

**Alternative Bases for Limiting Cannabis Production**

Mark A. R. Kleiman, BOTEC Analysis, UCLA

BOTEC Analysis Corp.

I-502 Project #430-5e

Final

June 28, 2013

## Table of Contents

Why limit production? _____	3
Options _____	4
Advantages of using THC as the basis of limitation _____	5
Design details _____	7
<i>Determining the total initial quota</i> _____	7
<i>Dividing the quota among licensees</i> _____	8
<i>“Banking” and trading</i> _____	8
<i>Adjusting the quota over time</i> _____	9
Monitoring and enforcement _____	9
Extensions _____	9
<i>Adjusting for chemicals with protective value</i> _____	9
<i>Using quota reduction as a regulatory penalty</i> _____	10
<i>Substituting a quota auction for taxation</i> _____	10

## Why limit production?

In designing a regulatory system to implement I-502, the Liquor Control Board needs to decide whether to limit licensed cannabis production, and, if so, on which one or more of several alternative bases. Production limits (“quotas”) are a standard feature of the regulation of Schedule II controlled substances such as oxycodone and are designed to discourage diversion by preventing the licit production of more product than can be licitly sold. In contrast, quotas play no part in alcohol regulation.

Overproduction of ordinary commodities is not a matter of public concern. If more of some good is produced than can be sold at current prices, prices will fall until the quantity demanded at some new price is equal to the quantity supplied at that price. If that new price is below the costs of production, firms will either voluntarily reduce the amount they make or simply go out of business when they run out of money.

There are, however, three distinct sets of reasons why production limits may be desirable for legal cannabis: They may (1) reduce the financial incentive for illegal sales out of state; (2) constrain consumption by subpopulations in which unfettered consumption is not desirable and (3) limit the brand, market, and political power of large producers.

There are three subpopulations whose consumption we may want to limit: underage users, dependent users, and out-of-state users. Nationally, those under the age of 21 account for about one quarter of cannabis consumption. Underage consumption would remain illegal under the proposed initiative. But, as the alcohol market demonstrates, intoxicants “leak” to minors from licit adult markets. This may actually be preferable, on balance, to continuing to have the underage market served by purely illicit production: the product reaching minors by diversion from legal channels will be tested for impurities and labeled for chemical content; diverters may be less likely to offer their young customers other drugs than dealers who buy from illegal growers; and there will be less money available to criminal organizations. To the extent that underage users are supplied by adults buying in the licit market, licit prices will influence the price at which cannabis is available to minors, and – since consumption by the young is of particular public concern and since young people tend to be more price-sensitive than adults – there is value in keeping those prices from falling dramatically.

The board should also confront the fact that some consumers will not be fully volitional participants in the market because their cannabis use is not entirely under their control. It is important to recognize that a substantial fraction of the market—measured in terms of volume rather than in terms of numbers of users—consists of people who have become dependent on cannabis. This is reflected in both in self-reports in national surveys and by admissions to drug treatment programs. The National Survey on Drug Use and Health estimates that nearly eight million people say that they have tried to cut down on their marijuana use in the last year, and between one and two million report “often” consuming more than they intend to consume. These problem users are disproportionately high-volume users. Heavy users (like younger users) tend to be more price-sensitive than occasional users, simply because cannabis forms a larger share of their household budgets. A price decrease would be expected to have a larger impact on heavy use than it would on casual, non-problem use.

The national market is many, many times larger than the Washington market. Licensed producers, processors, and sellers who find themselves with excess inventory may be tempted to divert some of that inventory for sale in the untaxed and unregulated market, including sale for shipment out of state. “Exports,” even if entirely accidental, might make the federal government less tolerant of allowing the state’s experiment to proceed. The risk of planned diversion by producers—deliberately growing for untaxed sale—will always be present; the risk of unplanned diversion to dispose of otherwise unsalable product may be greater in the early phases of the market, while participants are still learning about how much they can sell.

In addition to limiting total cannabis production, it might well be desirable to limit the production of each participant to prevent the domination of production by one or a few firms, as happens with mass-market beer, cigarettes, and breakfast cereals. Economies of scale might make it difficult for “craft” producers to compete with “mass-market” producers, perhaps limiting the variety of products available and stifling innovation. If early regulatory decisions push the market in the direction of concentration, those effects may prove to be enduring.

There are many reasons why state interests might be better served by an array including small-and medium-sized producers rather than one or a few giants. A competitive industry might be more responsive to customers. Large firms might be more likely to attract federal enforcement attention. Moreover, larger firms might be more capable than smaller ones of carrying out expensive marketing campaigns (with potentially unfortunate effects on drug abuse and sales to minors) and lobbying efforts (limiting the state’s capacity to effectively tax and regulate the cannabis industry). On the other hand, a more concentrated market might prove easier to regulate.

Of course limiting production has disadvantages as well as advantages. It will impose costs on producers, consumers, and regulators, and it might lead to attempts at evasion and the need for enforcement. It could also leave more room for wholly illegal production, partially frustrating the goals of shrinking the illicit market, protecting public health, and generating tax revenue. This memo will concentrate on alternative possible forms of quantity limitation.

## **Options**

Production limits could be defined in several different ways, including:

- Production area
- Power consumption
- Plant count
- Dry weight of salable product
- Amount of active chemicals in product delivered for processing
- Amount of effective THC in product delivered for processing

Combinations of these metrics are also possible. For example, limiting production by a single firm both in terms of area and number of plants. Limits could be different for indoor vs. greenhouse production, as with Canadian medical-marijuana regulations.

Some of these approaches require more precise definition or need to be adjusted for different circumstances. For example:

- What counts as “production area”? Should it be the entire area of the production facility, or the area behind the security cordon, or the area of the rooms dedicated to growing, or the area actually covered by plants? The LCB would need to decide whether greenhouse space and indoor space count equally, and how to count multiple levels of plants.
- Is “power consumption” measured in maximum power used (in watts), or by average power use (kilowatt-hours) over some period of time? What “power consumption” should be assigned to a greenhouse?
- What counts as a “plant”? A seedling? A cutting? A plant in the vegetative state? A flowering plant? And is the limit on the number of plants at any given time, or on the number of plants harvested during some time period?
- If “trim” and “shake” are processed into extracts, how does the dry weight of an extract count compared to the dry weight of flowers?
- What chemicals count as “active”? And do they all count the same, gram for gram?
- Does THC include THC-acid?

In addition to problems of definition, there may be problems of monitoring. Average power use can be read off the meter, but measuring installed capacity requires some additional equipment, and an unscrupulous producer could supplement power from an unmetered source. Choosing a production metric that makes monitoring easy saves public resources and enhances perceived fairness.

An ideal control system would not dramatically distort production choices or create opportunities for “gaming.” The mix of products on the market should be dictated by consumers’ desires and the technical facts of the production process, not by the need to squeeze as much value as possible into a given area or number of plants. Limiting power consumption would have attractive environmental results, but would also give greenhouse growers an edge, unless some power consumption were attributed to the use of the sun.

### **Advantages of using THC as the basis of limitation**

Cannabis comes in multiple forms and varieties: flowers, resin, “shake and trim,” extracts, and cannabis-infused products. Chemical content would seem to be the natural metric for comparing one lot with another: a gram of butane hash oil at 75% cannabinoid content is more like six grams of 12.5% flowers than it is like one gram of those flowers. Even if the rules require some admixture to convert an extract into an infused product, the result

could still be something with much more psychoactive potency per unit weight than what is described by the law as “useable marijuana.”

Cannabis contains many active chemicals, and, like wine, is valued for its aesthetic as well as its psychoactive properties. Still, THC is doubtless the main mood-altering chemical in cannabis, and most consumers buy cannabis in search of mood alteration. In the current markets—both illicit and medical—high-THC cannabis trades at a substantial premium to lower-THC cannabis. The relationship is not perfect, but price seems to be roughly proportional to potency, with some allowance for taste, appearance, and fragrance. There have been accusations that testing paid for by producers and sellers is sometimes “tilted” to increase reported potency and thus perceived value, but the trend toward increased potency is also reflected in enforcement data.

Since 2000, the markets have moved decisively toward higher-THC product. Some of this may reflect illegal producers’ desire to maximize value per unit of gross weight where gross weight is a factor in enforcement targeting and sentencing decisions, or to maximize the value that can be produced from a given number of plants, though the same phenomenon has been observed in the Dutch market. There are also demand-side factors at work; in addition to the influence of fashion (many consumers confuse THC content with quality), some people using cannabis for medical purposes may need, or think they need, high-potency material, and to the extent that very heavy users – the backbone of the market in terms of volume and revenue – develop tolerance to the effects of THC, they too may want very potent cannabis.

In addition to THC concentration, the presence of other cannabinoids, especially cannabidiol (CBD) help shape the effects of any cannabis preparation. CBD is a partial antagonist to THC (competing for CB-1 receptor sites) and may also offer independent “buffering” effects, with possible protective action with respect to anxiety, memory loss, and the risk of psychotic episodes. In the plant, CBD production competes with THC production for chemical precursors. As a result, cannabis breeders eager to maximize THC have tended to create strains with minimal CBD production; the ratio of THC to CBD in seized product has increased from an estimated 6:1 twenty years ago to 200:1 or more today. (Commercial testers report that some popular strains have no measurable CBD content.)

There are reasons to think that both high THC content and high THC:CBD ratios increase the risks of acute ill effects from cannabis use. This is both because greater concentration leads to a sharper “spike” of chemicals getting to the brain (when the drug is administered by lung rather than orally or topically) and because of the absence of CBD’s buffering effects. If this is so, regulations should be designed to discourage producers from continuing the trend of increasing THC content and THC:CBD ratios, or at least designed to avoid inadvertently encouraging the continuation of that trend. New and returning users, unfamiliar with the potency of much contemporary cannabis, might be especially vulnerable to episodes of anxiety, dysphoria, and even panic if most of the available product consists of very potent material.

Limiting the number of plants, dry weight produced, or area under cultivation could have precisely undesirable effect of furthering the trend toward high-THC, low-CBD

cannabis. Regulations making total effective THC a scarce resource producers need to conserve would, by contrast, create a desirable counter-pressure. The challenge to each producer would then be to squeeze as much consumer value as possible into each unit of THC, rather than to squeeze as much THC as possible into each bud, plant, or square foot.

A THC quota system is not the only approach to creating disincentives for excessive potency; taxes proportional to THC content (or, more aggressively, higher taxes per unit THC as THC concentration rises) could have the same effect. But such approaches would require changes in the I-502 legislation.

## **Design details**

### *Determining the total initial quota*

The problem of determining the first year's quota will confront any quantity-limitation approach, whether the limits are denominated in numbers of plants, dry weight of useable marijuana, production area, power consumption, or amount of THC.

The initial quota should be calculated to fully support the expected sales from Board-licensed retailers: not so low as to lead to shortages of licit product (keeping consumers in the parallel markets) nor so high as to drive licit prices down to levels that might lead to large increases in consumption or create financial incentives for illicit out-of-state sales (either via diversion before final retail sale or via "smurfing": purchase at retail in Washington for export to other states).

There is little reason to think that the mere existence of a licit source of supply will greatly increase the quantity of cannabis consumed compared to previous-year levels; even if there are substantial numbers of new users, few of them will quickly become heavy users, and it is heavy users that largely determine the total sales volume. Thus the estimate of current annual consumption levels now being developed by RAND can serve as a baseline for computing the projected sales of Board-licensed retailers in their first year of operation.

But it would be unduly optimistic to imagine that Board-licensed retailers could immediately capture the entire cannabis market. Some consumers accustomed to buying from medical "access points" or from purely illicit dealers will continue to patronize their current suppliers. Some of those who now grow their own plants (whether legally under medical recommendations or illegally) will maintain that practice. The RAND online survey of current daily or near-daily users will provide some insight as to their intentions with respect to the new Board-regulated market.

The rate at which current consumers convert from their existing sources into the Board-regulated market will depend on several factors: prices; perceived quality and reliability (certified testing and labeling should be advantageous); regulatory changes (e.g., a tightening of the medical-recommendation process would tend to deflect some consumers toward licensed retailers); and enforcement (illicit growers and sellers could be competitively disadvantaged by intensifying police activity against them). At first, licensed, regulated, and taxed retailers will almost certainly charge higher prices than the unlicensed, unregulated, and untaxed "access points" for medical marijuana. The medical system will be a larger factor of the market west of the Cascades (in the jurisdiction of the U.S. Attorney

for the Western District) than it will be in Eastern Washington, where federal authorities have been much more vigorous in suppressing open sales in the medical market. The rate at which current suppliers to the medical system (both growers and retailers) migrate to the commercial system will help drive the rate of consumer migration; a study could be made now of their intentions to help inform the estimate of initial volume for the Board-regulated system.

### *Dividing the quota among licensees*

The initial quota will need to be allocated among prospective market participants. Since the quota is a maximum rather than a minimum, the interest of the producer will always be to have a greater rather than a smaller quota, unless a greater quota is somehow made costly.

Economies of scale in production, distribution, and promotion might tend to push the market toward a small number of large producers. This outcome might be undesirable: it would tend to create an industry with greater marketing and lobbying power; disadvantage smaller, locally-owned producers; and favor large enterprises without local roots; or it may reduce consumer choice and product innovation. To the extent that the Board desires to avoid that outcome a quota system could be used to protect smaller enterprises, either by limiting the fraction of total permitted production allowed to any one producer or by setting aside some quota for producers below some specified size or sizes. There might, for example, be three quota levels, for small, medium, and large enterprises, with each producer of a given level having the same quota; allotting (say) half the quota to small and medium-sized producers would put an upper bound on market concentration at the producer level.

What might lead a firm to apply for a “small” or “medium” license rather than the largest license available? The Board could require each applicant to show that it has the capacity to produce the quota it applies for, or have stricter security rules and other regulatory requirements for larger producers than for smaller producers.

### *“Banking” and trading*

In addition to the market uncertainties confronting any business, cannabis growers confront technical uncertainty as to the outcomes of the growing process. They can decide how much production area to use or how many plants to grow, but both the gross weight of product and its cannabinoid content are subject to harvest-to-harvest variation only partly under the grower’s control. Thus an output-based quota system (whether denominated in gross weight or in THC content) could generate waste in the form of product that is made unsalable because it would exceed the producer’s quota.

To some extent, this problem could be managed at the firm level. If quota is allocated on an annual basis, then greater-than-planned early harvests could be compensated for by cutting back production later in the year. But of course that won’t work for the final growing cycle.



To avoid that problem, producers could be allowed to “bank” their quotas to a limited extent: either “saving” unused quota for a subsequent year or “borrowing” against future limits. In addition, firms could be allowed to buy and sell quota to each other, so that a firm that over-produced in any given year could purchase rights from a firm that had under-produced. Rules could be crafted to avoid pure “quota farming” by requiring that all firms have actual productive capacity and by limiting the share of a firm’s quota that it could buy or sell in any year.

### *Adjusting the quota over time*

Any estimate of first-year sales is subject to uncertainty. Any downward adjustment of quota in the first year would tend to bring strong protest from industry participants who relied on the initial quota to make investments. It is possible that strong sales toward the beginning of the year will demonstrate the need to adjust quotas upward.

In any case, sales in second and subsequent years will presumably be higher, unless the commercial system fails to gain market share from the two competing systems. It is possible, but not certain, that sales will also tend to rise over time, if the legal availability of cannabis of known composition increases total demand or reduces price. It will be necessary to derive an estimate of expected sales for each year, long enough before the start of the year to allow growers to plan for the coming year's production.

### **Monitoring and enforcement**

Since producers can sell only to licensed processors, and since each lot has to be registered, tested for chemical content, and entered into a tracking system, there would be no added burden on processors to compute and report the gross weight or total THC content of what they buy from each producer. If quotas are denominated in gross weight of usable marijuana, some formula would have to be derived to incorporate sales of “trim” and “shake” for extraction (whether extracts are approved for sale or allowed only as inputs to infused products).

Thus the tracking system can prevent over-quota sales within the licensed system; monitoring and inspection will be required to prevent sales outside that system, but those would be required whether quotas are imposed or not.

### **Extensions**

#### *Adjusting for chemicals with protective value*

Insofar as cannabidiol and other chemicals in the plant can be shown to have protective effects against the risks associated with large amounts of THC, a quota system of the type described here could be further adjusted by giving “credit” against THC quotas for the presence of protective chemicals. For example, instead of counting 100 grams of 10%-THC cannabis that also contains 3% CBD using up 10 grams’ worth of THC quota, it could be counted as only 7 grams (in effect, treating CBD as “negative THC” for quota purposes).

Such a system of measuring what might be called “effective THC content” would not have to be precisely accurate to have desirable incentive effects, but work would need to be done to determine which chemicals to count and how much to count them.

#### *Using quota reduction as a regulatory penalty*

Any regulatory system must deal with the problem of rule-breaking, both deliberate and inadvertent. Some violations, or patterns of violations, are serious enough to demand license revocation. But a less serious violation calls for less drastic sanctions. In principle, these could be monetary penalties, but the more common response is a temporary suspension (usually measured in days or weeks) of a seller’s license.

Such an approach could be applied to marijuana retailers without the need for much adaptation. However, the temporary shutdown of a processor would be more complex, since the retailers supplied by that processor might find themselves damaged through no act of their own. But the really difficult problem would come in the case of licensed producers, because there’s no way to tell plants to stop growing, or to stop needing water and light. So unless the Board can erect and defend a system of monetary penalties significant enough to act as deterrents, misconduct by growers will remain difficult to deter, let alone to prove and punish.

A quota system provides a potential solution to that problem; loss of production quota (like temporary closure) imposes a fairly easily quantified economic penalty. Assuming that the overall target previously set still seemed optimal, the Board could either redistribute the quota across the remaining producers as a bonus or (if permitted by law) offer it for sale in the “trading” market described above.

#### *Substituting a quota auction for taxation*

I-502 writes into law a set of tax rates. If, as predicted by the RAND Drug Policy Research Center, legal-production costs (and retail prices) tend to fall over time toward levels much lower than current illicit or medical-market prices, those taxes may become inadequate to sustain the goals of preventing an upsurge in heavy use or use among juveniles and of preventing illicit export to other states. Monitoring and inspection should be adequate to keep the problem of diversion from licensees to a minimum; the likely problem would be “smurfing.” The problem of “smurfing” would arise only if retail prices in Washington fell below wholesale prices for illicitly-produced cannabis in other markets.

In principle, prices could be maintained at any desired level by adjusting tax rates. But that approach puts an unreasonable strain on the legislative process.

An alternative to adjusting tax rates would be to use a quota system to limit volumes, allowing prices to adjust in the market. In effect, quota-holders would collect as economic “rent” the difference between the cost of production and the market price. That difference would be reflected in the price to purchase credits in the quota-trading market.

There seems no clear reason why that economic rent should accrue to quota-holders rather than to taxpayers. Moreover, the Board would likely face a flood of applications for

now-valuable quota rights, and have to choose between denying new applicants to protect existing quota-holders and cutting back on everyone's quotas to allow new entrants.

The economically logical alternative would be to conduct an annual auction for quota rights. (Again, this could be done no matter what units were used to set quotas.) Assuming an efficient auction mechanism, the entire economic surplus would be captured as revenue for the taxpayers of Washington without any need for frequent tax-rate changes.

A quota auction could also provide a mechanism to discourage an outcome where a few large producers dominate the market. Requiring producers with large quotas to pay a premium, above market, price during the auction could counter-balance economies of scale and allow smaller "craft" producers to remain competitive with large firms.

The annual decision on how much quota to issue would balance several aspects of the public interest: the interest in keeping prices high enough to reduce drug abuse, use by minors, and diversion; the interest in keeping prices low enough to discourage illicit production and sale; and the state's revenue needs. Calculating a precisely optimal value would be virtually impossible. But it would also be unnecessary: small deviations from optimality would have only small social costs, while getting the answer roughly right would present a manageable estimation problem once the initial disruption caused by the creation of a legal market had passed and year-to-year market changes had therefore slowed down.