THC potency to rise by 1.4%. The effect of *legal* on the potency-normalized price is lower (in absolute value) than the effect found on price, without the normalization. The treatment *retail* results in a drop in the quality adjusted price by more than 16%.

The DiD-M estimates are very different from the OLS estimates of the TWFE model. Further, they display relatively high standard errors. This result, combined with the OLS results as well as the diagnostic tests suggest a high heterogeneity in treatment effects either across time or units. For example, the DiD-M estimate of the treatment *legal* on the black-market price for cannabis is around -7.7%. Comparing it with model (1) suggests that the TWFE estimates are driven down by some heavily weighted units featuring an outstandingly high decrease in the price. On the effect of *retail*, the DiD-M estimate is even from a different sign as the OLS. Since the diagnostic tests indicate all weights associated to this regression are positive, it suggests that while the price in some heavily weighted units decreases notably, it rises in other units *post-legalization*. The results on the two other outcomes reaffirm this intuition.

These results are not surprising. Legalization policies vary in their implementation

Table 1: Difference-in-difference estimates of the effects of legalization on the price of black market cannabis

	(i)		(ii)		(iii)
	OLS (1)	DiD-M (2)	OLS (3)	DiD-M (4)	DiD-M (5)
Treatment					
legal	-0.1950***	-0.0774	-	-	-
	(0.0405)	(0.0732)			
retail	-	-	-0.1901**	0.0689	0.0684
			(0.0319)	(0.1451)	(0.1467)
Fixed effects (OLS)					
State	✓	-	✓	-	-
Year	✓	-	✓	-	-
N	9,460	-	9,460	-	-
n	-	115	-	139	139
Switchers	-	10	-	6	6

Standard errors in parentheses are clustered at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports OLS and DiD-M estimates of the coefficients for indicators of the legalization of recreational cannabis (*legal*) and the operation of legal retail sales (*retail*) on the logarithm of price. Columns (1) and (2) provide the results for specification (i) respectively using OLS and the DiD-M estimator introduced by de Chaisemartin and d'Haultfœuille (2020). Column (3) and (4) describe the OLS and DiD-M results for specification (ii). Column (5) gives the result of the DiD-M estimation of the impact of *retail* without controlling for legalization. Other covariates in the OLS models are state and year fixed effects. DiD-M estimates are robust to dynamic treatment effects. For OLS estimates, N indicates the number of observations. For DiD-M estimates, n gives the number of entities compared in the model while *switchers* is the number of treated entities.

Table 2: Difference-in-difference estimates of the effects of legalization on the THC potency of black market cannabis

	(i)		(ii)		(iii)
	OLS	DiD-M	OLS	DiD-M	DiD-M
	(1)	(2)	(3)	(4)	(5)
Treatment					
legal	0.0038	-0.0581	-	-	-
	(0.0073)	(0.0645)			
retail	-	-	0.0143**	0.0345	0.0345
			(0.0056)	(0.0382)	(0.0331)
Fixed effects (OLS)					
State	✓	-	✓	-	-
Year	✓	-	✓	-	-
N	7,901	-	7,901	-	-
n	-	80	-	119	119
Switchers	-	9	-	6	6

Standard errors in parentheses are clustered at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports OLS and DiD-M estimates of the coefficients for indicators of the legalization of recreational cannabis (*legal*) and the operation of legal retail sales (*retail*) on the logarithm of THC potency. Columns (1) and (2) provide the results for specification (i) respectively using OLS and the DiD-M estimator introduced by de Chaisemartin and d'Haultfœuille (2020). Column (3) and (4) describe the OLS and DiD-M results for specification (ii). Column (5) gives the result of the DiD-M estimation of the impact of *retail* when controlling for legalization. Other covariates in the OLS models are state and year fixed effects. DiD-M estimates are robust to dynamic treatment effects. For OLS estimates, N indicates the number of observations. For DiD-M estimates, n gives the number of entities compared in the model while *switchers* is the number of treated entities.

Table 3: Difference-in-difference estimates of the effects of legalization on the quality adjusted price of black market cannabis

	(i)		(ii)		(iii)
	OLS	DiD-M	OLS	DiD-M	DiD-M
	(1)	(2)	(3)	(4)	(5)
Treatment					
legal	-0.1526***	-0.0361	-	-	
	(0.0247)	(0.1335)			
retail	-	-	-0.1622***	0.0509	0.0509
			(0.0193)	(0.1238)	(0.1264)
Fixed effects (OLS)					
State	✓	-	✓	-	-
Year	✓	-	✓	-	-
N	7,029	-	453	-	-
n	-	80	-	119	119
Switchers	-	9	-	6	6

Standard errors in parentheses are clustered at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports OLS and DiD-M estimates of the coefficients for indicators of the legalization of recreational cannabis (*legal*) and the operation of legal retail sales (*retail*) on the log-difference between the price and the THC potency. Columns (1) and (2) provide the results for specification (i) respectively using OLS and the DiD-M estimator introduced by de Chaisemartin and d'Haultfœuille (2020). Column (3) and (4) describe the OLS and DiD-M results for specification (ii). Column (5) gives the result of the DiD-M estimation of the impact of *retail* without controlling for legalization. Other covariates in the OLS models are state and year fixed effects. DiD-M estimates are robust to dynamic treatment effects. For OLS estimates, N indicates the number of observations. For DiD-M estimates, n gives the number of entities compared in the model while *switchers* is the number of treated entities.

and hence can lead to very different legal markets – on thereby different prices and levels of quality. The 15% the tax rate on retail sales for cannabis in Colorado reflects the liberal spirit in which the reforms were implemented. The state of Washington, where the cannabis market is more regulated, taxes retail cannabis as high as 37%. Other examples of regulations affecting the legal retail market include – but are not limited to – limits on the number of licenses awarded, stricter or laxer traceability controls at the upstream level or regulations on personal home growing. Section 3.2 provides detail on responses by state.

Price decreases in response to *legal* and *retail* suggest both legality and availibility matter. Almost half of the price variation could be attributed to recreational cannabis being legal (without retail sales being regulated). This could be the result of several phenomena. The illegal retailers could anticipate the upcoming competition from the legal retail market and lower their prices. Alternatively, legalization could in theory lower the risk for illegal producers of being detected,[16] which would lower the costs of producing black-market cannabis, producers being subject to lower risk of sanctions. This could imply less seizures, resulting in lower marginal costs, as well as lower investment in infrastructures, which need not be as *hidden* as under prohibition, reducing fixed costs. Lower fixed costs would atomize the supply for cannabis on the illegal market and enhance competition.

Once legal retail competition is introduced, the average potency of black market cannabis rises. This could be explained by consumers going to the black market to find high potency products unavailable legally. Another explanation could be that the black market strategically responds to the legal competition by rising the quality of its products.

As a check for recreational cannabis causing these market responses, I check for effects of unsuccessful legalization ballots on the black-market price and THC potency; and find none (see Appendix D).

The results regarding the effect of *legal* on the quality adjusted price are in contradiction with the results on price and THC potency. If the average THC potency is unchanged *post-legalization*, while the price for cannabis significantly decreases, one should expect the price normalized to THC potency to decrease in the same proportion as the – unnormalized – price. This suggests the possibility of heterogeneous effects on equilibrium prices and average quality, depending on the type of product, i.e. whether the product is low-quality or *premium*. To enquire this, section 3.3 proposes to analyze the effects of legalization across different product categories, constructed based on their THC potency.

[16]Under prohibition almost any production is illegal, which makes detection relatively straightforward compared to a *post-legalization* environment where law enforcement would have to distinguish illegal from legal businesses.

3.2 Heterogeneous effects of legalization on black-market prices and quality

Following the results of the previous section on the disparity between the TWFE and DiD-M, I estimate the effects of legalization policies across states using a stacked difference-in-difference model, as in Deshpande and Li (2019) and Cengiz et al. (2019). It relies on a comparison of each treated unit to its own set of controls. Here each unit of control is a never treeated unit and is only included in one set.

I build the sets of control states based on similarities in terms of electricity prices[17] and climate, which I proxy by the latitude, the average yearly rainfall, as well as the average temperature.[18] Most of the non-labor inputs involved in cannabis growing are electricity and water (Caulkins, 2010; Mills, 2012). Uniform and THC-rich production of cannabis requires stable lighting conditions as well as up to 0.21 gallons of water per square foot per day (Zheng et al., 2021).[19] The variations across states in the quantity of electricity – i.e. lighting, and heating / air-conditioning – required for indoor growing is captured by the latitude and the average temperature. Rainfall does measure the accessibility of water. The local black market for cannabis is also likely to be affected by the proximity to Mexico and Canada, the measure of which is encompassed in the latitude. For the states Colorado, Maine, Massachusetts,[20] Oregon and Washington, I gather a group of five to six states that are the closest in terms of electricity prices and climate.[21] [22] The composition of these control groups is provided in Table 4.

To determine the effects of legalization policies across states, I estimate the following equation:

$$y_{isgt} = \theta_{sg} + \psi_{gt} + \beta_{Lg}\mathbf{L}_{sgt} + \epsilon_{isgt} \tag{2}$$

[17] As of 2020, in cents per kilowatt hours, retrieved online from the U.S. Energy Information Administration, https://www.eia.gov/electricity/state/.

[18] Average yearly temperature and rainfall over the period 2000-2020, from the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental information, Climate at a Glance: National Mapping, published July 2022, retrieved on July 13, 2022 from https://www.ncei.noaa.gov/cag/.

[19] For indoor production, at the peak of the growing season. This figure drops to 0.2 gallons for outdoor production. As a comparison, wheat requires 0.19 and maize / corn 0.17 gallons of water per square foot per day at the peak growing season.

[20] Since Maine and Massachusetts are very similar neighbors and legalized cannabis the same year, I use the same control group for the two of them.

[21] I compute the average distance in percentage of every potential control state to every treated state in terms of precipitation, temperature, latitude and electricity prices. I then average these distances and

1. select in the set of controls the states whose average distance to a given treated state are below 15%,
2. should some states be selected in several control groups, I assign them to the group of the treated unit for which they are the closest.

[22] Although being treated in the data, I discard the states which I either do not observe sufficiently *post-legalization* or for which I cannot find comparable controls.

Table 4: Composition of control groups for each treated state

Group of states	Treated	Controls
1	Colorado	Montana, North Dakota, South Dakota, Utah, Wyoming
2	Maine	Connecticut, New Hampshire, New Jersey, New York, Rhode Island
3	Massachusetts	Connecticut, New Hampshire, New Jersey, New York, Rhode Island
4	Oregon	Iowa, Kansas, Minnesota, Nebraska, Wisconsin
5	Washington	Illinois, Indiana, Missouri, Ohio, Pennsylvania, West Virginia

where the notation builds on the notation used in the TWFE model in (1). y_{isgt} is the outcome of observation i in state s, which belongs to the group of state g, at time t, θ_{sg} and ψ_{gt} are state and time fixed effects, $\mathbf{L}_{sgt}$ indicates the legalization status in state s of group g at time t. The parameters β_{Lg} are to be estimated and differ across groups of states.

The results of estimating model (2) are presented in tables 5 through 7. In each table, the first column refers to the effects of *legal* and the second to *retail*. The results confirm the intuition from the TWFE and DiD-M effects: the effects of legalization on black-market cannabis prices and THC potency vary sensibly across treated states. While policy entails large drops in prices around 30% in the states of Oregon and Washington, it is not necessarily the case in other jurisdictions. For instance, the black-market price of cannabis remains stable in Maine after *legal* is implemented. This heterogeneity is reflected in the results on the quality adjusted price and THC potency. While *legal* entails a significant rise in THC potency by 10% in Massachusetts, there is no effect in Oregon and a negative effect in Maine.

Responses also feature heterogeneity across time. Appendix E explores this feature.

3.3 Heterogenous quality responses to legalization

Following the contradictory results on the effects of legalization policies on the black market outcomes, I divide observations into three categories, depending on their THC potency.[23] These are reported in Table 8 and enable me to investigate whether responses in price differ between products, based on their THC potency.

I estimate the following variation from the TWFE model described by equation (1), which consists in distinguishing the effects of policies on the price of products, depending

[23]This classification follows the classification of the Ontario Cannabis Store.

Table 5: Stacked difference-in-difference estimates of the effects of legalization on the price of black market cannabis

	legal (1)	*retail* (2)
Colorado	-0.146***	-0.196***
	(0.0421)	(0.0549)
Maine	0.0382	-
	(0.0423)	
Massachusetts	0.00397	-0.0480***
	(0.0288)	(9.18e-15)
Oregon	-0.354***	-0.291***
	(0.0118)	(0.0150))
Washington	-0.296***	-0.259***
	(0.0436)	(0.0368)
Average effect	-0.151***	-0.183***
	(0.0188)	(0.0177)
Fixed effects		
State × group	✓	✓
Year × group	✓	✓
N	4,718	4,718

Standard errors in parentheses are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports the OLS estimates for the stacked difference-in-difference model described in equation (2). The first column reports the estimates for the local effects of *legal*, while the second column reports the ones relating to the variable *retail*.

Maine is not observed after the *retail* treatment.

Table 6: Stacked difference-in-difference estimates of the effects of legalization on the THC potency of black market cannabis

	legal (1)	*retail* (2)
Colorado	0.0298***	0.0296**
	(0.00806)	(0.0108)
Maine	-0.0567***	-
	(0.00961)	
Massachusetts	0.104***	0.0609***
	(5.30e-14)	(7.82e-15)
Oregon	-0.00678	-0.000304
	(0.0120)	(0.0130)
Washington	0.0217**	0.0249***
	(0.00876)	(0.00829)
Average effect	-0.00735	0.0248***
	(0.00484)	(0.00504)
Fixed effects		
State × group	✓	✓
Year × group	✓	✓
N	3,943	3,943

Standard errors in parentheses clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table reports the OLS estimates for the stacked difference-in-difference model described in equation (2). The first column reports the estimates for the local effects of *legal*, while the second column reports the ones relating to the variable *retail*.

Maine is not observed after the *retail* treatment.

Table 7: Stacked difference-in-difference estimates of the effects of legalization on the quality adjusted price of black market cannabis

	legal (1)	*retail* (2)
Colorado	-0.243***	-0.297***
	(0.0210)	(0.0307)
Maine	0.122***	-
	(0.0319)	
Massachusetts	-0.0919***	-0.109***
	(7.24e-14)	(9.54e-16)
Oregon	-0.236***	-0.173***
	(0.0318)	(0.0339)
Washington	-0.215***	-0.205***
	(0.0301)	(0.0289)
Average effect	-0.106***	-0.176***
	(0.0158)	(0.0153)
Fixed effects		
State × group	✓	✓
Year × group	✓	✓
N	3,943	3,943

Standard errors in parentheses clustered at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
This table reports the OLS estimates for the stacked difference-in-difference model described in equation (2). The first column reports the estimates for the local effects of *legal*, while the second column reports the ones relating to the variable *retail*.
Maine is not observed after the *retail* treatment.

Table 8: Product classification based on THC potency

Category	Total THC content	Anticipated potency
1	12-16.99%	medium
2	17-20%	strong
3	>20%	very strong

on their *category*.

$$y_{ist} = \theta_s + \psi_t + \sum_{j=1,2,3} category_{ist,j} + \sum_{j=1,2,3} \beta_{Lj} \mathbf{L}_{st} \times category_{ist,j} + \epsilon_{ist} \tag{3}$$

To the notations defined earlier, I add $category_{ist,j}$ which is an indicator of the observation belonging to category $j = 1, 2, 3$. The estimation results are presented in Table 9. They suggest heterogenous price responses on the black market.

Table 9: Difference-in-difference estimates on the effects of legalization on price of black market cannabis by category

	legal (1)	*retail* (2)
medium	-0.107***	-0.155***
	(0.0367)	(0.0338)
strong	-0.149***	-0.136***
	(0.0155)	(0.0278)
very strong	-0.0336	-0.0354
	(0.0672)	(0.0481)
Fixed effects		
State × category	✓	✓
Year	✓	✓
Category	✓	✓
N	7,219	7,219

Standard errors in parentheses are clustered at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
This table reports the OLS estimates for the TWFE model described in equation (3).

Cannabis liberalization is associated with a decrease in the price of medium to strong potency cannabis. Strong potency products observe a moderate drop in price after legalization – 14.9% – which accentuates after legal retail sales are implemented – retail sales being responsible for a drop in price by 13.6%. Medium potency products see their price decrease by 15.5% after the implementation of retail sales. Assuming the demand for these medium range products does not decrease, this feature suggests that price differentiation is relatively important in the market for medium range cannabis products.

On the other end, the most potent products see no significant change in their price after legalization. This fact, along with the general observation that THC potency rises when the legal market is introduced, suggests that differentiation on premium products would be mostly based on quality, rather than price.

These results are averages over treated states. To refine them, I estimate the following variation from equation (2), which consists in distinguishing the effects of policies on the price of products, depending on their *category*. This requires to further refine the different groups of comparison, classifying observations not only by cohort of states but also product category.

$$y_{isgct} = \theta_{gc} + \psi_{gct} + \beta_{Lgc}\mathbf{L}_{sgct} + \epsilon_{isgct} \tag{4}$$

To the notations defined earlier, I add the subscript c which is an indicator of the observation belonging to category $c = 1, 2, 3$.

The estimation results are presented in Table 10. Except in Massachusetts and Maine, policies entail decreases in the price of medium and strong potency cannabis products. Treatment effects on very strong potency products are more sporadic and do not seem to follow any general rule.

4 Uncovering consumer preferences for cannabis: evidence from the state of Washington

The first part of the chapter shows legalization reforms have caused reactions in the black-market prices and THC potency, the former being subject to large drops while the latter rise moderately. In the state of Washington, legalization policies are associated with decreases in prices for illegal cannabis by 25 to 30%. This decrease is driven by the products of medium to strong potency, while the very strong types of cannabis see their price unchanged *ex-post*. Meanwhile, the THC potency rises by more than 2%. This strategic reaction supports the intuition that legalization atomizes the supply for black market cannabis and reduces its production and distribution costs, through changes in risk. Further, the heterogeneity of price responses depending on the product category suggest some selection of the black market products towards higher potency.

Yet, the underlying mechanisms responsible for these effects are not clear and the analysis requires more structure to assess the extent to which legalization weakens the illegal market. This part of the analysis is all the more important since the effects observed are heterogeneous across states, time and product categories: strategic responses of the black market are complex.

Modeling consumers preferences for legal and illegal cannabis, both before and after legalization is necessary to fully understand the effects of legalization on consumption. Under prohibition, consumers who wish to use cannabis necessarily turn to the black market. They

Table 10: Difference-in-difference estimates on the effects of legalization on price of black market cannabis by state and category

		legal (1)	*retail* (2)
Colorado	medium	-0.164*	-0.216***
		(0.0827)	(0.0693)
	strong	-0.221***	-0.272***
		(0.0404)	(0.0613)
	very strong	0.405***	-0.262***
		(2.63e-09)	(1.32e-09)
Maine	medium	0.201***	-
		(0.0420)	
	strong	0.0312	-
		(0.0372)	
	very strong	-	-
Massachusetts	medium	0.0495	-
		(0.0562)	
	strong	0.0204	-
		(0.0269)	
	very strong	-0.336***	-
		(0.0438)	
Oregon	medium	-0.271***	-0.236***
		(0.0576)	(0.0765)
	strong	-0.177***	-0.119***
		(0.0184)	(0.0207)
	very strong	-1.070***	-1.070***
		(0.225)	(0.225)
Washington	medium	-0.202***	-0.269***
		(0.0483)	(0.0447)
	strong	-0.160***	-0.0900***
		(0.0371)	(0.0297)
	very strong	-0.0265	-0.0815
		(0.231)	(0.222)
Average effect	medium	-0.0773**	-0.240***
		(0.0285)	(0.0375)
	strong	-0.101***	-0.160***
		(0.0171)	(0.0237)
	very strong	-0.206***	-0.471***
		(0.0651)	(0.105)
Fixed effects			
State × group × category		✓	✓
Year × group × category		✓	✓
N		4,820	4,820

Standard errors in parentheses are clustered at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
This table reports the OLS estimates for the stacked difference-in-difference model described in equation (2). The first column reports the estimates for the local effects of *legal*, while the second column reports the ones relating to the variable *retail*.
Maine and Massachusetts are not observed after the *retail* treatment. (Maine is simply not observed, while categories could not be determined in Massachusetts for this period of time).

purchase cannabis if their indirect utility derived from cannabis consumption is positive. This utility depends on the market price, observed quality – measured by THC potency – and unobserved heterogeneity, as well as individual characteristics. Legalization introduces a new option in the consumers' choice set. This legal alternative for consuming cannabis is valued differently than black-market cannabis, involving both new potential cannabis consumers and former black-market consumers joining the legal market.[24] Yet, legalization does not automatically pair with the disappearance of the illegal market. Some consumers might remain on the black market *post-legalization*, in particular if the legal market is not attractive (see Chapter 1). Preferences for illegal cannabis are therefore a significant piece of information to understand preferences and choices for legal cannabis.

This section relies on a random utility discrete choice model, applied to cannabis consumption choices in the state of Washington and specifically accounting for preferences for quality. I estimate own- and cross- elasticities of consumer participation in the legal and the illegal markets with respect to both price and quality. These document substitution patterns[25] and enable to retrieve structural estimates for marginal costs of producing and distributing cannabis on both the legal and the illegal markets. Modeling the competition between these enables me to calibrate the black market's best-response function to changes in price and THC potency of the legal product. This counterfactual exercise highlights the importance of THC potency as a tool to regulate the cannabis market.

[24] Chapter 1 provides a theoretical framework on general equilibrium dynamics and detail consumer selection in partial equilibrium *post-legalization*. In particular, it shows that under partial equilibrium, legalization, by introducing a new option and expanding consumers' choice set, increases the overall demand for cannabis.

[25] Future versions of this work will include random coefficients for the sensitivities to price and quality. One limit of the simple logit model presented in this version relates to the *independence of irrelevant alternatives* (IIA) hypothesis, which forces restriction on substitution patterns. A change in one attribute of a given option yields the same change in the probability of all other options. For example, if the price of legal cannabis decreases, it entails the same decrease in the probability of choosing illegal cannabis and not choosing to consume cannabis, while one could expect a proportion of new users lower than the proportion of illegal cannabis consumers turning to the legal market. Hence, the logit model would overestimate the rise in demand following an improvement in one attribute of the legal option. Following an improvement in one attribute of the illegal option, one should similarly expect a higher decrease in the market share of legal cannabis than in the market share of the outside option. The model is therefore likely to underestimate the share of the outside option, as it is the case in Appendix G. For the same reasons, one could also expect the share of the legal market to be overestimated by the model. Expanding the choice set, because of the IIA, clearly affects the ability of the logit to properly predict counterfactual market shares. This lack of precision could be reinforced by the fact that β_1 is constant. Yet, since the predicted market shares in appendix are the result of averaging individual market shares over a large population, this explanation seems less likely to be the main driver.

4.1 The demand for recreational cannabis

4.1.1 Model

I consider the following discrete choice model, where an agent $i \in \mathcal{I} = \{1, \ldots, N\}$, living in the Metropolitan Statistical Area (MSA) $m = 1, \ldots, M$ at time $t = 1, \ldots, \mathrm{T}$, decides whether to consume cannabis or not. Under prohibition, available products exclusively come from the black market. After legalization, agents who wish to consume cannabis choose between two differentiated products: illegal ($j = 1$) and legal ($j = 2$) cannabis. Not consuming cannabis is considered the outside option ($j = 0$). Formally, the indirect utility is given as follows.

$$u_{ijt} = \beta_{pj} p_{jmt} + \beta_{qj} q_{jmt} + \beta_{Xj} X_{imt} + \xi_j + \Delta\xi_{jmt} + \epsilon_{ijmt} \tag{5}$$

where ϵ_{ijmt} is some agent-good-market specific idiosyncratic term, known to agent i but unknown to the econometrician. I assume ϵ is an independent Extreme Value Type I variable. $\beta_j = (\beta_{pj}, \beta_{qj}, \beta_{Xj})$ is a vector of parameters to be estimated. The utility derived from choosing the outside option $j = 0$ is normalized to $u_{i0t} = \epsilon_{i0t}$, for all consumers i and on all markets m and periods t.

The indirect utility derived from cannabis consumption depends on a number of factors, including the price p_{jmt} and the THC potency q_{jmt}, observed for cannabis of type $j = 1, 2$ in market m and period t, as well as individual demographic and health characteristics (represented by the vector X_{imt}).

The value derived by agents when purchasing legal cannabis is different from the value derived when purchasing black-market cannabis. The product fixed effect ξ_j and the random variable $\Delta\xi_{jmt}$ account for these effects. In particular, $\Delta\xi_{jmt}$ relates to shocks in the valuation of consumers in market m and period t for unobserved characteristics of product j.

In my model, the extent to which individual preferences affect the utility derived from illegal consumption are policy invariant. Data limitation, namely the fact that I do not observe the type – legal *versus* illegal – of cannabis consumed *ex-post*, makes this assumption necessary. Hence, the change in consumer choices is not caused by a change in preferences *per se*. It is rather the result of the birth of a retail market for legal cannabis, individual (ϵ_{ijmt}) and market-good ($\Delta\xi_{jmt}$) specific shocks, as well as changes in market prices and THC potencies.

Time and product specific variables also affect the benefit of consuming cannabis. At the time of its legalization, cannabis had been prohibited for almost a century; it is still prohibited in most states. While legalization is the result of evolving social norms, it is also likely to have accelerated the change towards acceptance of cannabis consumption; social stigma fading with time. This effect is captured in the the random variable $\Delta\xi_{jmt}$.

To ease the exposition, the market-product-specific terms are regrouped under the notation

$$\delta_{jmt} \equiv \beta_{pj} p_{jmt} + \beta_{qj} q_{jmt} + \xi_j + \Delta\xi_{jmt}$$

and the mean conditional valuation of individual i for good j in market m and period t is defined as

$$\bar{u}_{ijmt} \equiv \delta_{jmt} + \beta_{Xj} X_i.$$

Let $y_{it} = j$ if agent i chooses the option j on market m in period t. Then, under the standard logit assumptions, the conditional probability that individual i chooses j, $\mathscr{s}_{ijmt}$, is

$$\mathscr{s}_{ijmt} = P\left(y_{imt} = j | p_{mt}, q_{mt}, X_{imt}; \beta, \xi, \Delta\xi_{mt}\right) = \frac{\exp\left(\bar{u}_{ijmt}\right)}{1 + \sum_{k=1,2} \exp\left(\bar{u}_{ikmt}\right)}. \tag{6}$$

The market share of product j is then the probability that an individual consumes j, averaged over her characteristics X_{imt}; formally $s_{jmt} = \int \mathscr{s}_{ijmt} \mathrm{d}F_X(X_{imt})$. As underlined by Berry et al. (1995), under the logit assumptions, the market-product-specific term δ_{jmt} is equal to $\ln\left(s_{jmt}\right) - \ln\left(s_{0mt}\right)$.

Besides, the conditional own- and cross-price elasticities of these market shares are

$$\eta^p_{ijkmt} = \frac{\partial \mathscr{s}_{ijmt}}{\partial p_{ikmt}} \frac{p_{ikmt}}{\mathscr{s}_{ijmt}} = \begin{cases} \beta_{pj} p_{jmt}(1 - \mathscr{s}_{ijmt}) & \text{if } j = k \\ -\beta_{pj} p_{kmt} \mathscr{s}_{ikmt} & \text{otherwise.} \end{cases} \tag{7}$$

The average price elasticities are therefore given by

$$\eta^p_{jkt} = \begin{cases} \beta_{pj} p_{jmt}(1 - s_{jmt}) & \text{if } j = k \\ -\beta_{pj} p_{kmt} s_{kmt} & \text{otherwise.} \end{cases} \tag{8}$$

Symmetrically, one can define the conditional own- and cross-quality elasticities as

$$\eta^q_{ijkmt} = \frac{\partial \mathscr{s}_{ijmt}}{\partial q_{ikmt}} \frac{q_{ikmt}}{\mathscr{s}_{ijmt}} = \begin{cases} \beta_{qj} q_{jmt}(1 - \mathscr{s}_{ijmt}) & \text{if } j = k \\ -\beta_{qj} q_{kmt} \mathscr{s}_{ikmt} & \text{otherwise.} \end{cases} \tag{9}$$

which yields average elasticities given as follows

$$\eta^q_{jkt} = \begin{cases} \beta_{qj} q_{jmt}(1 - s_{jmt}) & \text{if } j = k \\ -\beta_{qj} q_{kmt} s_{kmt} & \text{otherwise.} \end{cases} \tag{10}$$

4.1.2 Estimation

I estimate consumer valuations for black-market and legal cannabis for *pre-* and *post-legalization*. For both periods, I observe whether individuals used cannabis or not. The subset of agents surveyed in the two periods are denoted respectively by $\mathcal{I}_{pre}$ and $\mathcal{I}_{post}$. Analogously, the corresponding time periods belong to the subsets $\mathcal{T}_{pre}$ and $\mathcal{T}_{post}$.

No recreational cannabis is legally available under prohibition. Therefore I assume that any consumer before legalization is provided by the black market. The log-likelihood of the model for all subjects $i \in \mathcal{I}_{pre}$ living in periods $t \in \mathcal{T}_{pre}$ under prohibition is

$$\mathcal{L}\left(\delta_1, \beta_{X1}\right) = \sum_{\substack{i \in \mathcal{I}_{pre} \\ t \in \mathcal{T}_{pre}}} \mathbb{1}_{[y_{imt}=1]}\left(\delta_{1mt} + \beta_{X1} X_i\right) - \ln\left(1 + \exp\left(\delta_{1mt} + \beta_{X1} X_i\right)\right)$$

The BRFSS data does not distinguish legal from illegal cannabis consumption. Directly evaluating equation (6) during the *post-legalization* period does not enable to disentangle s_{i1mt} from s_{i2mt}. Instead, it only allows to estimate the conditional probability that individual i consumes cannabis $s_{i1mt} + s_{i2mt}$. The log-likelihood of the model for all subjects $i \in \mathcal{I}_{post}$ consuming $j = 1, 2$ in periods $t \in \mathcal{T}_{post}$ *post-legalization* is

$$\begin{aligned}\mathcal{L}\left(\delta, \beta_X\right) = \sum_{\substack{i \in \mathcal{I}_{post} \\ t \in \mathcal{T}_{post}}} &\left\{\mathbb{1}_{[y_{imt}>0]} \times \ln\left(\exp\left(\delta_{1mt} + \beta_{X1} X_i\right) + \exp\left(\delta_{2mt} + \beta_{X2} X_i\right)\right)\right. \\ &\left. - \ln\left(1 + \exp\left(\delta_{1mt} + \beta_{X1} X_i\right) + \exp\left(\delta_{2mt} + \beta_{X2} X_i\right)\right)\right\}\end{aligned}$$

The log-likelihood of the demand for legal and illegal cannabis on the whole sample is simply given by the sum of the log-likelihood functions of the demands for cannabis under prohibition and legalization.

$$\begin{aligned}\mathcal{L}\left(\delta, \beta_X\right) = &\sum_{\substack{i \in \mathcal{I}_{pre} \\ t \in \mathcal{T}_{pre}}} \mathbb{1}_{[y_{imt}=1]}\left(\delta_{1mt} + \beta_{X1} X_i\right) - \ln\left(1 + \exp\left(\delta_{1mt} + \beta_{X1} X_i\right)\right) \\ &+ \sum_{\substack{i \in \mathcal{I}_{post} \\ t \in \mathcal{T}_{post}}} \left\{\mathbb{1}_{[y_{imt}>0]} \times \ln\left(\exp\left(\delta_{1mt} + \beta_{X1} X_i\right) + \exp\left(\delta_{2mt} + \beta_{X2} X_i\right)\right)\right. \\ &\left. - \ln\left(1 + \exp\left(\delta_{1mt} + \beta_{X1} X_i\right) + \exp\left(\delta_{2mt} + \beta_{X2} X_i\right)\right)\right\}\end{aligned} \tag{11}$$

The parameters $\{\delta_{1mt}, \delta_{2mt}\}$, β_{X1} and β_{X2} are to be estimated by Maximum Likelihood (ML).

I assume that the sensitivity parameters β_j are policy invariant, i.e. the parameter $\beta_1 = (\beta_{p1}, \beta_{q1}, \beta_{X1})$ remains unchanged after the implementation of legal retail sales. This implies that the choice of consumers – and substitution between illegal and legal cannabis –

is solely driven by the introduction of a new option, everything else being equal on the black market. [26] Although this assumption imposes some restriction on consumer preferences, it is necessary to allow for identification of consumer sensitivity to characteristics. This paper is the first to estimate preferences for legal and illegal cannabis simultaneously.

The estimates for β_{pj} and β_{qj}, $j = 1, 2$ are retrieved from a standard two-step estimation procedure, which follows Nevo (2001), where the estimates for δ_{jmt} are regressed on the prices p_{jmt} and THC contents q_{jmt}:

$$\hat{\delta}_{jmt} = \beta_{pj} p_{jmt} + \beta_{qj} q_{jmt} + \xi_j + \Delta\xi_{jmt} \tag{12}$$

Potential correlation between prices and unobserved characteristics threaten the consistent estimation of (12). To correct for this source of endogeneity, I therefore use the price on the upstream market as an instrument on the legal price and the proximity to British Columbia as an instrument on the black-market price. Details on instrumental variables are presented in Appendix F.

While strategic responses in prices are immediate, adjustements in quality take time. A natural cannabis crop cycle is a year long. In artificial environments, heavily controlled with refined hydroponic infrastructures and lighting, plants can flower up to 6 times a year. However, changing plants in crops or improving the quality of one's crops otherwise requires relatively more investment and time. For this reason, I consider that quality at time t is predetermined and does not require instruments.[27]

4.1.3 Results

The ML estimates for the parameters β_{Xj}, $j = 1, 2$ from equation (11) are provided in Table 2.6. Unsurprisingly, female and older individuals derive less utility from cannabis consumption, while tobacco smokers are more likely to consume cannabis. The coefficients regarding both products are relatively similar. Interestingly, when market definition does

[26] Under the logit assumption, the own-price elasticity of illegal cannabis only changes through price and quantity, β_{p1} remaining identical. Under prohibition, it is indeed given by $\eta^{pre}_{i1t} = \beta_{p1} p_{1mt} \dfrac{1}{1 + \exp(\bar{u}_{i1mt})}$ After legalization, it becomes $\eta^{post}_{i1mt} = \beta_{p1} p_{1mt} \dfrac{1 + \exp(\bar{u}_{i2mt})}{1 + \exp(\bar{u}_{i1mt}) + \exp(\bar{u}_{i2mt})}$ The same applies for quality elasticities.

[27] In practice, legal local retailers could adjust the quality of their products by sourcing them from different suppliers. In this case, the assumption of exogeneous quality would be violated. However, the State of Washington does not allow any importation of recreational cannabis, whether at the retail level or upstream. Hence such adjustments in quality are limited to the extent of the existence and availability of higher quality products, whose production should be authorized by the WSLCB. Further, in this version the legal and the illegal sectors are for now each modeled as one representative agent, with average THC potency reflecting the whole panel of products at the grower level. Regarding illegal sellers, these are more likely to be subject to sticky contracts, since most transactions on the black market happen between individuals who are already acquainted (Caulkins and Pacula, 2006).

Table 11: Estimated coefficients for individual preferences for cannabis (first-stage ML results)

	(1)	(2)	(3)
X_1			
age	-0.0445***	-0.0405***	-0.0545***
	(0.00106)	(0.000937)	(0.000636)
female	-0.556***	-0.546***	-0.646***
	(0.0365)	(0.0356)	(0.0257)
smoke100	1.554***	1.437***	1.483***
	(0.0408)	(0.0394)	(0.0284)
X_2			
age	-0.0435***	-0.0408***	-0.0482***
	(0.00151)	(0.00135)	(0.000914)
female	-0.530***	-0.518***	-0.515***
	(0.0513)	(0.0499)	(0.0319)
smoke100	1.551***	1.447***	1.601***
	(0.0570)	(0.0556)	(0.0348)
Market definition			
MSA × year	✓	-	-
MSA only	-	✓	-
year only	-	-	✓
N	55,100	55,100	80,948

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Sensitivity of cannabis consumption to price and quality

	(1)	(2)
Sensitivity parameters		
β_{p1}	-0.0291***	-0.0353***
	(0.000790)	(0.000642)
β_{p2}	-0.0353***	-0.0451***
	(0.000821)	(0.000390)
β_{q1}	0.386***	-
	(0.0241)	-
β_{q2}	0.0223***	-
	(0.00556)	-
Average price elasticities		
Prohibition		
η^p_{11}	-0.252	-0.305
	(0.0283)	(0.0343)
Legalization		
η^p_{11}	-0.227	-0.275
	(0.0545)	(0.0661)
η^p_{22}	-0.510	-0.565
	(0.249)	(0.245)
η^p_{12}	0.0293	0.0343
	(0.0284)	(0.0320)
η^p_{21}	0.0200	0.02623
	(0.0220)	(0.0273)
Average quality elasticities		
Prohibition		
η^q_{11}	6.187	-
	(0.544)	-
Legalization		
η^q_{11}	6.237	-
	(0.682)	-
η^q_{22}	0.419	-
	(0.0349)	-
η^q_{12}	-0.562	-
	(0.521)	-
η^q_{21}	- 0.0244	-
	(0.0254)	-
Quality included	✓	-

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

not include MSA, individual distaste for black-market cannabis with regards to age and gender is intensified. In this case, observations related to individuals living in relatively rural – and expectedly more conservative – are actually included in the estimated sample; which could explain this result. This underlines the importance of accounting for geographic disparities.

The market shares predicted by model (1) for the sample are generally consistent, although the model seems to over-estimate the extensive margin of cannabis consumption (see Appendix G).

Table 12 presents the average price elasticities computed using equation (8) and the second-step estimation results (equation 12). I use here the specification (1) of the first-stage model (i.e. with year × MSA fixed effects). Obtained average own-price elasticities for the extensive margin of black-market cannabis consumption are generally between -0.2 and -0.3. Price elasticity of participation to the legal market lies around -0.5. These values are consistent with the results of Jacobi and Sovinsky (2016) on the elasticity of participation to the black market for cannabis, as well as with Hollenbeck and Uetake (2021) estimating higher sensitivity of individuals to the price of legal cannabis than what the literature had measured on the illegal market (see for example Davis et al., 2016). I also find exacerbated sensitivity to quality on the black market (elasticities around 6) relatively to the legal market. On the illegal market, quality is not certified and hence more volatile than on the legal market, which could explain this result. Finally, subtitution between the legal and the illegal sectors following changes in price is very limited, with cross-price elasticities between 0.02 and 0.03. This is not the case for changes in quality. In particular, the THC potency on the legal market rising by 10% causes a 5.62% drop in the demand for illegal cannabis. This suggests quality as a viable tool for the legal market to compete against the black market and drive it out of business.

4.2 The supply for legal and black-market cannabis *post-legalization*

The supply is shared between two sectors: a legal one and an illegal one. The legal sector is composed by a limited number of licensed businesses, which have to comply to local regulations. Prices are affected directly by fiscal requirements, as well as indirectly by licensing, which impacts market concentration. Further, quality and traceability regulations inflate prices by two channels. Resulting cost inflation forces legal retailers to set higher prices. Consumers' willingness to pay for higher quality products enables legal retailers to raise prices. At the other end of the spectrum, the illegal sector abides to no rule. Its price and quality (here measured by THC potency) are set according to the production and distribution costs, the costs related to the business exposure to sanctions, as well as competition dynamics with the legal sector. For the sake of simplicity and due to data limitation, the legal and the illegal sectors are respectively modeled as one representative firm selling a single product. Extensions will account for market concentration within each sector. *Post-legalization*, the legal and the illegal sectors compete playing a two-stage game in which (i) they simultaneously choose their levels of quality q and (ii) given the chosen levels of quality, they choose prices simultaneously. This assumption is consistent with the cannabis one year long crop cycle and the sticky adjustment of THC potency.

The profit function of sector $j = 1, 2$ on market m in period t is

$$\Pi_{jmt}(p_{jmt}, q_{jmt}) = [p_{jmt} - c_j(q_{jmt})] \, s_{jmt} \left(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt} \right)$$

where $\boldsymbol{q}_{mt} = (q_{jmt}, q_{kmt})$ and $\boldsymbol{p}_{mt} = (p_{jmt}, p_{kmt})$, $k \neq j$. Sector j maximizes its profit with respect to price in the second period, $\frac{\partial \Pi_{jmt}}{\partial p_{jmt}} = 0$.

$$s_{jmt} \left(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt} \right) + [p_{jmt} - c_{jmt}(q_{jmt})] \frac{\partial s_{jmt}}{\partial p_{jmt}} \left(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt} \right) = 0 \tag{13}$$

An equilibrium in quality is such that, for the quality of the other sector being given, sector j maximizes its profit. Thus, the level of quality q_{jmt} that maximizes the profit of sector j is such that

$$- c_j'(q_{jmt}) s_{jmt} \left(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt} \right) + [p_{jmt} - c_{jmt}(q_{jmt})] \frac{\partial s_{jmt}}{\partial q_{jmt}} \left(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt} \right) = 0 \tag{14}$$

Retrieving the marginal cost function is necessary in order to analyze counterfactual quality choices. Assume the marginal cost of product j is a function of product quality q_{jmt}, geographical-sector fixed effects θ_{jm}, and a market-time-specific shock ω_{jmt} and can be written as follows[28]

$$\ln c_j \left(q_{jmt}; \theta_{jm} \right) = \gamma_{0j} + \gamma_{1j} q_{jmt} + \omega_{jmt} \tag{15}$$

Under this specification and using the results from the demand estimation $\frac{\partial s_{jmt}}{\partial q_{jmt}} = \beta_{qj} s_{jmt}(1 - s_{jmt})$, I evaluate condition (14) which becomes

$$\ln p_{jmt} = \ln \left(\frac{1}{\beta_{qj}(1 - s_{jmt})} + 1 \right) + \gamma_{0j} + \gamma_{1j} q_{jmt} + \omega_{jmt} \tag{16}$$

The estimation results are presented in Table 13. Their interpretation in terms of marginal cost for medium, strong and very strong products is provided in Table 14.

[28] This is in line with the empirical literature on quality (see Crawford et al., 2019; Fan and Yang, 2020, for example): it specifies quality-convex marginal costs (γ_j is expected to be positive) and hence a profit function concave in quality.

On the black market, estimates for γ_1 confirm that the marginal cost on the illegal market is convex in quality. Prices for medium to very strong cannabis vary from 7.92 USD/g to 8.45 USD/g on the black market under prohibition. Under legalization, the estimate for the baseline parameter γ_0 decreases, but the quality coefficient γ_1 increases. The former confirms the intuition of a drop in operation costs *ex-post*, which could in theory be due to lower risk for illegal suppliers to be detected and hence arrested. The latter is more difficult to interpret. One explanation could be that more potent products *post-legalization* are further differentiated, which inflates their cost. This possibility is consistent with the results of section 3.3, which suggest that black-marker suppliers differentiate very strong products by improving their quality.[29]

On the legal market, the estimated function predicts the marginal cost for medium range cannabis above 19 USD/g. The marginal cost of very strong cannabis would be at 16.7 USD/g. The high difference between these estimates and the ones relating to black-market cannabis reflects the cost burden of the quality and traceability controls implemented by the Washington State Liquor and Cannabis Board. Besides, the value for γ_1 on the legal market being negative is counter-intuitive and requires more investigation. One cause could be the legal market in the State of Washington being subject to heavy traceability regulations. Producing high potency cannabis requires environments where growing conditions are stable and as a result compliance to regulations is easier, hence less costly.

4.3 Policy implications

Throughout the last decade, one of the main objectives of governments legalizing cannabis has been killing the black market. Implementing a legalization policy exclusively aiming at evicting the black market could result in a higher price than under full deregulation and still be successful (see the First Article). In this case, the rise in demand subsequent to legalization can be moderated through a price effect. A government willing to control the demand for cannabis would therefore wish for the prices on the legal market to be relatively high.

Using the estimates on consumers' sensitivity and substitution patterns with regards to price and quality, I compute the black market best response functions, as well as the variations in the demand for legal and illegal cannabis, to changes in price and quality of the legal good. A wide range of policies enable the government to manipulate the price and quality on the legal market, using the three following tools:

(i) imposing an oligopoly structure for the legal retail market through licensing (and eventually setting a limited number of awarded licenses),

[29] While for the sake of simplicity this version proxies quality solely using the THC potency, other aspects might come at play and be paired with increases in THC potency.

Table 13: Marginal cost functions

	BLACK MARKET	LEGAL MARKET
Prohibition		
γ_0	1.95***	-
	(0.0211)	-
γ_1	0.00820***	-
	(0.000958)	-
Legalization		
γ_0	0.939***	3.29***
	(0.0780)	(0.0614)
γ_1	0.0595***	-0.0211***
	(0.00454)	(0.00223)
Geographical f.e.	✓	✓
Year f.e.	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(ii) taxing legal cannabis,

(iii) submitting the legal sector to quality or traceability requirements and controls.

The two first tools have somehow straightforward impacts on the price, while the latter results in a shift in the marginal cost function. Better quality and traceability may also imply higher investments, both before and after the licensing and production phases, resulting in higher cost to enter the legal market. Disentangling entry- *versus* non entry-induced oligopolistic structures is actually a difficult task to undertake. The extent to which state governments limit the number of licenses *per se* is often unclear; so is the cost of complying to the standards imposed on legal retailers prior to being allowed to enter the market. Cannabis in the state of Washington is heavily taxed. On average, tax rates in this sample, which are a compound of state and local taxes, are around 40%, which heavily inflates prices. As underlined in the previous section, this is amplified by strict quality enforcement driving up marginal costs.

I model reaction prices from the black market and show that improvements in quality are essential for a government aiming at eradicating the black market. I consider the price response of the black market to several legalization scenarios, implying variations in both price and quality of the legal product. Since quality adjustments are sticky, on the short

Table 14: Marginal costs for medium, strong and very strong cannabis

	Black market	Legal market
Prohibition		
medium (14.5% THC)	7.92	-
strong (18.5% THC)	8.18	-
very strong (22.5% THC)	8.45	-
Legalization		
medium (14.5% THC)	6.06	19.77
strong (18.5% THC)	7.69	18.17
very strong (22.5% THC)	9.75	16.70

Prices in USD/g come from the estimation results presented in Table 13.

run the black market only reacts in price.[30] Differentiating equation (13) yields

$$\alpha_{pj}\mathrm{d}p_{jmt} + \alpha_{qk}\mathrm{d}q_{kmt} + \alpha_{pk}\mathrm{d}p_{kmt} = 0$$

where

$$\begin{aligned}
\alpha_{pj} =& 2\frac{\partial s_{jmt}}{\partial p_{jmt}}(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt}) + [p_{jmt} - c_{jmt}(q_{jmt})]\frac{\partial^2 s_{jmt}}{\partial p_{jmt}^2}(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt}) \\
\alpha_{qk} =& \frac{\partial s_{jmt}}{\partial q_{kmt}}(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt}) + [p_{jmt} - c_{jmt}(q_{jmt})]\frac{\partial^2 s_{jmt}}{\partial p_{jmt}\partial q_{kmt}}(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt}) \\
\alpha_{pk} =& \frac{\partial s_{jmt}}{\partial p_{kmt}}(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt}) + [p_{jmt} - c_{jmt}(q_{jmt})]\frac{\partial^2 s_{jmt}}{\partial p_{jmt}\partial p_{kmt}}(\boldsymbol{q}_{mt}, \boldsymbol{p}_{mt})
\end{aligned}$$

The best-responses adjustments of sector j to price and quality changes in sector k are hence respectively given by $-\frac{\alpha_{pk}}{\alpha_{pj}}$ and $-\frac{\alpha_{qk}}{\alpha_{pj}}$. Best-responses can then be interpolated lin-

[30] Further versions of this work will include responses in terms of quality, to investigate long run strategic responses of the black market, which should account for consumers' preferences and substitution patterns with respect to quality.

early from the observed *post-legalization* equilibrium using the estimates for the sensitivity parameters[31] and the marginal cost function. This equilibrium is such that the price and potency on the black market are 8.41 USD/g and 17.68% on average. The price and potency on the legal market are 15.47 USD/g et 20.28%. These baseline price and potency equilibria on the legal and the black markets correspond to the average price and potency observed in the data, *post-legalization*. The corresponding marginal cost on the black market is 7.32 USD/g. It is computed from the estimated marginal cost function.

Reaction prices of the black market, from the observed *post-legalization* scenario are represented in Figure 3 by different colors as functions of the price (on the vertical axis) and THC potency (on the horizontal axis) of cannabis set on the legal retail market. The dashed line represents the isoquant of level 7.32, i.e. combinations of price and THC potency of legal cannabis such that the best-response price of the black-market is at marginal cost. Points in green, South-East of the isoquant, are eviction scenarios: the black market does not survive. The darker the color, the lower the price. In the opposite direction, points in red represent reaction prices above the marginal cost.

To eradicate the black market, the legal sector needs to invest in quality improvements. Even with a price at 5 USD/g, which is well below its current marginal cost, the legal sector cannot eradicate the illegal sector, unless it raises THC potency to 22%. Further, setting a high quality on the legal market – for instance a 24% THC potency – enables to get rid of the black market while setting high prices, and thereby curbing use.

5 Conclusion

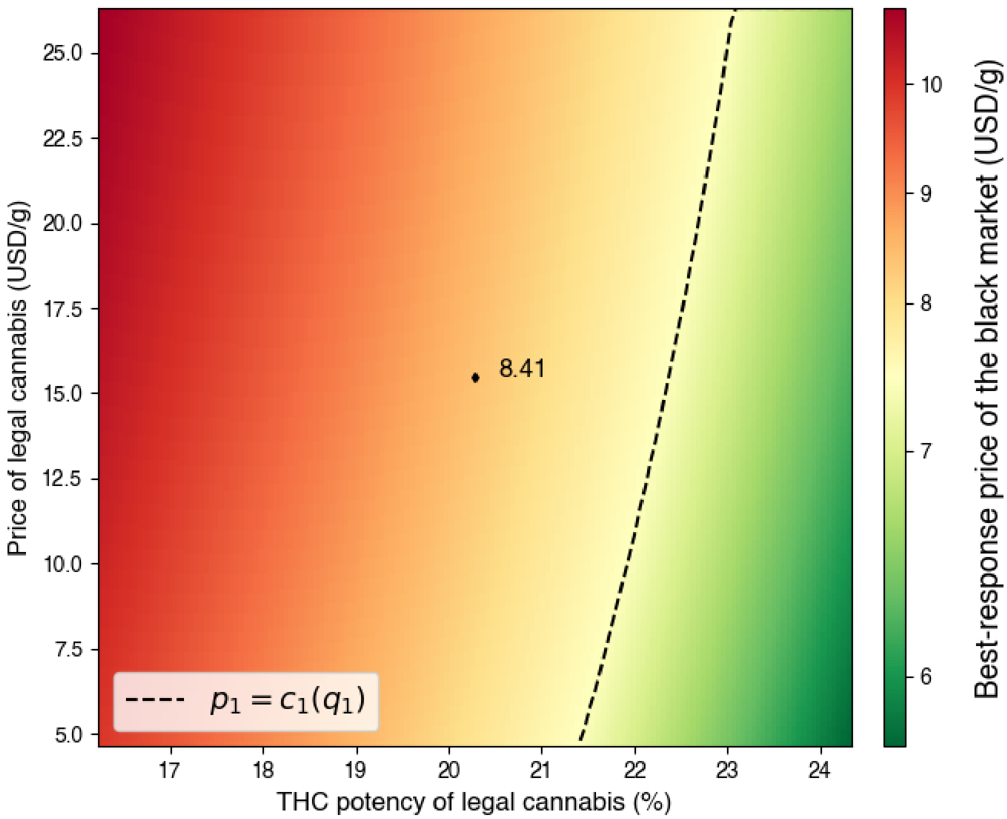
The literature on cannabis has covered illegal consumption behavior under prohibition. Recent papers have documented the legal sector, covering strategic interactions of legal firms with respect to policy and demand sensitivity to prices. Yet, to my knowledge, no previous work has covered market interactions across the legal and the illegal sectors.

Another contribution of this paper relates to the estimation of preferences with respect to quality; a dimension that has been overlooked by the literature. This second contribution is made possible by the exploitation of original crowd-sourced data on black-market prices, that includes information on cannabis strains.

I first focus on average price and quality responses to legalization policies. Reduced-form estimation highlights equilibrium changes on the black market, where prices decrease by up to 20% while THC potency increases by up to 1.4% on average. This effect in price is heterogenous across different levels of THC potency. These results suggest legalization

[31] Recall that under logit assumptions, the first derivatives of market shares are simply given as $\frac{\partial s_{jmt}}{\partial q_{jmt}} = \beta_{qj} s_{jmt}(1 - s_{jmt})$, $\frac{\partial s_{jmt}}{\partial q_{kmt}} = \beta_{qj} s_{jmt} s_{kmt}$, $\frac{\partial s_{jmt}}{\partial p_{jmt}} = \beta_{pj} s_{jmt}(1 - s_{jmt})$, $\frac{\partial s_{jmt}}{\partial p_{kmt}} = \beta_{pj} s_{jmt} s_{kmt}$.

Figure 3: Short-run best-response of the black market with respect to changes in the price and the THC potency of legal retail cannabis



Notes: At equilibrium *post-legalization*, the price and potency on the black market are 8.41 USD/g and 17.68% on average. The corresponding marginal cost is 7.32 USD/g. The price and potency on the legal market are 15.47 USD/g and 20.28%. From this equilibrium, I compute the reaction of the black market to changes in price and THC potency on the legal market. On the short run, the black market responds to the legal market by adjusting its price solely. These reaction prices are obtained iteratively using linear interpolation and differentiating equation (13).

enhances competition on the global market for cannabis; and that retailers' strategy does not only depend on prices, quality matters.

Understanding the role of quality and how it interacts with price responses motivate the evaluation of a random utility demand model that accounts for quality. Estimation yields measures for sensitivity with respect to both price and THC potency. In particular, it presents substitution patterns between the legal and the illegal sectors with respect to both price and quality. While consumers substitution with regards to price is very low, sensitive substitution based on quality presents it as a viable policy tool. Counterfactual analysis computes the black market best-response functions and show that price competition solely can drive illegal retailers out of business, but at the cost of traceability standards.

Eradicating the black market has been a common objective displayed by governments promoting legalization. Yet, the social optimality of underlying outcomes remains to be discussed in further versions of this work. Besides, as is standard in the literature, this work restricts market prices and levels of quality to single dimensions. Further research should account for quantity discounts in price as well as other dimensions for quality, such as product diversity and availibility.

References

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

Auriol, Emmanuelle, Alice Mesnard, and Tiffanie Perrault, "Weeding out the Dealers? The Economics of Cannabis Legalization," *CESifo*, 2020. Working Paper No. 8645.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890.

Brandeis, Jason, "The Continuing Vitality of Ravin v. State: Alaskans Still Have a Constitutional Right to Possess Marijuana in the Privacy of Their Homes," *Alaska Law Review*, 2012, *29*, 175.

Brinkman, Jeffrey and David Mok-Lamme, "Not in my backyard? Not so fast. The effect of marijuana legalization on neighborhood crime," *Regional Science and Urban Economics*, 2019, *78*, 103460.

Carrieri, Vincenzo, Leonardo Madio, and Francesco Principe, "Light cannabis and organized crime: Evidence from (unintended) liberalization in Italy," *European Economic Review*, 2019, *113*, 63 – 76.

Caulkins, Jonathan P, "Estimated cost of production for legalized cannabis," *RAND, Drug Policy Research Center. http://www. rand. org/content/dam/rand/pubs/working_papers/2010/RAND_WR764. pdf*, 2010.

Caulkins, Jonathan P and Rosalie Liccardo Pacula, "Marijuana markets: Inferences from reports by the household population," *Journal of Drug Issues*, 2006, *36* (1), 173–200.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Červenỳ, Jakub and Jan C van Ours, "Cannabis prices on the dark web," *European Economic Review*, 2019, *120*, 103306.

Chang, Tom Y. and Mireille Jacobson, "Going to pot? The impact of dispensary closures on crime," *Journal of Urban Economics*, July 2017, *100* (C), 120–136.

Crawford, Gregory S, Oleksandr Shcherbakov, and Matthew Shum, "Quality overprovision in cable television markets," *American Economic Review*, 2019, *109* (3), 956–95.

Davis, Adam J, Karl R Geisler, and Mark W Nichols, "The price elasticity of marijuana demand: Evidence from crowd-sourced transaction data," *Empirical Economics*, 2016, *50* (4), 1171–1192.

de Chaisemartin, Clément and Xavier d'Haultfœuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–96.

de Chaisemartin, Clément and Xavier d'Haultfœuille, "Two-way Fixed Effects and Differences-in-Differences Estimators with Several Treatments," Technical Report, National Bureau of Economic Research 2022.

Deshpande, Manasi and Yue Li, "Who is screened out? Application costs and the targeting of disability programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–48.

Di Forti, Marta, Diego Quattrone, Tom P Freeman, Giada Tripoli, Charlotte Gayer-Anderson, Harriet Quigley, Victoria Rodriguez, Hannah E Jongsma, Laura Ferraro, Caterina La Cascia et al., "The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-GEI): a multicentre case-control study," *The Lancet Psychiatry*, 2019, *6* (5), 427–436.

Dills, Angela K, Sietse Goffard, and Jeffrey Miron, "The Effects of Marijuana Liberalizations: Evidence from Monitoring the Future," Working Paper 23779, National Bureau of Economic Research September 2017.

Diplock, Jordan, Darryl Plecas, and Fire Chief Len Garis, "Targeting marihuana growing operations in British Columbia: a summary report highlighting current research rindings," *Centre for Public Safety & Criminal Justice Research, University of the Fraser Valley, Abbotsford, BC, Canada*, 2013.

Dragone, Davide, Giovanni Prarolo, Paolo Vanin, and Giulio Zanella, "Crime and the legalization of recreational marijuana," *Journal of Economic Behavior & Organization*, 2019, *159*, 488 – 501.

Fan, Ying and Chenyu Yang, "Competition, product proliferation, and welfare: A study of the US smartphone market," *American Economic Journal: Microeconomics*, 2020, *12* (2), 99–134.

Galenianos, Manolis and Alessandro Gavazza, "A structural model of the retail market for illicit drugs," *American Economic Review*, 2017, *107* (3), 858–96.

Gavrilova, Evelina, Takuma Kamada, and Floris Zoutman, "Is legal pot crippling Mexican drug trafficking organisations? The effect of medical marijuana laws on US crime," *The Economic Journal*, 2019, *129* (617), 375–407.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár, "On Estimating Multiple Treatment Effects with Regression," *arXiv preprint arXiv:2106.05024*, 2021.

Goodhart, Charles and Jonathan Ashworth, "Canadian Legalization of Cannabis reduces both its cash usage and 'Black' Economy," 2019.

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Hansen, Benjamin, Keaton Miller, and Caroline Weber, "The taxation of recreational marijuana: Evidence from Washington state," Technical Report, National Bureau of Economic Research 2017.

Hollenbeck, Brett and Kosuke Uetake, "Taxation and market power in the legal marijuana industry," *The RAND Journal of Economics*, 2021, *52* (3), 559–595.

Hull, Peter, "Estimating treatment effects in mover designs," *arXiv preprint arXiv:1804.06721*, 2018.

Jacobi, Liana and Michelle Sovinsky, "Marijuana on main street? Estimating demand in markets with limited access," *American Economic Review*, 2016, *106* (8), 2009–45.

Jacobson, Mireille, "Baby booms and drug busts: Trends in youth drug use in the United States, 1975–2000," *The Quarterly Journal of Economics*, 2004, *119* (4), 1481–1512.

Miller, Austin M., Robert Rosenman, and Benjamin W. Cowan, "Recreational marijuana legalization and college student use: Early evidence," *SSM - Population Health*, 2017, *3*, 649 – 657.

Mills, Evan, "The carbon footprint of indoor Cannabis production," *Energy Policy*, 2012, *46*, 58–67.

Morris, Robert G., Michael TenEyck, J. C. Barnes, and Tomislav V. Kovandzic, "The Effect of Medical Marijuana Laws on Crime: Evidence from State Panel Data, 1990-2006," *PLOS ONE*, March 2014, *9*, 1–7.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

Thomas, Danna, "License quotas and the inefficient regulation of sin goods: Evidence from the Washington recreational marijuana market," *Available at SSRN 3312960*, 2019.

van Ours, Jan C. and Jenny Williams, "Cannabis prices and dynamics of cannabis use," *Journal of Health Economics*, 2007, *26* (3), 578 – 596.

Zheng, Zhonghua, Kelsey Fiddes, and Liangcheng Yang, "A narrative review on environmental impacts of cannabis cultivation," *Journal of Cannabis Research*, 2021, *3* (1), 1–10.

Appendix

A Cannabis laws in the U.S.

Along with the societal changes and the increase in cannabis use associated to the seventies, numerous states proceeded to a wave of decriminalization. California, Colorado, Maine, Minnesota, Mississippi, Nebraska, New York, North Carolina, Ohio, Oregon and Washington declassified possession of small amounts of cannabis (usually up to 1 ounce) to a misdemeanor. In 1975, Alaska declared possession of small amounts of cannabis to be protected under state constitutional right to privacy. However, the intensification of the *War on Drugs* in the eighties left this liberalization process stagnating. The rising concerns about the efficiency of this costly war led to a new wave of policy changes at the edge of the 21st century. Initiated by a second wave of decriminalization laws and the first laws in favor of medical use, this liberalization movement accelerated in the last decade. In 2012, Colorado and Washington states passing bills legalizing recreational use of cannabis after a referendum. From 2014 onward, these states would be imitated by seventeen other American states and the District of Columbia. Legalization policies implemented so far are quite diverse. As of February 2023, while nineteen states and the District of Columbia have legalized the use of recreational cannabis, possessing this commodity remains a felony in other states such as Arizona. Not only the legal status differs across states, but sanctions and fine levels are far from uniform between two states having the same cannabis laws. For instance, Arizona state law would not provide any guideline for punishment regarding small amounts of cannabis; possessing up to 2 pounds of cannabis entails a risk of incarceration of up to 2 years and a maximum fine of USD 150,000. In Alabama, possessing any amount of cannabis is punishable by up to 1 year of incarceration along with a maximum fine of USD 6,000. In contrast, Virginia sets a threshold for possession of small amounts at 1/2 oz and sanctions it by no more than 30 days of incarceration and a fine of USD 500 on a first offense. Possessing up to 42.5 g in Minnesota would only entail a risk of a USD 200 fine. Such diversity being observed across a single territory makes the United-States a preferred case of study for analyzing the impacts of cannabis policies.

The table below is borrowed from Auriol et al. (2020). It provides a global overview of the state of american cannabis regulation, highlighting its disparity. For each state, column 2 reports the year during which cannabis was decriminalized. Column 3 provides the year of the first ballot to implement a *Medical Marijuana Law* (MML), while the fourth column records the year during which such a law was passed. Analogously, the year of the first recreational cannabis legalization ballot is recorded in the fifth column; its passing is given in column 6. The last column reports the year of the first legal retail sales of cannabis.

Dashes materialize the absence of the event described in the corresponding column.

State	Decrim.	1st MML ballot	MML	1st rec. ballot	Rec.	Retail
AL	-	-[a]	2021	-	-	-
AK	1975[b]	1998	1998	2000	2014	2016
AZ	-	1996	2010	2016	2020	2021
AR	-[c]	2012	2016	2022	-	-
CA	1975	1996	1996	1972	2016	2018
CO	1975	2000	2000	2012	2012	2014
CT	2011	-[a]	2012	-[d]	2021	2023
DE	2015	- [a]	2011	-	-	-
D.C.	2014	1998	2010	2014	2014	-[e]
FL	-[f]	2014	2016	-[g]	-	-
GA	-[f]	-	-[h]	-	-	-
HI	2020	- [a]	2000	-	-	-
ID	-	-	-	-	-	-
IL	2016	- [a]	2013	-[d]	2019	2020
IN	-[i]	-	-	-	-	-
IA	-	-	-	-	-	-
KS	-	-	-	-	-	-
KY	-[f]	-	-[j]	-	-	-

[a] *Medical Marijuana* was not on the ballot: instead, it was signed into law after legislative approval.

[b]Alaska issued a cannabis decriminalization bill on May 16, 1975, which is two weeks before the famous *Ravin* decision, protecting the possession of small amounts under constitutional privacy rights, was issued. Decriminalization of cannabis came into effect on June 5, 1975. The timeline of cannabis policy in Alaska then becomes fuzzy: further decriminalization was billed in 1982, then cannabis was recriminalized in 1990, decriminalized in 2003, then recriminalized in 2006; while the *Ravin* caselaw would still interact with the criminal state law (Brandeis, 2012). Legalization approved in 2014 ended this confusion.

[c]Although cannabis use remains a crime under state law, it is decriminalized locally.

[d] The recreational use of cannabis was not on the ballot: instead, it was signed into law after legislative approval.

[e]Implementation still pending.

[f] Although cannabis use remains a crime under state law, it is decriminalized locally.

[g] A cannabis legalization initiative was expected to be on the ballot in November 2022 and is now expected for November 2024 ("Marijuana on the ballot", *Ballotpedia*. Retrieved online June 2022 and February 2023, `https://ballotpedia.org/Marijuana_on_the_ballot`)

[h]A bill was passed in 2015, legalizing the use of *light cannabis*, i.e. cannabis products featuring low THC potency (see Georgia General Assembly, `https://www.legis.ga.gov/legislation/42674`).

[i]Decriminalized in Marion County as of 2019 (see `https://web.archive.org/web/20190930193952/https://www.wthr.com/article/marion-county-will-no-longer-prosecute-simple-marijuana-cases`).

[j]A *Medical Marijuana* bill was presented to the House of Kentucky in January 2020. It is presently under evaluation by the Senate Judiciary Committee (Kentucky General Assembly, *House Bill 136*; retrieved online

State	Decrim.	1st MML ballot	MML	1st rec. ballot	Rec.	Retail
LA	2021	- [a]	2015[k]	-	-	-
ME	1975	1999	1999	2016	2016	2020
MD	2014	- [a]	2013	2022	2022	-[l]
MA	2008	2012	2012	2016	2016	2018
MI	2018	2008	2008	2018	2018	2019
MN	1976	- [a]	2014	-	-[m]	-
MS	1978	2020	2020	-	-	-
MO	2014	2018	2018	2022	2022	2023
MT	-[f]	2004	2004	2020	2020	2022
NE	1979	-[n]	-	-	-	-
NV	2016	1998	1998	2006	2016	2017
NH	2017	- [a]	2013	-	-	-
NJ	-	-[a]	2010	2020	2020	2022
NM	2019	-[a]	2007	[d]	2021	2022
NY	1977	- [a]	2014	-[d]	2021	2022
NC	1977	-	-	-	-	-
ND	2019	2016	2016	2018	-	-
OH	1975	- [a]	2016	2015	-	-
OK	-[o]	2018	2018	-[p]	-	-
OR	1973	1998	1998	2012	2014	2015
PA	-[f]	- [a]	2016	-	-	-
RI	2012	- [a]	2005	-[d]	2022	-
SC	-	-	-	-	-	-

3rd December 2020, url: https://apps.legislature.ky.gov/record/20rs/hb136.html).

[k]Although *Medical Marijuana* was signed into law in 2015, it was unlawful to inhale cannabis until 2019 (see https://www.mpp.org/states/louisiana/overview-of-louisianas-medical-cannabis-law/).

[l]Expected July 2023.

[m]In January 2023, the Minnesota House of Representatives introduced bill HF 100, which plans the legalization and regulation of adult-use cannabis (Minnesota House of Representatives, *HF 100*; retrieved online 8th February 2023, url:https://wdoc.house.leg.state.mn.us/leg/LS93/HF0100.0.pdf.

[n] A *Medical Marijuana* ballot is expected to be on the ballot in November 2022 ("Marijuana on the ballot", *Ballotpedia*. Retrieved online June 2022, https://ballotpedia.org/Marijuana_on_the_ballot).

[o]A cannabis decriminalization initiative is expected to be on the ballot in November 2022 ("Oklahoma State Question 812, Marijuana Decriminalization Initiative (2022)", retrieved online on Ballotpedia; url: https://ballotpedia.org/Oklahoma_State_Question_812,_Marijuana_Decriminalization_Initiative_(2022)).

[p] A cannabis legalization initiative was expected to be on the ballot in November 2022 and is now expected for March 2023 ("Marijuana on the ballot", *Ballotpedia*. Retrieved online June 2022 and February 2023, https://ballotpedia.org/Marijuana_on_the_ballot)

State	Decrim.	1st MML ballot	MML	1st rec. ballot	Rec.	Retail
SD	-	2006	2020	2020	-[q]	-
TN	-	-	-	-	-	-
TX	-[f]	-	-	-	-	-
UT	-	2018	2018	-	-	-
VT	2013	- [a]	2004	- [d]	2018	2020
VA	-	-	-	-[d]	2021	- [r]
WA	2012	1998	1998	2012	2012	2014
WV	-	-	2017[s]	-	-	-
WI	-[f]	-	-	-	-	-
WY	-[t]	-	-	-	-	-

B Data cleaning and processing

B.1 Geographical matching and aggregation

Geographical markets were defined using the Metropolitan and Micopolitan Statistical Areas (MMSA) division established by the US Census.

While this information is directly available in the BRFSS data, matching it with the THMQ and the WSLCB data required some processing.

Observations in the THMQ data are given by city – sometimes county or general area – and state. I geocoded these observations and cleaned their associated addresses by scraping Open Street Map's Nominatim. The cleaned, detailed, addresses provided me with the county for each location.

To match geographical areas with prices listed in the WSLCB *seed-to-sale* data, I first follow the same procedure as in Hollenbeck and Uetake (2021); which consists in retrieving the list of license applications from the Washington State Liquor and Cannabis Board website, as well as their history though the Internet Wayback Machine. As previously, I then clean the addresses obtained and assign them to counties by scraping Open Street Map's Nominatim.

The lists of detailed locations obtained was then merged with the US Census list of statistical divisions.

[q]The recreational use of cannabis was legalized by the 2020 ballot. However, in 2021, the South Dakota Supreme Court ruled the amendment responsible for the legalization of recreational as unconstitutional.

[r]Expected in 2024.

[s]Although a bill regulating medical use of cannabis was signed in April 2017, Medical Marijuana Laws were not implemented in West Virginia before 2019.

[t]Expected to be on the ballot in 2024 ("Marijuana on the ballot", *Ballotpedia*. Retrieved online June 2022, https://ballotpedia.org/Marijuana_on_the_ballot

B.2 Associating strains with THC potencies

The THMQ data provides information on strains. The dataset I collected accounts for more than 2,000 different values of *strain*. To exploit this information, I scraped the strain repertory of `leafly.com` from which I recovered the THC potency, plant type (indica, sativa or hybrid), as well as the different appellations of each strain. I matched this list with the THMQ data. When possible, I used exact matching on strain names and alternative appellations. I paired remaining observations to the repertory items to which they were the closest, in terms of Jaro-Wrinkler distance. I discarded pairs for which the Jaro-Wrinkler metric was less than 75%.

C Robustness to staggered policy adoption

This appendix provides the results of the diagnostic test `twowayfeweights` described in de Chaisemartin and d'Haultfœuille (2020), which I apply to check the robustness of the TWFE estimator to treatment heterogeneity in my data.

Table 16 reports the percentage of negative weights associated to ATT estimates, as well as in brackets the sum of these negative weights and in braces (for the single treatment cases) the minimal value of the standard deviation of the treatment effect across the treated groups and time periods under which $\hat{\beta}^{fe}$ is compatible with a data generating process (DGP) where the average of those ATT estimates is 0, which I further denote $\underline{\sigma}_{fe}$. These are computed for three specifications, assuming:

(i) the effect of the legalization of recreational use was computed without controlling for the implementation of retail sales;

(ii) the effect of the implementation of retail sales was computed without controlling for the legalization of recreational cannabis use;

(iii) the effect of the implementation of retail sales was computed controlling for the legalization of recreational cannabis use.

Columns (1)-(3) describe the results for the outcome variable being the logarithm of price per once, columns (4)-(6) for the logarithm of the THC potency and columns (7)-(9) for the logarithm of quality adjusted price, which is the difference between the two previous outcomes.

There are no negative weights under specifications (i) and (ii). Therefore, for these single treatment cases, should all the ATT effects be of the same sign, the TWFE estimator

Table 16: Diagnostic tests for specifications (i)-(iii)

	Price			THC			Quality adjusted price		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
legal	0 %		46.20%	0%		46.84%	0%		46.84%
	[0.00]		[-1.00]	[0.00]		[-1.00]	[0.00]		[-1.00]
	{0.95}		-	{0.01}		-	{0.68}		-
retail		0%	0%		0%	0%		0%	0%
		[0.00]	[0.00]		[0.00]	[0.00]		[0.00]	[0.00]
		{1.63}	-		{0.12}	-		{1.33}	-

This table reports the percentage of all ATT estimates that display a negative weight, as well as in brackets the sum of negative weights attached to the TWFE estimators and in braces (for the single treatment cases) the minimal value of the standard deviation of the treatment effect across the treated groups and time periods under which $\hat{\beta}^{TWFE}$ is compatible with a data generating process where the average of those ATT estimates is 0. These figures are obtained running the `twowayfeweights` Stata command described in de Chaisemartin and d'Haultfœuille (2020). The outcome variable in columns (1)-(3) is the logarithm of the price per ounce. In columns (4)-(6), it is the logarithm of the THC potency and in columns (7)-(9) the logarithm of the quality adjusted price, i.e. the difference between the logarithms of the price per ounce and the THC potency. Columns (1), (4) and (7) relate to the specification (i) where the effect of the legalization of cannabis use for recreational purposes was estimated without controlling for the implementation of retail sales. Columns (2), (5) and (8) relate to the specification (ii) where the effect of implementing retail sales was estimated without controlling for the legalization of recreational cannabis use. Columns (3), (6) and (9) relate to the specification (iii) where the effect of the legalization of cannabis use for recreational purposes was estimated while controlling for the implementation of retail sales.

has the same sign as the causal effect. Moreover, the value $\underline{\sigma}_{fe}$ is relatively large for these specifications. It is reasonable to state the average of the ATT effects is unlikely zero.[32]

Under specification (iii), more than 40% of the weights related to the estimate of the parameter for *legal* are negative, which raises the issue of contamination between the two treatments.

Table 17 reports the results of the same diagnostic tests as above, for the following specifications:

(i′) the effect of the legalization of recreational use prior to the implementation of retail sales, i.e. on the subset of observations for which *retail* is zero;

(ii′) the effect of the implementation of retail sales was once recreational cannabis use is legal, i.e. on the subset of observations for which *legal* is one.

[32] As shown later in table 1, the effect of policies on THC potency is relatively small. Hence, the seemingly small values for $\underline{\sigma}_{fe}$ obtained in the case of THC potency remain large enough such that the average ATT is unlikely zero.

Table 17: Diagnostic tests for specifications (i′) and (ii′)

	Price		THC		Quality adjusted price	
	(1)	(2)	(3)	(4)	(5)	(6)
legal	0 %		0%		0%	
	[0.00]		[0.00]		[0.00]	
	{2.13}		{0.06}		{1.42}	
retail		28.25%		32.74%		28.89%
		[-0.38]		[-0.38]		[-0.38]
		{0.02, 0.05}		{0.005, 0.01}		{0.01, 0.03}

This table reports the percentage of all ATT estimates that display a negative weight, as well as in brackets the sum of negative weights attached to the TWFE estimators and in braces the minimal value of the standard deviation of the treatment effect across the treated groups and time periods under which $\hat{\beta}^{TWFE}$ is compatible with a data generating process where the average of those ATT estimates is 0 (first element) or of opposite sign (second element, if any). These figures are obtained running the `twowayfeweights` Stata command described in de Chaisemartin and d'Haultfœuille (2020). The outcome variable in columns (1)-(2) is the logarithm of the price per ounce. In columns (3)-(4), it is the logarithm of the THC potency and in columns (5)-(6) the logarithm of the quality adjusted price, i.e. the difference between the logarithms of the price per ounce and the THC potency. Columns (1), (3) and (5) relate to the specification (i′) where the effect of the legalization of cannabis use for recreational purposes was estimated on the subset for which $retail = 0$. Columns (2), (4) and (6) relate to the specification (ii′) where the effect of implementing retail sales was estimated on the subset for which $legal = 1$.

While all weights related to the treatment *legal* are positive, nearly a third of weights for the treatment *retail* are negative, although their sum is around -0.4. This raises the concern of a possible average zero treatment effect or of a treatment effect of opposite sign for the treatment *legal* – relative to the results from estimating the TWFE model.

D Avorted cannabis reforms

The TWFE results in Section 3 suggest that legalizing cannabis and regulating its market yields a sustainable decrease of the black-market price and a rise in product THC potency. To support the argument of a causal effect of legalization and retail sales for recreational cannabis on the black-market equilibrium price and potency, I provide TWFE results on avorted legalization attempts.

These attempts are modeled using two variables:

- *no successful ballot* describes a situation where a state has put the legal use of recreational cannabis on the ballot but this initiative never resulted in legalization;

- *failed ballot* describes a ballot initiative that was not followed by the legalization of recreational cannabis within two years.

Table 18: OLS estimates of the TWFE model of the effects of unsuccessful legalization attempts

	Price		THC		Quality adjusted price	
No successful ballot	0.0319		0.000496		0.0627***	
	(0.0220)		(0.00366)		(0.0164)	
Failed ballot		-0.118		-0.0113		-0.0651
		(0.108)		(0.0118)		(0.0888)
N	8,373	8,373	7,272	7,272	7,272	7,272

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (2) and (3) of Table 18 provide results regarding prices, columns (4) and (5) focus on THC potency, while the two last columns give estimates for price relative to potency. Line 2 reports the coefficient obtained from regressing the binary indicator *no successful ballot* on the outcomes of interest. Line 3 gives the estimates from regressing *failed ballot* on the outcomes of interest. Lines 4 and 5 specify the fixed effects used.

I find in general no significant effect of failed cannabis ballots on the black-market price and quality of cannabis. The exception is the effect of having no successful ballot on the quality adjusted price, the understanding of which would require further investigation.

E Heterogenous responses to legalization reforms

This section aims at providing comparison between several states on how the equilibrum price and quality on the black market for cannabis evolve after legalization. I generalize the TWFE model described previously and use an event-study type of analysis, in which I compare states affected by legalization to states that were never treated in my data.[33]

This exercise also allows to distinguish short-term effects from long-term effects of cannabis policies on the illegal markets. Shedding light on the permanence of such responses provides hints about the temporality of the market responses of both the illegal supply and the demand; and how fast one they would adapt to structural changes in the cannabis market.

[33]Restricting the analysis to the five groups described in section 3.2 drastically reduces the number of observations, which is why I conduct separate regressions, each time comparing one treated state to all never treated states.

Consider the following econometric model:[34] [35]

$$y_{ist} = \sum_{\tau=-q}^{m} \beta_\tau D_{st}^\tau + \theta_s + \psi_t + \epsilon_{ist} \tag{17}$$

The D_{st}^τ are a series of "event-time" binary variables that equal one when the legalization policy is implemented τ quarters away in state s; formally:

$$D_{st}^\tau \equiv \begin{cases} \mathbb{1}\left[3(\tau-1)+1 \leq t-e_s \leq 3\tau\right] & \text{, if } \tau \geq 1 \\ \mathbb{1}\left[3\tau \leq t-e_s \leq 3(\tau+1)-1\right] & \text{, if } \tau \leq -1 \end{cases} \tag{18}$$

with e_s being the time at which legalization came into effect in state s.

The coefficients $(\beta_\tau)_{\tau\in\{-q,...,m\}}$ estimate the time path of the average price per once of cannabis before ($\tau = -q, ..., -1$) and after ($\tau = 1, ..., m$) legal recreational use of cannabis is implemented ($t = e_s$), conditional on state- and year- fixed effects. Legalization being randomly implemented, conditional on the fixed effects, implies that legalization should not be preceded on average by any geographical-specific trend in average cannabis prices. Formally:

$$\beta_\tau = 0, \forall \tau < 0 \tag{19}$$

I estimate the model described by equation (17) using ordinary least squares, including a set of event-time binary variables along with binary variables for the state and year-fixed effects. Standard errors are clustered at the state level, to correct for eventual intra-state correlation. In the presence of geographical fixed effects, all the coefficients β_τ are perfectly collinear. For this reason, I restrict the estimation to a window covering 12 months before and up to 24 months after the date of policy implementation;[36] formally $\tau \in \{-4, ..., -1, 1, ..., 8\}$. Further, I impose $\beta_{-1} = 0$, so that all post-treatment coefficients should be thought as treatment effects.

My data is an unbalanced panel, in which some states are more represented than others, and covers dates until February 2019. Estimating model (17), I compare the effects of legalization on price and potency in states for which I have a sufficient number of observations before and after the policy change: Colorado, California, Maine, Massachusetts, Nevada, Oregon and Washington. Figures 4 and 5 describe the results of these estimations.

[34] The variables y_{ist}, θ_s, ψ_t and ϵ_{ist} follow the same notation as in equation (1).

[35] Given the number of observations, I chose to use fixed effects at the year level, in contrast to a finer level. This decision is also motivated by the fact that most ballots are voted in November, which would cause month or quarter effects to be correlated with the binary variables describing legalization policies.

[36] Because the treatments *legal* and *retail* are likely contaminated, as previously, I restrict the analysis of the effect of legalization on subsamples for which retail sales have not yet been implemented. Conversely, I restrict the analysis of the implementation of retail sales on the subsample for which legalization has already taken place.

The results clearly show that responses to legalization differ from one state to the next. While there is a clear immediate and substantial decrease in prices in Oregon, Washington and Massachusetts, the effect is more mitigated in Maine and seems smaller and somehow delayed in Colorado. Further, the dynamics of the price effects seem different from one state to the next: some states seem to endure lasting drops in prices while other feature a more temporary shock. On THC potency, tendencies are more mitigated and difficult to interpret.

F Instrumental variables

Estimating equation (12) requires instruments on prices, which are likely correlated to the unobservable heterogeneity $\Delta\xi_{jmt}$ and thereby endogenous.

Instruments on black-market prices

I exploit the geographical proximity between the State of Washington and British Columbia. The instrumental variables on the black-market prices are the driving distance to the nearest border point in British Columbia, computed using Google Maps API, the annual exchange rate between the US and the Canadian dollars, as well as an interaction between these two variables. The Canadian province has indeed been a significant cannabis producer, the sector especially thriving at the turn of the 21st century, in terms of both size and sophistication (Diplock et al., 2013). Assume the composition of local markets are subject to their distance to British Columbia. In this case, relative geographical position affects local black-market prices. Further, as highlighted by the results of Section 3.1 and Table 9, the reaction of black-market prices to policy changes varies across product categories.

Instruments on legal prices

The WSLCB data includes information on upstream transactions. Each retail item is associated with detailed information on the wholesale batch from which it originates. I use the upstream price associated to p_{2t}, denoted p_{2t}^{up}, as an intrument on the price p_{2t}.

In the state of Washington, commercial prices are set freely by retailers, who decide of the profit margin they obtain from re-selling the upstream product. Therefore, the upstream price of a given product influences its retail price. Note that the legal cannabis industry is regulated by the WSLCB. Independent cannabis growers, processors and retailers can apply for state business licenses. Retailers are not allowed to hold a processor or grower license simultaneously. The number of licenses awarded is controlled by the state: licenses are attributed to qualified applications based on a lottery. T he density of retail stores

Figure 4: Dynamic effect of the legalization of recreational cannabis on its black-market price: comparing Colorado, Maine, Massachusetts, Oregon and Washington

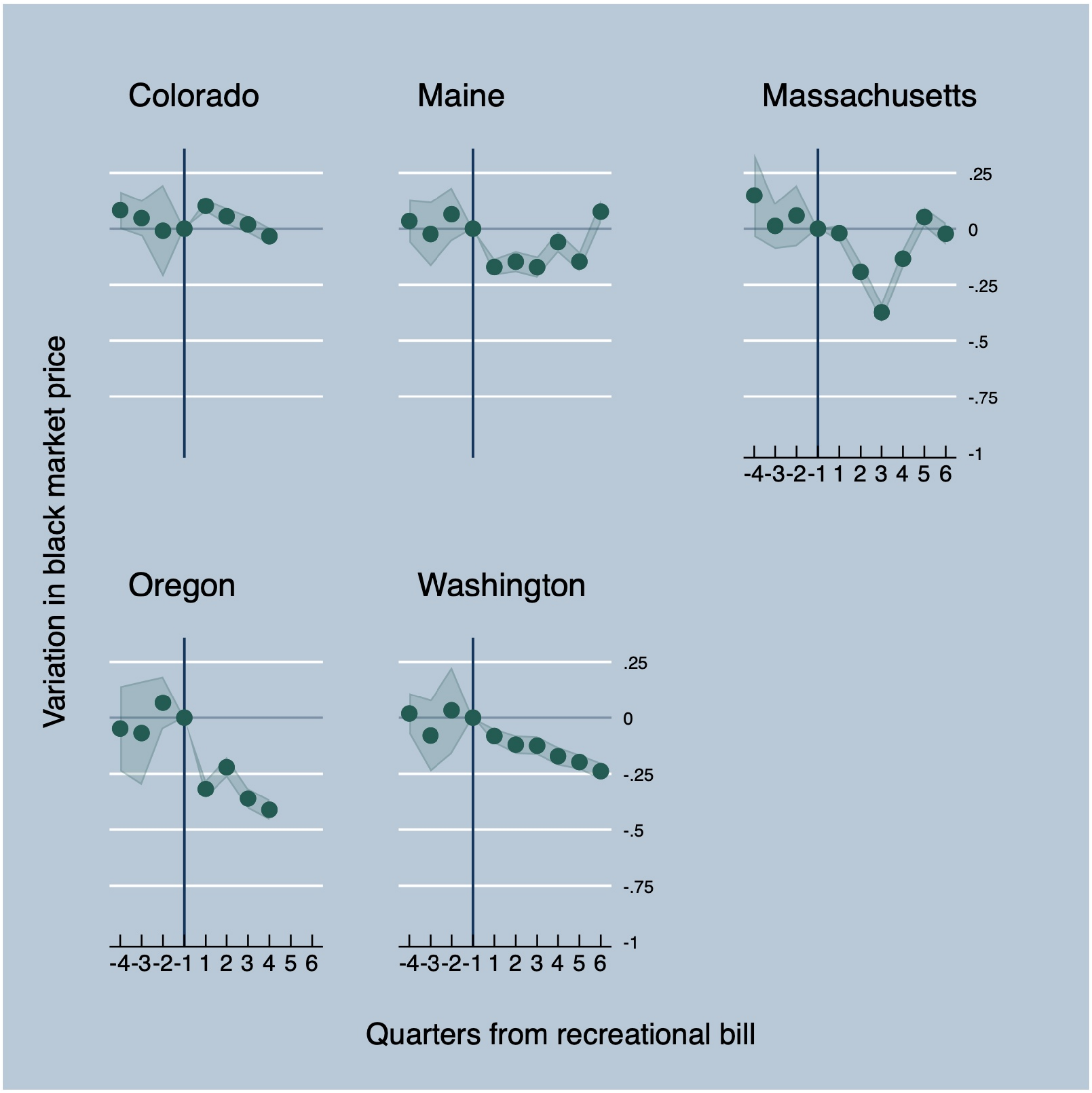
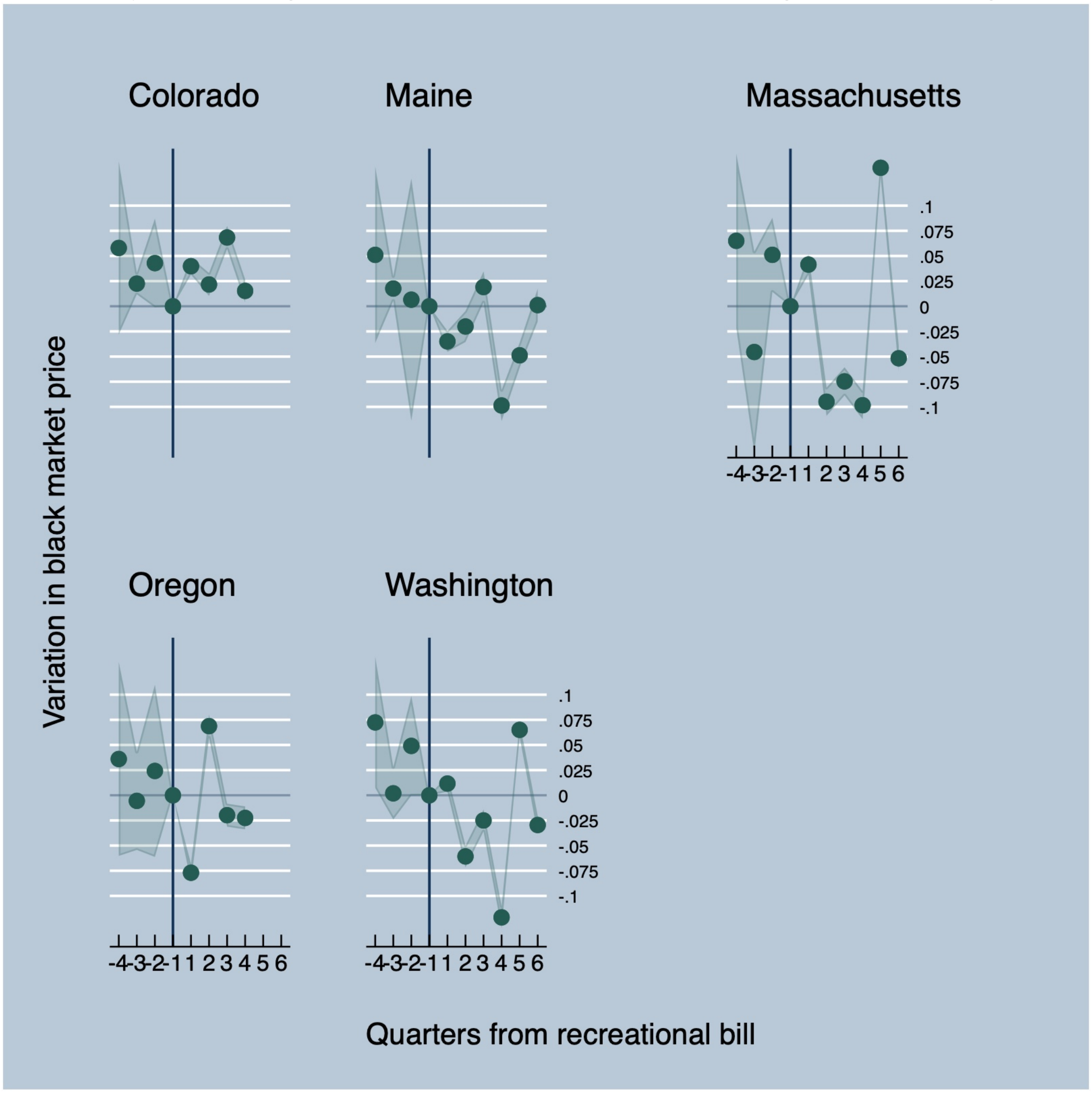


Figure 5: Dynamic effect of the legalization of recreational cannabis on its black-market THC potency: comparing Colorado, Maine, Massachusetts, Oregon and Washington



is not constant across locations. Further, retail sales of cannabis are subject to relatively high (37%) state taxes, as well as further local taxes. These, combined with the oligopoly structure of the market enable the government to manipulate the retail prices. Besides, the extent to which prices are inflated by policy varies from one location to the next. These features of the WSLCB regulation enable one to discard the concern of upstream prices being perfectly collinear with retail prices.

G First-stage estimation: predicted market shares

Table 19: Observed and estimated extensive margins of cannabis consumption

Good	Under prohibition		After legalization	
	s_j	$\hat{s}_j$	s_j	$\hat{s}_j$
0	94.29%	94.12%	89.29%	85.58%
1	5.712%	5.885%	10.71%	7.213%
2	-	-		7.210%