Big Data on a Big New Market:
Insights from Suppliers and Customers in Washington State's Legal Cannabis Market

Jonathan P. Caulkins, Yilun Bao, Imane Fahli, Yutian Guo, Krista Kinnard, Mary Najewicz, & Lauren Renaud

Abstract:
Introduction: Eight U.S. states have legalized large-scale commercial production of cannabis for non-medical use. All require some form of "seed-to-sale" tracking system, which creates a unique opportunity for researchers to study market activity at a heretofore unimagined level of detail. These legal markets also create a range of new matters for policy makers to address.
Data: Publicly available data were obtained on approximately 45 million individually-priced items purchased in the 35 million retail transactions that took place during the first 2+ years of Washington State's legal cannabis market. Product-related measures include product type (flower, extract, lotion, liquid edible, etc.), product name, price, and potency with respect to multiple cannabinoids. Items sold can be traced back up the supply chain through the store to the processor and producer, to the level of identifying the specific production batch and mother plant, and the firm that performed the product testing, as well as test results.
Method: A variety of data visualization methods are employed to locate and describe spatial-temporal patterns of multiple correlated attributes (e.g., price and potency) broken down by product. Text-analytics methods are used to subdivide the broad category of "extracts for inhalation" into more homogeneous sub-categories. Difference-in-difference methods are used to investigate variation between college towns and the rest of the state across natural experiments created by the beginning and ending of summer vacation.
Results: Prices are falling steadily and proportionally at the processor and retailer levels. Potency is high, averaging over 20% for flowers and over 60% of extracts, the two largest product categories; it appears that more than half of the sales of extracts for inhalation are in forms suitable for dabbing, as opposed to vaping or smoking. THC potency is highly variable, CBD-products have minimal market share, and although a wide range of edibles are sold, they account for a modest share of consumer spending. There are some indications that edibles may be relatively more preferred by college students.
Discussion: Preliminary results suggest that the state-legal cannabis market in Washington is highly dynamic in terms of pricing, products, and organization.

Acknowledgements: This paper was developed in close collaboration with a RAND Drug Policy Research Center team led by Beau Kilmer and has benefitted considerably from their comments and suggestions.

Keywords: Cannabis, extracts, potency, prices, legalization, drug policy

Correspondence to: Jon Caulkins, Heinz College, Carnegie Mellon University, Pittsburgh, PA, 15237, USA. Email: caulkins@andrew.cmu.edu

Introduction

Over the past several years, policies surrounding the production, sale and use of cannabis have become less strict in many countries in Europe and the Americas. In November 2014, voters in the U.S. States of Colorado and Washington approved propositions making them the first jurisdictions to state-legalize not just home cultivation, possession, and use, but also large-scale commercial production, distribution, and sale of cannabis products for recreational use. After a period of regulatory design, the first licensed stores opened in January 2014 (in Colorado) and July 2014 (in Washington).

In Washington, the State Liquor and Cannabis Board (LCB) licenses and regulates the producers, processors, laboratories and retailers. The LCB also manages a "seed-to-sale" database tracking cannabis products from growers to laboratory testing, processors and ultimately retail stores. This tracking provides a wealth of data that can inform policy makers of the overall dynamics of the market, elucidate the demand for certain products and consumer markets (*Medical Marijuana FAQ's*, 2016; *Marijuana retailers - medical endorsement*, 2016; *Weekly Marijuana Report, December 12, 2016*, 2016).

Considerable interest from policymakers has focused on the potency of cannabis products and their retail prices. Even before legalization, high-potency products were becoming increasingly popular (Mehmedic et al., 2010; Ben Lakhdar et al., 2016). Policymakers are concerned about the possible health and addictive impacts of increases in potency, especially considering the variety of methods in which higher potency products can be consumed, including dabbing, vaping and edibles (Loflin M., Earleywine M., Krauss, et al., 2015). A recent RAND study shows that for traditional cannabis flowers ("usable marijuana" in Washington State parlance), potency positively affects price (Smart et al., in submission). Additionally, multiple studies have also looked at price sensitivity in both illicit and legal markets (Pacula, 2014; Ben Lakhdar et al., 2016). Some analysis has been completed on the cannabis market in Washington, including estimating the market demand for legal cannabis (Kilmer, 2013). While this past research looks into price and potency, much of their information was based on the illicit marijuana market or the very early stages of legal markets.

The seed-to-sale tracking system is not a sample; it provides the universe of data on all legal commerce in the state, which has generated about 25GB of data over the past three years. This analysis takes advantage of the depth and breadth of these data to explore details of this new legal market that were previously difficult to understand, which, in turn, can inform research that aids policy makers in decision making. This paper explores three trends in particular: 1. Partitioning of the broad product category "extracts for inhalation" into more insightful subgroupings, 2. The relationship between wholesale and retail prices for usable marijuana and extracts for inhalation, and 3. Evaluation of seasonality in sales in college towns, as a window into consumption patterns of this particularly interesting consumer market segment. These analyses help form a preliminary understanding of marijuana product types, market price mechanism and a special group of consumers and provide clues for future analysis.

Data and Measures

Washington State's "seed-to-sale" tracking system was designed to capture all transactions and conversions of marijuana products as they move from producers to processors to labs and retail stores.

The standard unit of analysis in this paper is perhaps most properly called an "item-entry" not a "transaction" because one purchase can produce multiple observations in the data set (Smart et al., in submission). For example, if a customer bought two grams of one type of cannabis flower and one gram of another at the same time, those would show up as separate observations. However, multiple copies of the same item can appear within a single observation in this data set. For example, if that person bought two separate one gram packages of the first type of flower, that could appear as a single $20 observation with a "usable weight" of 2 grams and a '2' in the "weight" field which, for retail transactions indicates the number of items in that item-entry. Nonetheless, for brevity we will abbreviate "item-entry" to "item" in the sequel.

Each observation reports the price paid to the seller by the buyer, whether the buyer is a retail consumer, store owner, processor, etc. In July 2015, Washington changed from a 25% tax at each step of the production process to a single retail excise tax of 37% (*Taxes Due on Marijuana*; 2016). The pre-July 2015 observations include those taxes, while the post-July 2015 data do not. Following Smart et al. (in submission), we bump up retail prices after July 2015 by 37% to consistently reflect the effective cost to the buyer. Prices are expressed in dollars per gram, calculated as the sale price divided by the usable weight of the cannabis. Price mark ups are calculated by comparing the price per gram a retailer paid for a product and the price it was sold for to a consumer.

Potency is defined as the amount of "Total THC" in the product. This variable is calculated as plain THC, which is not decarboxylated, plus 0.877 times THC-A, which is decarboxylated (Smart et al., in submission).

The data set contains a variable "inventorytype" that distinguishes thirty-three product types. This analysis primarily focuses on the most common three: "usable marijuana" which refers to traditional flower with minimal processing, "edibles" which refers to cannabis infused food and drink products, and "extracts for inhalation" which refers to a wide range of processed cannabis products, including but not limited to wax, kief, shatter, oils, and distillates for portable vaporizers.

One comment on methods. We frequently do not test below for the statistical significance of differences for the simple reason that the data represent the *universe* of all legal transactions in the state. They are not a sample drawn from some larger population. So, for example, when we report that the proportion of all extracts in two college towns (Pullman and Airway Heights) that are of the shatter/wax/dab variety is much higher (~65%) than in some other places (24-35%), we do not compute a p-value to assess whether those differences are statistically significant. Those proportions are simply facts, not sample statistics.

Analysis and Results

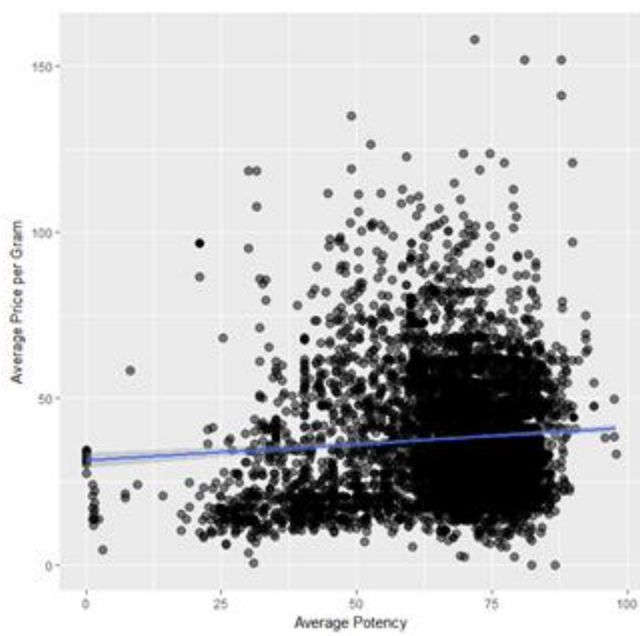
Partitioning Extracts for Inhalation

One prominent trend in the Washington legal cannabis market observed by Smart et al. (in submission) among others is the increasing proportion of the market that is made up of extracts for inhalation (hereinafter "extracts"). While as of 2016 about 75% of transactions continue to be flower transactions, the second largest part of the market is extracts at 12% of transactions. This differs from some other legal marijuana markets such as Colorado where extracts appear to be less popular (Daniulaitye, 2015).

Smart et al. (in submission) analyze trends in price and the relationship between potency and price for usable marijuana (i.e., flower) products but not extracts, in part because the extracts category is a heterogeneous amalgam of different product types. Below, we extend portions of that analysis to extracts for inhalation both as a broad class and broken down into some constituent parts.

Washington's seed-to-sale database does not distinguish between different extract products. For example, cartridges and wax are both listed as inhalants, though they differ in terms of price, potency and modality of use (Morean M.E., 2015; Krauss and Sowles, 2015). Figure 1 plots the average price vs. average potency for each product type within the extract category in the month of June 2016. There is one plotting point for each unique product name, not one point for each item sold. Few patterns can be discerned; therefore, we try to exploit the presence or absence of key words in the free text product name field to find more homogeneous sub-groups within this broad product category (after excluding the relatively small number of high-CBD observations, defined here as those with greater than 1% CBD, which tend to be more expensive and have lower THC potency than do the THC products).

Figure 1: Statewide Average Price and Potency Levels of Extract Products in June 2016



Based on the literature, inspection of store menus, and reading the most common product names, we identified eight subcategories that seemed usually to be distinct (i.e., relatively few product names spanned categories). For each we developed a set of search words. E.g., “shatter” or “budder” for shatter type products. It is important to note that we used a search function that accounted for a “wildcard” in front and behind the search word. This insures that when we search on “wax” we can grab products called “earwax” and “wax-pucks”.

In general, only products whose names included search words for one category and not search words associated with any other category were categorized. However, special rules were written for some common overlaps. E.g., “hash oil” was placed with other oils, and “cartridge oil” with other cartridge observations. However, a product that has both “hash” and “cartridge” in its name is left uncategorized since it is unclear what kind of product it is. Some examples include “X-Tracted Hashplant Wax .5g”, “Fire Alien OG Live Resin Wax (0.5g)”, and “DD Purple Jolly Rancher Dab Oil .5g”. (Each of these examples had fewer than 20 transactions in June 2016.)

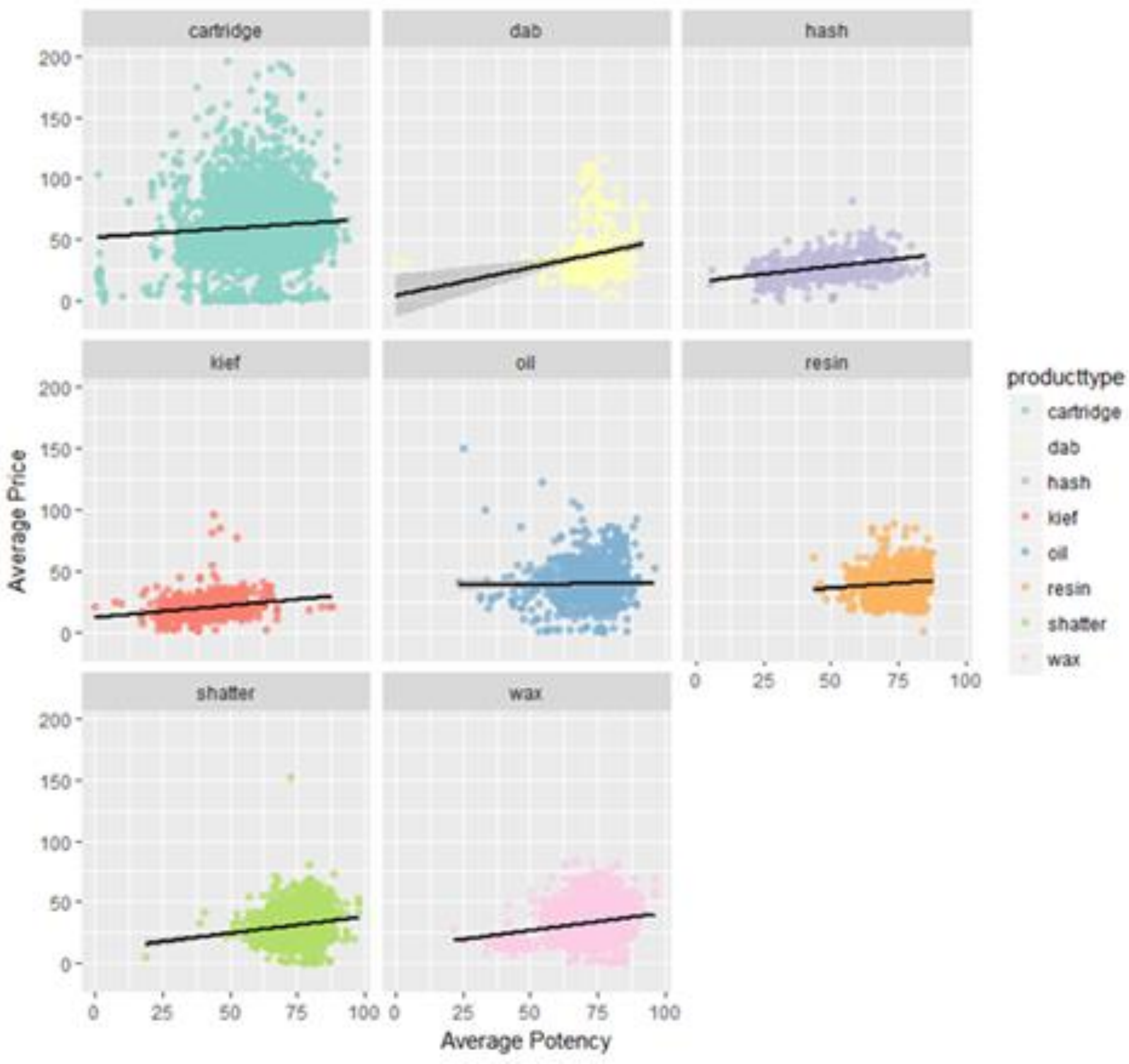
Approximately 74% of all product names (accounting for about 63% of all extract transactions) could be placed within a specific product type, with the remaining observations in the “other” group. See Table 1.

Table 1: Categorization schema and three most popular product names within each category based on June 2016 retail sales

Category	**Search word(s)**	**Three most common products within that category**	**Proportion of obs.**
Cartridge	cart, vape, pen, vc, refill	The Clear Cartridge; Liberty Reach, 0.5g PURE Vaporizer Cartridge, Blue Dream; Willy's Wonder .5ml Cartridges	22%
Oil	oil, rso, eso	Jesus (.5g) Oil; Berry Haze (.5g) Oil; Pineapple Super Silver Haze (.5g) Oil	3%
Hash	hash	Bubble Hash .5g; Sugar Hash - 1g; Monk Hash - 1g	1%
Kief	kief, keif	Kief, BSH Kief 1g; BSH Kief 1g	3%
Dab	dab	Lucid Dabs (1g); Dabulators 0.25g; Dabz - Mt Rainier #10;	1%
Wax	wax, budder	Wax 1g; Blue Dream SugarWax; Supergirl Wax 1g	19%
Shatter	Shatter, crumble	1g Girl Scout Crumble (grow state); Concentrate: BHO Shatter 1g; Wa Woo Cookie Shatter	11%
Resin	resin, rosin	Pineapple Express Live Resin (.5g); Tangie Live Resin (.5G); Middlefork Live Resin (.5g)	4%
Uncategorized	NA	Dutch Hawaiian Frost R.IO6013z 0.5g Atomizer; The Clear Concentrate; Jedi Kush	36%

Figure 2 replicates the display in Figure 1 for the eight separate categories of inhalant products. The average price per gram for an inhalant product is around $55. This is substantially higher than the average price per gram for flower products. Smart et al. (forthcoming) reports average potency for flower products of just over 20%, vs. 67% THC for extracts. With this new categorization some patterns in price and potency now emerge.

Figure 2: Average price and potency for extract products broken down by extract category, June 2016



We explored trends over time in price and potency for all nine categories including “other” (not shown). Broadly speaking, categories that looked similar in Figure 2’s price vs. potency scatter plot also tended to have similar trends in price and potency over time. For example, hash and kief both experienced decreases in potency over time. Another example is that wax, shatter, and resin all appear to have high potency rates that are stable over time.

Based on these patterns, and also the way these products are often presented in store menus, three larger groupings emerge. It seems sensible to group cartridge and oil observations together as they are similar products since cartridges are usually filled with oil or a distillate. Another grouping combines wax, shatter, dabs, and resin as they are products that are similar in price and have the highest potency. Similarly, hash and kief are grouped together due to similarities in products, price, and potency. Subsequent analysis uses these three larger groupings, as described in Table 2.

Table 2: Extract Market broken down by product category for June 2016 (observations with matched product names only)

	Average Price	Price Q1	Price Q3	Average Potency	Potency Q1	Potency Q3	% of Market
cartridge	$ 79.56	$65.46	$86.78	68	61	78	34%
hash/kief	$ 23.57	$16.92	$28.21	41	30	50	6%
oil	$ 43.13	$29.92	$54.25	72	68	78	5%
wax/shatter/ resin/dab	$ 30.54	$23.36	$35.77	73	70	78	55%

The largest proportion of the extract market appears to be comprised of products such as dabs, shatter, resin, and wax (55% of the categorized observations, which include roughly two-thirds of all extract observations), with the 2nd largest proportion being cartridge/oil products at approximately 39% (Table 2).

The dab/wax/shatter/resin group is also the fastest growing. In June of 2015 the average number of transactions per store per day of a wax/shatter/resin product was 5. One year later, in June 2016, that average had jumped to 17. Cartridge/oil transactions also increased, but more in keeping with the rate of growth in Washington's legal cannabis market overall (from 7 to 12 transactions per store per day).

This illustrates the potential value in partitioning extract observations. The number of extract transactions overall grew by a substantial 50% (from 15 to 30 transactions per store per day), but if someone who had particular concerns about the health consequences of dabbing only had access to that figure, they would have underestimated the 340% growth in the dab/shatter/wax/resin submarket which includes products that are particularly associated with dabbing.

Returning to the matter of trends over time, for all three of these aggregate groupings prices declined rapidly until the summer of 2015 and then begin to level out at around $30 per gram for products such as wax, shatter, and resin. Prices for cartridges/oil fell the most.

Average potency of the cartridge/oil category has been steadily increasing from 50% THC to closer to 75% THC over the last 2 years, but potency for the other two categories peaked and has since been decreasing slightly, albeit at quite different levels. Wax/shatter/resin potency levels remain around 65% THC, whereas hash/kief have significantly lower potency.

Trends over time in wholesale and retail prices for usable marijuana and extracts
Retail prices of cannabis products in Washington State have fallen significantly since legalization (Smart et al., in submission; Caulkins, 2017). This may be the result of some combination of new production and processing technologies, mastery of existing processes, economies of scale, and greater competition, including at the retail level. Policymakers pay special attention to prices because they directly impact tax revenue and indirectly affect health outcomes by influencing consumption. The elasticity of demand for marijuana is thought to be somewhere in the neighborhood of -0.5, meaning that 10% decline in retail price would increase consumption by 10% (Gallet, 2014; Pacula and Lundberg, 2014).

Arguments have been made that legalization should greatly reduce cultivation costs (Kilmer et al., 2010; Caulkins et al., 2016). Indeed, wholesale prices have been falling throughout the period of policy liberalization dating at least to the so-called “Ogden memo” (Ogden, 2009), and so well before formal state-legalization of production for a non-medical market. Before policy began to liberalize, high-quality domestically-produced cannabis sold at wholesale for around $6,000 per pound (WSIN, 2008). By November 2016, the Cannabis Benchmarks spot index – a commodities price index for the American market - had fallen below $1,400 per pound. Aphria, a Canadian producer, reports production costs of $400 - $700 per pound in US dollars at current exchange rates. Aphria now produces 60 grams per square foot in its greenhouses. If production costs fell to that typical of tomatoes grown in greenhouses – roughly $4 per square foot – that would be only $30 per pound.

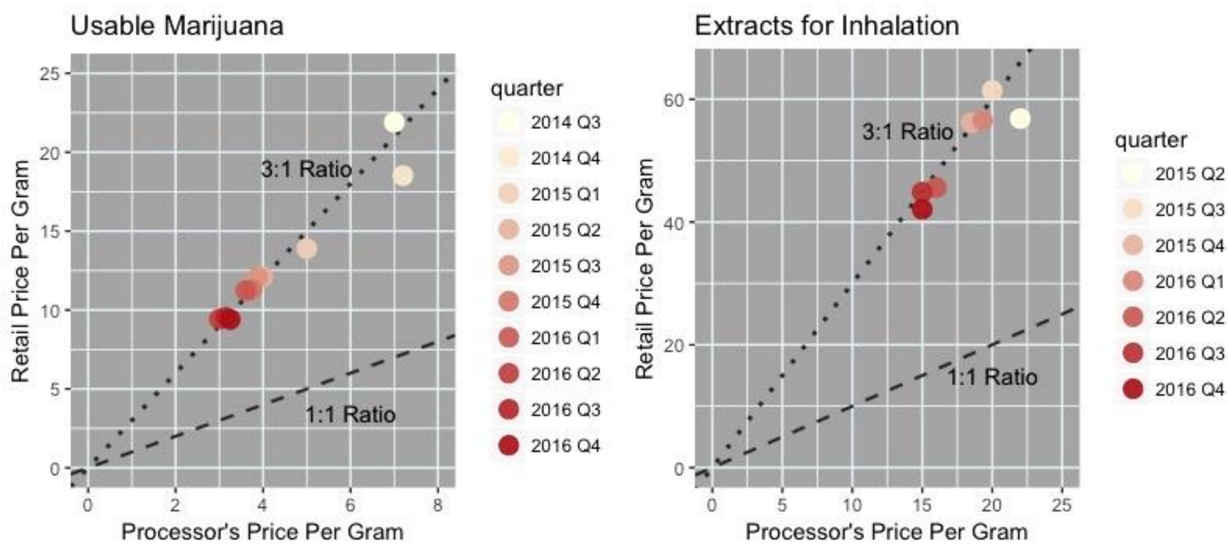
What is unclear at present is how such a radical decline in production costs, if it were to transpire, would affect (pre-tax) retail prices. There is a venerable history of asking how changes in prices further up the distribution chain may percolate down to affect retail prices because it bears directly on estimates of how effective interdiction and high-level enforcement may be at suppressing consumption (Reuter and Kleiman, 1988; Kleiman, 1992). Caulkins (1990, 2007) laid out two extreme models that may bracket the actual relationship. Under the additive model, retail price equals wholesale price plus a constant, so if wholesale prices fall by $1 per gram then so will retail prices. Under the multiplicative model, retail prices are a fixed multiple of wholesale prices, so if wholesale prices fall by 10% then so will retail prices.

Washington State’s seed-to-sale data provide a unique opportunity to examine this relationship directly because they follow each individual unit of cannabis from producer through processor down to retail sale, with prices recorded at each step along the way. By contrast, past analyses of markups were often based on law enforcement data sets that could be vulnerable to selection biases if, say, retail enforcement focused on places or types of the drug that tended to have higher prices and wholesale enforcement focused on places or types with lower prices. For example, if retail prices pertained to sinsemilla in the interior of the United States and wholesale price data came from Mexican "commercial grade" marijuana near the Southwest border that might artificially inflate measured markups.

Technical details about Washington's seed-to-sale database complicate finding the price a producer (farmer) was paid by a processor for a particular unit of cannabis, so the present analysis just compares the retail price per gram the customer paid the retail store to the (wholesale) price per gram the retailer paid the processor for both usable marijuana and extracts for inhalation.

The following graphs plot wholesale and retail average prices by quarter, with a 45-degree dashed line, a line representing a constant 3:1 ratio of retail to processor prices and a linear regression line. Figure 1 presents price data for usable marijuana, since the market was legalized. Figure 2 presents price data for extracts for inhalation starting from Q2 2015 when this category began to gain significant market share.

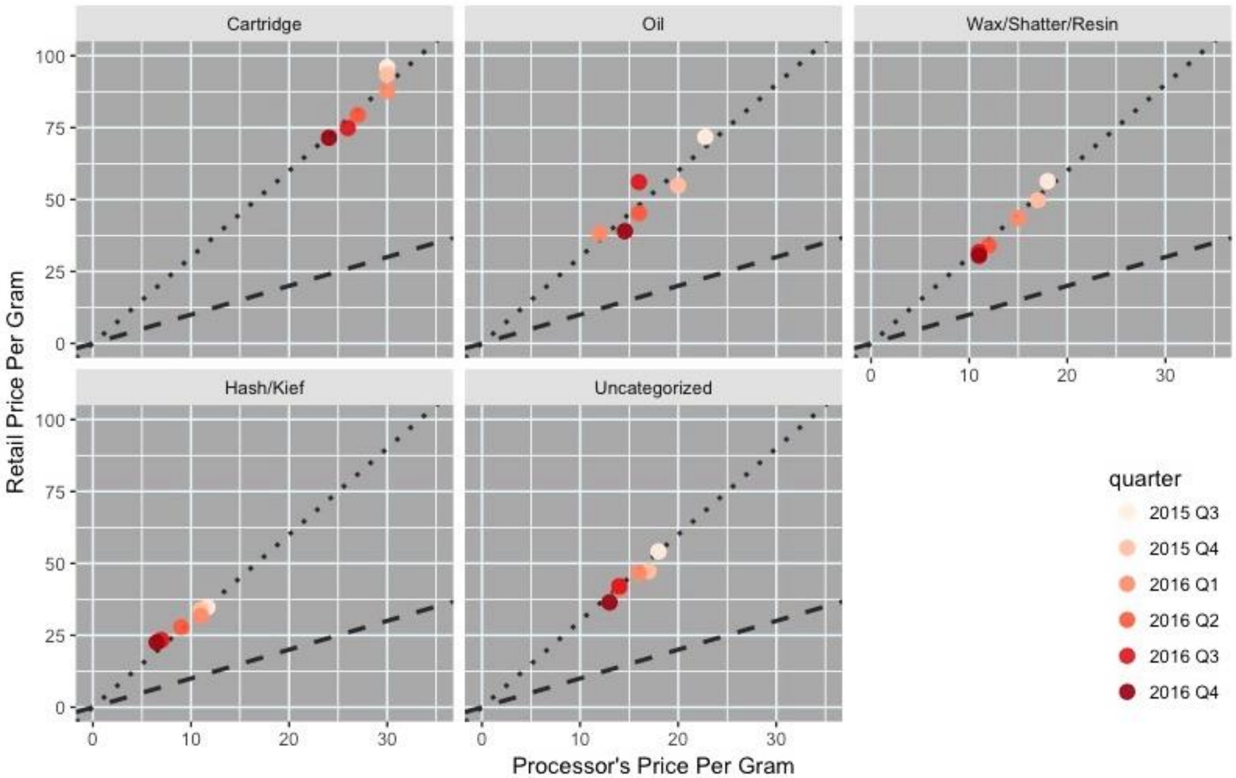
Figure 3: Relationship between retail and processor prices for usable marijuana (left panel) and extracts (right panel)



Prices tended to fluctuate initially after legalization so the lighter dot, representing Q3 2014 for usable marijuana and Q2 2015 for extracts for inhalation, is away from the trend that seems to form later. As the market matures, prices steadily decreased by more than half for usable marijuana and a third for extracts for inhalation. However, the ratio of retail price to wholesale price consistently maintained a roughly 3:1 ratio for both product types.

As noted above, extracts for inhalation is a heterogeneous category encompassing products from kief to wax and shatter, so Figure 3 replicates the plot after partitioning extracts into several more homogeneous sub-categories. Although these sub-categories vary in both price levels and the degree to which prices decreased over time, the 3:1 ratio still holds reasonably well for all of them.

Figure 4: Relationship between processor prices and retail prices for Extracts for Inhalation, per gram, by subcategory



Hence, to date, price trends in Washington State's legal cannabis market appear to have been more consistent with a multiplicative model. It is tempting to leap to a conclusion that if, as expected, continued declines in production costs further depress wholesale prices that decline might translate into proportional reductions in retail prices. However, it is important to remember that no causal inference can be drawn from the stable-to-date ratio about the causal relationship between wholesale and retail prices. Prices at the two market levels may have declined in lock step simply because parallel forces (efficiency gains, increased competition, etc.) operated independently but with equal force to date at the wholesale and retail market levels.

Analysis of seasonal variation in sales in a college town

Market patterns observed in Washington State today do not represent a post-legalization equilibrium; the market is still evolving very rapidly. College students are often seen as bellwethers of future consumption trends. For example, reports that college students' were rapidly switching from Dell to Apple computers may have triggered a sell-off of Microsoft stock (Elmer-DeWitt, 2010). So this section searches for hints about possible future consumption patterns by contrasting legal cannabis sales in so-called "college towns" with other places, when school is in session versus summer holidays when most students are away. Purchasing patterns of college students are also of direct interest because of particularly concerns about how cannabis may affect brain development in younger people (Hall, Wayne, and Lynskey, 2016).

College towns – meaning smaller, relatively isolated towns with large student populations – could present a useful natural experiment since college students might be expected to account for a sizable share of demand during the academic year and many if not most leave those towns during the summer. This allows for comparison of the market between the times when students are present and when they are absent.

This analysis focuses on Pullman, which has the greatest estimated ratio of college students to other residents. Pullman is home to Washington State University (WSU), with an enrollment of 20,043 students, while the US Census (2010) records a population in the town of Pullman of just 29,799 (*Academic Calendar,* 2016; *American Fact Finder,* 2017; *Housing and Resident Life,* 2016; *Quick Facts,* 2016;) .

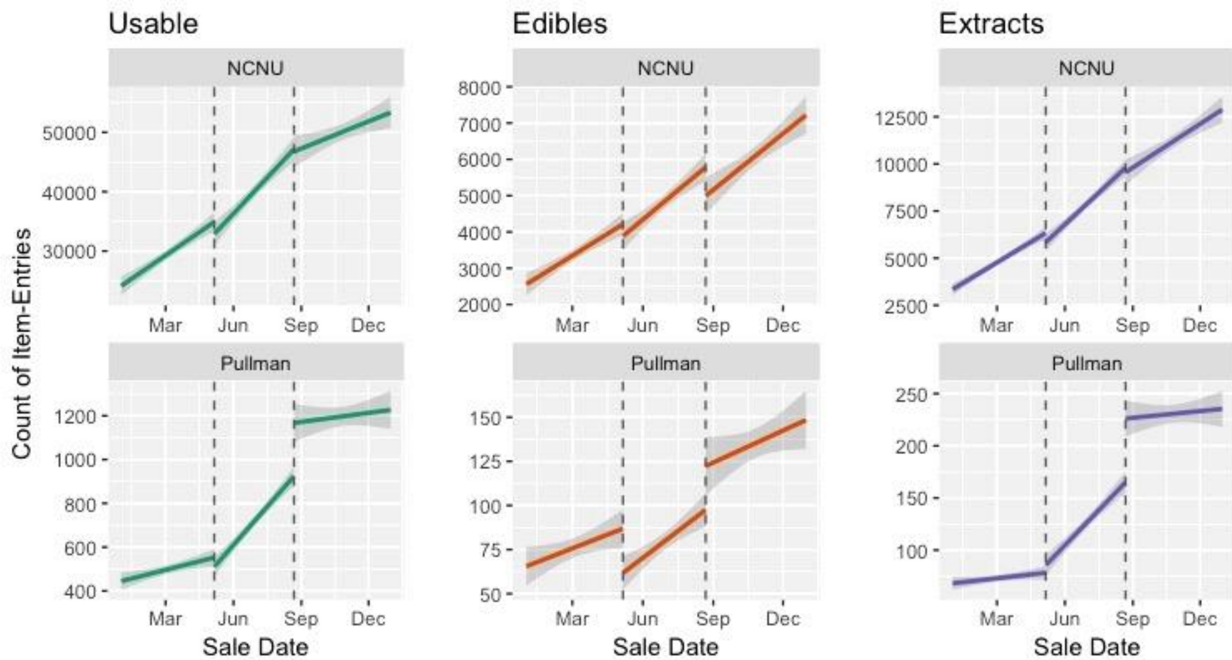
Pullman is compared to a group of non-college, non-urban towns (NCNU) towns. These towns have no colleges or have just small schools with a low student to nonstudent ratio, and have a population below 90,000 people.[1] The analysis encompasses January 2016 to December 2016, which captures a spring semester, summer vacation and fall semester, and avoids irregularities that were present at the initial opening of the new market and the change to the tax structure which took place on July 1st, 2015.

Overview of 2016 Transactions by Each Product Type

Figure 5 compares purchases in NCNU (top panel) and Pullman (bottom) for the major product categories over time. The vertical dashed lines demarcate WSU's summer break. Linear trend lines are drawn separately for the spring, summer, and fall semesters without connected spline points to allow for discontinuous jumps at semester transitions.

Figure 5: Sales trends in Pullman (bottom) and Non-College, Non-Urban (NCNU) towns (vertical dashed lines demarcate summer term, when many students are away)

[1] We considered comparing Pullman to a single nearby town with similar population that does not have a college. However, such a town was not found. Though Clarkston, which is only about thirty miles from Pullman, is similar in size and does not itself have a college, there is a college just across the border in Idaho, where cannabis sale remains illegal. Therefore stores in Clarkston may directly or indirectly be supplying those students.



Sales in NCNU towns generally increase steadily over time, as they do statewide, without any strong discontinuity in level or trend at the beginning or end of summer. The overall trend in sales is also increasing in Pullman. However, edible sales fell when students left for summer vacation while extracts rose slightly. For all three product types, there is a noticeable jump in sales when students return for fall semester. (Note the difference in y-axis scales for all NCNU collectively vs. Pullman.)

Simply finding that stores sell less when they have fewer customers would not itself be terribly interesting. Therefore, we focus further analysis on the *proportion* of products purchased that are of one type or another. Under a null hypothesis that students are similar to non-students in their purchasing patterns, the product mix would be consistent in both spring and summer.

<u>Before and after comparison in Pullman only</u>

The last day of final exams in WSU's spring 2016 semester was Friday May 6th, but simply comparing sales on Thursday the 5th and Saturday the 7th would suffer from three problems: (1) Sales vary by day of week, with Saturdays consistently being busier than Thursdays, (2) Sales on a single day in a small city are subject to some general variability, and (3) Semester breaks are not that sharp. Some students leave early, some stay later, and exams themselves may alter purchase and use of intoxicants.

Instead, we contrast two 28-day windows, with the spring window ending three weeks before the end of final exams to avoid the periodic effect around April 20th and the symmetric summer window starting three weeks later and observe some differences. For example, extracts' share of all sales rose from 9.7% in the pre-graduation window of Mar 19, 2016 to Apr 15, 2016 to 13.7% in the summer window of May 27, 2016 to Jun 23, 2016.

If no students remained in the summer and students accounted for two-thirds of cannabis purchases during the school year, that would suggest that 13.7% of cannabis purchases made by Pullman's year-round residents are extracts and the corresponding proportion for students is 7.7% since (2/3) * 7.7% + (1/3) * 13.7% = 9.7%. In other words, when students bought cannabis at a state licensed store, they would be only about 7.7% / 13.7% = 0.56 times as likely to buy an extract as are Pullman's year-round residents. (Note: That is not the same as saying 0.56 times as likely as the general population because residents of college towns ("townies") may not be representative of the general non-student population statewide.)

The actual algebra is slightly more complicated because not every student leaves Pullman for the summer. There appear to be no official data, but calls to Washington State University's Department of Student Affairs and Enrollment suggest that about r = 25% of students remain in Pullman for the summer. Any given product type's share among purchases for students (S) and non-students (A) can still be found by solving the following equations:

$$f \cdot r \cdot S + (1 - f) \cdot \mathrm{A} = (f \cdot r + 1 - f) \cdot P_{summer}$$
$$f \cdot S + (1 - f) \cdot \mathrm{A} = P_{spring}$$

where

f = Students' share of purchases during the spring,
r = Proportion of students who remain in Pullman during the summer,
P_{summer} = Observed market share of that product during the summer, and
P_{spring} = Observed market share of that product during the spring

Table 3 solves these equations assuming that students account for 55% of purchases during the school year and 25% of students remain in Pullman for the summer. The proportions for the three main product types (usable marijuana, extracts and edibles) sum to slightly less than 100% because there are some other, minor product types (e.g., "mixed marijuana"). These point estimates would suggest that edibles account for almost twice as great a share of students' consumption as they do for other residents, and extracts' corresponding share is much smaller.

Table 3: Product categories broken down by students' and other residents' in the basic scenario

	P_{spring}	P_{summer}	**S** *(student)*	**A** *(resident)*	**S/A** *(ratio of S to A)*
Usable	0.732	0.739	0.723	0.743	0.97
Edible	0.107	0.086	0.137	0.070	1.95
Extract	0.097	0.137	0.039	0.167	0.23
Other	0.065	0.038	0.102	0.019	5.36

These ratios are, however, sensitive to the somewhat arbitrary assumptions about students' share of purchases during the school year (f) and the proportion of students who remain in Pullman during the summer (r). Table 4 shows how each of these quantities vary if we consider alternative values of those two parameters within the ranges of $0.4 \leq f \leq 0.7$ and $0.2 \leq r \leq 0.25$.

Table 4: Each product category broken down by students' and other residents' in the sensitivity analysis with $0.4 \le f \le 0.7$ and $0.2 \le r \le 0.25$

	S *(student)*		**A** *(resident)*		**S/A** *(ratio of S to A)*	
	Min	Max	Min	Max	Min	Max
Usable	0.717	0.727	0.741	0.746	0.97	0.98
Edible	0.123	0.156	0.062	0.077	1.78	2.10
Extract	0.002	0.065	0.154	0.182	0.01	0.38
Other	0.085	0.126	0.009	0.027	4.09	9.48

The upshot of that sensitivity analysis is that the qualitative conclusion that students seem less likely to buy extracts and more likely to buy edibles is robust for a range of those parameters, but the extent of the differences depend considerably on those parameter values. Furthermore, seasonality could threaten this comparison. Perhaps all sorts of people, student and non-student, tend to buy more extracts in summer and more edibles in spring. We try to address that problem with a difference-in-difference estimator that compares changes from spring to summer in Pullman and other college towns with corresponding changes over the same time period in non-college towns.

Using Regression to Compare Pullman to Non-College, Non-Urban Towns

A difference-in-differences model allows for comparison of the change over time between a treatment and a control location. Here the treatment is the presence of college students. Pullman is the treated location and the control location is non-college, non-urban (NCNU towns. The outcome is the proportion of sales on a given day that are of a given product type, with three separate regressions for the three product types (usable, extracts, and edibles). A dummy variable is included to indicate whether that day is a Friday or a Saturday (days for which sales are generally higher statewide) but we also run the model for a one-week window three weeks before the end of the semester and three weeks after, so that both the treated and control periods have one observation for each type of day of the week.

The empirical model used is:

$$Y_{ct} = \alpha + \beta \text{ Pullman}_c + \gamma \text{ Spring}_t + \delta_{DD}(\text{Pullman}_c \times \text{Spring}_t) + \pi \text{ Fri\_Sat} + \varepsilon_{ct}$$

Y_{ct}: Sales proportion of a certain type of products in location c on date t;
Pullman_c: Indicator for whether the location is Pullman;
Spring_t : Indicator for the Spring Semester;
Fri_Sat : Indicator for the day being a Friday or Saturday;
δ_{DD} : Coefficient for the difference in differences;

Under the null hypothesis that students have similar purchasing habits as other residents, the proportion of purchases that are of a particular type should not change whether students are present or absent. (The same would not be true for the absolute numbers of sales, which is why the outcome is the proportion of sales by product type, not the sales levels.)

For the assumptions underpinning the difference-in-differences model to hold, the trends in the proportion of products must be similar between Pullman and NCNU towns during the summer (when there is no treatment) and different in the spring (when the treatment occurs). In Figure 5 above it appears there are indeed similar sales trends in the summer in both groups. However, that parallel does break down if higher order polynomials, not just straight lines, are used to depict those trends. (See Figure 6 below for a quadratic fit.)

The results in Table 5 suggest that stores in Pullman tend to sell a greater proportion of edibles and fewer extracts in spring, when students are present, than in summer, relative to the pattern seen in other locations.

Table 5: Impact of Students' Presence on the Proportion of the Products Sold

	Y1 (% Usable)		**Y2** (% Extract)		**Y3** (% Edible)	
(Intercept)	0.721		0.14		0.102	
	(0.0084)		(0.0047)		(0.0058)	
Pullman	-0.0188		-0.006		-0.024	**
	0.0113)		(0.0063)		(0.0078)	
Spring	0.017		-0.001		-0.009	
	0.0113)		(0.0063)		(0.0078)	
Fri_Sat	-0.021	*	0.004		0.018	**
	(0.0088)		(0.0050)		(0.0061)	
Pullman_Spring	-0.025		-0.038	***	0.037	**
	(0.016)		(0.009)		(0.012)	

<u>Replication with Other College Towns</u>

To assess whether effects observed in Pullman are generalizable, parallel analysis could be performed for the towns of Ellensburg (population 18,000), which is home to Central Washington University's 11,000 students, and Airway Heights (population 6,000) which is 13 miles from Eastern Washington University's 13,000 students. (The store in Airway Heights is the closest one to Eastern Washington University.)[2]

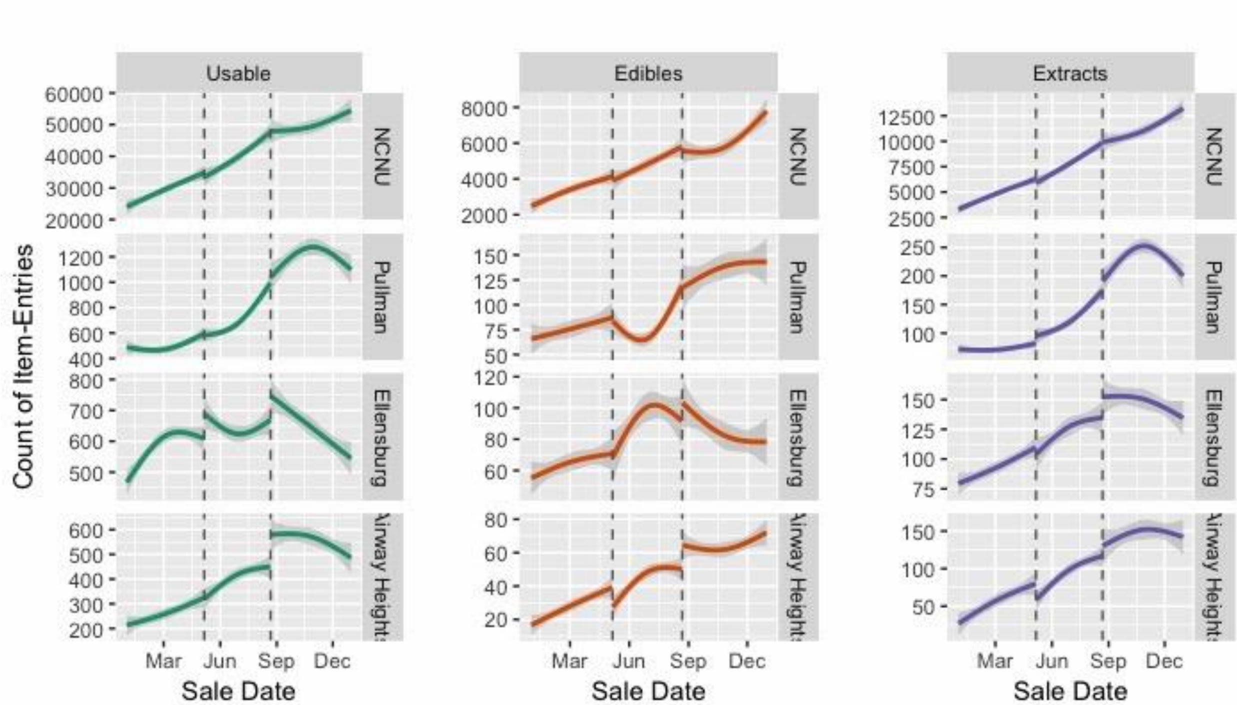
However, Figure 6 shows that the patterns in these other towns do not appear consistent with those seen in Pullman. For example, Figure 6 illustrates that while sales of edibles decrease in Pullman during the summer, they increase in Ellensburg. There also does not appear to be much of a difference in summer between Airway Heights and NCNU.

Furthermore, the trends within the untreated period (summer) in the treated locations (Ellensburg and Airways Heights) do not closely parallel the summer trends in the control locations (NCNU towns). That violates the assumptions that must hold for the difference-in-differences regression analysis to be valid.

There are reasons why the results for Pullman may be valid and generalizable to college students, but nonetheless not be replicable for these two other college towns. For instance, both Central Washington University and Eastern Washington University follow a quarter not a semester system. However, that some initial attempts to shift time windows and otherwise address those differences did not permit a successful replication gives us pause. The effect observed in Pullman may be a Pullman effect, not a college town effect, and hence may arise from some idiosyncratic factors, such as decisions by stores in Pullman as to what types of products to stock, and not be successfully capturing differences in student vs. non-student purchase patterns.

Figure 6: Quadratic fits to sales trends in Non-College, Non-Urban (NCNU) towns and three college towns (The graph omits sales May 6, 2016 to June 9, 2016 and also August 22, 2016 to September 21, 2016 because the three schools have different academic schedules, and these buffers allow for comparison of time periods when classes are either in session or not in session at all three universities)

[2] Other candidates for parallel analysis seem less promising either because of their proximity to larger metropolitan areas and/or because they are Christian colleges whose students may have different patterns of consumption of intoxicants.



Variation in the Particular Types of Extracts Used in College Towns

Since the assumptions underpinning the regression model are violated for the other college towns, we close by stepping back to basic descriptive statistics. As noted above, extracts are a heterogeneous category, and trends that hold for the category broadly may not hold for all subcategories. Thus, it is of interest to ask whether the "Pullman effect" of students' perhaps shunning extracts holds for all types of extracts.

Table 6 shows the proportion of total extract sales by type of extract in spring vs. summer of 2016 for each of the three college towns, for the non-college non-urban group (NCNU), and also for urban cities[3]. It appears that consumers in Pullman purchase more wax, shatter, and resin products and fewer cartridge and oil products than do consumers in both large cities and small towns. Figure 7 plots these proportions day-by-day with a second order polynomial fit. It shows a noticeable drop in the purchases of wax, shatter, and resin in Pullman in the summer that is not seen in urban areas or small towns. Additionally, this figure indicates lower use of those products in urban areas than in the rest of the state.

This detailed breakdown suggests that even if the "Pullman effect" discussed above does mean that college students are less likely to use extracts overall, they may not necessarily be less likely to use wax/shatter/resin and, hence, be less likely to dab.

[3] This analysis considers "big cities" to be those with a population over 90,000. They are Seattle, Spokane, Tacoma, Vancouver, Bellevue, Everett, Yakima, and Renton.

Table 6: Proportion of Extract Market by Location and Product Subcategory

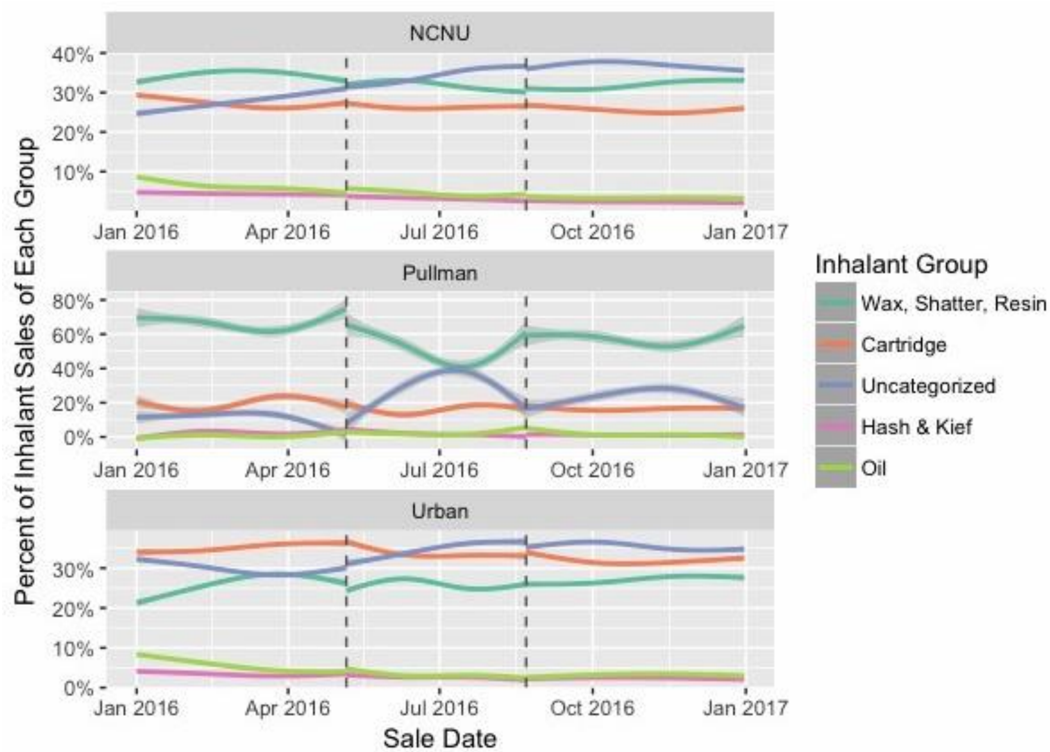
Average from Jan 1 2016 to Apr 30, 2016

	Urban	NCNU	Pullman	Ellensburg	Airway Heights
Wax, Shatter, Resin	26.2	34.6	66.7	68.2	35.4
Cartridge	35.2	27.2	19.4	16.6	18.0
Oil	5.5	6.2	0.5	1.3	18.0
Hash & Kief	3.4	4.4	2.1	2.7	3.6
Uncategorized	29.7	27.6	11.3	11.1	25.0

Average from Jun 1, 2016 to Jul 30, 2016

	Urban	NCNU	Pullman	Ellensburg	Airway Heights
Wax, Shatter, Resin	26.0	32.2	46.9	31.7	73.0
Cartridge	33.2	26.2	15.8	21.3	16.2
Oil	3.0	4.2	1.6	6.4	1.2
Hash & Kief	2.6	3.1	1.6	2.9	2.9
Uncategorized	35.2	34.3	34.0	37.7	6.7

Figure 7: Proportion of Extract Usage by Location



Discussion

It has been common in the past to differentiate cannabis flower from more concentrated products produced by traditional "mechanical" extraction methods (termed herbal cannabis vs. resin in much of the world and marijuana vs. hashish in North America). Sometimes there were also special names for higher potency flower products (sinsemilla vs. commercial grade in North America and skunk or hydroponic vs. general herbal in Europe).

However, that typology fails badly in Washington's legal market. Almost all of the herbal product is high-potency (averaging over 20% THC), and hash and kief account for a quite small share of the rapidly growing extract market segment, which is dominated by solvent-based extraction (particularly if one views supercritical C02 as a solvent). Furthermore, edibles constitute a non-trivial share of sales, and there are other categories (e.g., marijuana "mixes").

It may be that post-legalization, researchers and policy makers will need to think in terms of a family of cannabis products, akin to how we think of opioids, new psychoactive substances (NPS), and amphetamine-type stimulants (ATS), not a single drug "marijuana" the way it is possible to think of cocaine and cocaine markets.

This paper took a step toward creating a typology for one important part of that family, namely extraction products. There are at least two broad reasons why the legal industry is embracing extracts. First, just as prescription pills can bring opioids to the masses who shun injecting drugs, extract-based products can reach customer segments that dislike smoking, perhaps because of the negative connotations of tobacco smoking. Second, now that the exaction machinery does not expose owners to the risk of arrest and seizure, there is no reason to discard all of the THC contained in leaves and other parts of the plant where the THC exists in less concentrated form. Since most of the plant's weight is in leaves, not flowers, a considerable share of the intoxicant appears in parts of the plant that could not easily be brought to market before legalization.

Our guess is that Washington State today manifests only the beginning not the culmination of the changes in product form and marketing that legalization will bring. It is not hard to imagine additional product forms emerging that might shake up any current typology, e.g., "bundled" products that combine cannabis with alcohol, tobacco, or other intoxicants. So we likewise view the present effort as just the beginning, not the culmination, of attempts to partition the rapidly evolving cannabis product space.

The curiously stable 3:1 ratio of retail to wholesale prices merits further analysis and replication. It appears to hold across a range of products, but obviously all of our data come from just one jurisdiction and time period. Perhaps the stability of that ratio reflects not something intrinsic to the products or their production technologies, but rather something unique to Washington State's regulatory practices, or even the oddities of tax law (Section 280(e)) during this strange time when the cannabis business is state-legal but still fully prohibited under federal law.

The topic, though, is of clear importance. If the multiplicative relationship between wholesale and retail prices were sustained going forward and, as is expected, production costs continue to fall, that might portend quite low pre-tax retail prices in the future. However, we stress that the analysis above is purely descriptive and does not in any way demonstrate a causal relationship between declines in wholesale prices and corresponding declines in retail prices.

Put simply, the results of the college town analysis are a bit of a puzzle. We have the paradox that students leaving Pullman for the summer appear to have a not insubstantial effect on the mix of products sold, but not an overwhelming impact on the volume of products sold. One can spin stories for why that might occur. For example, the undergraduates, who are mostly under the age of 21 and so cannot purchase directly from the state-licensed stores, might still purchase most of their cannabis through the black market. (Washington State is thought to have large-scale purely black market production for export to other states that have not yet legalized.) But smoking cannabis in a dormitory might be hard to conceal from residential hall monitors, so they might obtain (via friends who are over 21) edibles from the state-licensed stores for use in dorms. And if the students who remain over the summer live in off-campus housing because dormitories close over the summer, that might account for a decline in demand for edibles during summer months.

Another story is that Washington State University students who leave Pullman for the summer might continue to create demand for sales from stores in Pullman if their friends who stay behind for the summer are buying in Pullman and sending the product back to the students in their home towns or wherever they spend the summer. Indeed, if Washington State University (the school in Pullman) has a greater proportion of out of state students than do either Central or Eastern Washington University, and students going elsewhere within the state buy locally over the summer, that purchase-by-proxy effect might help explain why seasonal patterns appear to differ in Pullman vs. Ellensburg and Airway Heights.

In sum, the natural experiment created by students leaving a college town for the summer may not be quite as "clean" an exogenous shock as one might have hoped at first. There may be other natural experiments—such as the effect of legal changes in neighboring states on sales in border towns—that might be worth exploring.

Conclusions and Further Work

At the highest level, the conclusions of this analysis are clear. Legalization induces dramatic changes in cannabis markets, and seed-to-sale monitoring systems offer a valuable window into those important changes. That suggests that states might want to design these systems in ways that facilitate not only routine administrative functions that apply to individual licensees (collecting taxes, ensuring compliance with testing requirements, detecting diversion, etc.), but also the monitoring of market-level aggregates that are of interest to public health.

Although we did not delve into particular technical aspects of the data system, in many ways Washington State's present system is not well-designed to serve those functions. For example, the system collects information from many activities throughout the market, and the database does not make it clear how the data stored in one table connect to data another table. This vastly complicates the ability to actual track product from "seed-to-sale" or back.

The other obvious and significant limitation of these data is the absence of individually-identifying information on the customer. If customers were known, then changes in consumption patterns between college age purchasers and older people would be immediately obvious. One could also look at trajectories of purchasing patterns over time for individuals, and could investigate the extent to which a smaller number of frequent purchasers account for a disproportionate share of sales.

Indeed, it may be that the next generation of market analysis after exploiting seed-to-sale regulatory monitoring systems will be studying data from private companies' "frequent buyer" programs since they, unlike the state, will want to connect different transactions associated with the same repeat customer.

Nevertheless there are a variety of additional analyses that could profitably be pursued with this or parallel seed-to-sale data systems. For one, the partitioning analysis could be replicated with edibles, a category that is in some respects as heterogeneous as extracts, including various solid products (cookies, caramels, granola, candy cubes, peanut butter cups, etc.) and liquids (lemonade, punch, droppers, and sprays). Analysis of weight, potency, and hence potency-adjusted price per is more complicated, but might inform later federal regulatory decisions of concerning psychoactive food additives.

Testing practices are another topic of interest. In Washington State, producers get to pick the lab that will test their products, and there is no auditing of reported test results by any independent third party. This creates incentives for labs to produce results that are desirable from the producers' perspective, such as inflating measured THC or repeatedly sampling until a batch passes a quality control test (e.g., concerning mold or pesticide residue). Malfeasance cannot be seen directly in these data, of course, but it is clear that some producers shop around, trying out various testing labs, and it would be interesting to see if switches from one lab to another seem more associated with tangibles like price and location or with test results.

More generally, aspects of the supply chain above retail can be investigated in ways that have not previously been possible, such as yields per unit area and price markups between growers and processors, not just between processors and retailers. The structure of the industry and the database complicate those analyses, but their novelty might warrant that additional effort.

Correlating these fine-grained sales data with public health outcome data on traffic crashes, emergency room visits, and the like is of considerable interest. These data record down to the minute when a sale was recorded so they may expose correlations that are hidden in data reported only monthly or quarterly. That detail might be particularly valuable for trying to tease out substitution or complementarity with alcohol, e.g., if crimes commonly associated with alcohol go up or down on immediately after a store changes its menu of cannabis prices.

References

Ben Lakhdar C., G. Vaillant N., Wolff F.-C. Price elasticity of demand for cannabis: does potency matter? Addict Res Theory 2016; 24: 300-312.

Caulkins, J.P. (1990). The Distribution and Consumption of Illicit Drugs: Some Mathematical Models and Their Policy Implications. Doctoral dissertation, MIT, Cambridge, MA.

Caulkins, Jonathan P. 2007. Price and Purity Analysis for Illicit Drugs: Data and Conceptual Issues. Drug and Alcohol Dependence. 90S, S61-S68. DOI 10.1016/j.drugalcdep.2006.08.014.

Caulkins, Jonathan P., Beau Kilmer and Mark A.R. Kleiman (2016). Marijuana Legalization: What Everyone Needs to Know, 2nd Edition. Oxford University Press.

Caulkins, Jonathan P. (2017). Recognizing and Regulating Cannabis as a Temptation Good. *International Journal of Drug Policy*, 42:50-56.

Daniulaityte R., Nahhas R. W., Wijeratne S., Carlson R. G., Lamy F. R., Martins S. S. et al. "Time for dabs": Analyzing Twitter data on marijuana concentrates across the US. Drug Alcohol Depend 2015; 155: 307-311.

Elmer-DeWitt, Philip (2010). Big Macs on campus. *Fortune*. Downloaded on April 21st, 2017 from http://fortune.com/2010/08/07/big-macs-on-campus/.

Gallet, Craig A. (2014). Can price get the monkey off our back? A meta-analysis of illicit drug demand. Health Economics, 23:55-68, published online in 2013 at DOI: 10.1002/hec.2902.

Hall, Wayne, and Michael Lynskey. Evaluating the public health impacts of legalizing recreational cannabis use in the United States. Addiction (2016).

Kilmer, Beau, Jonathan P. Caulkins, Gregory Midgette, Linden Dahlkemper, Robert J. MacCoun, and Rosalie Liccardo Pacula. *Before the grand opening: measuring Washington State's marijuana market in the last year before legalized commercial sales*. Rand Corporation, 2013.

Kilmer, Beau, Jonathan P. Caulkins, Rosalie Liccardo Pacula, Robert MacCoun, Peter Reuter (2010). Altered State? Assessing how marijuana legalization in California could influence marijuana consumption and public budgets. RAND OP-315-RC, Santa Monica, CA

Kleiman, Mark AR. (1992) Against Excess: Drug Policy for Results. Basic Books, New York.

Krauss M. J., Sowles S. J., Mylvaganam S., Zewdie K., Bierut L. J., Cavazos-Rehg P. A. Displays of dabbing marijuana extracts on YouTube. Drug Alcohol Depend 2015; 155: 45-51.

Loflin M., Earleywine M. A new method of cannabis ingestion: the dangers of dabs? Addict Behav 2014; 39: 1430-1433.

Mehmedic Z., Chandra S., Slade D., Denham H., Foster S., Patel A. S. et al. Potency trends of en ‐ THC and other cannabinoids in confiscated cannabis preparations from 1993 to 2008. J Forensic Sci 2010; 55: 1209-1217.

Morean M. E., Kong G., Camenga D. R., Cavallo D. A., Krishnan-Sarin S. High school students' use of electronic cigarettes to vaporize cannabis. *Pediatrics* 2015; **136**: 611-616.

Ogden, Davie W. 2009. Memorandum for selected United State Attorneys on investigations and prosecutions in states authorizing the medical use of marijuana. Available from the U.S. Department of Justice Archives at https://www.justice.gov/archives/opa/blog/memorandum-selected-united-state-attorneys-investigations-and-prosecutions-states.

Pacula, Rosalie Liccardo, and Russell Lundberg. "Why changes in price matter when thinking about marijuana policy: A review of the literature on the elasticity of demand." Public health reviews 35, no. 2 (2014): 1.

Reuter, Peter and Mark Kleiman Risks and Prices: An Economic Analysis of Drug Enforcement, in Morris and Tonry (eds.) Crime and Justice: An Annual Review of Research Volume 7, Chicago: University of Chicago Press. 1986, pp.289-340.

Smart, Rosanna, Jonathan P. Caulkins, Beau Kilmer, Steve Davenport, & Greg Midgette (in submission). Variation in cannabis potency and prices in a newly-legal market: Evidence from 30 million cannabis sales in Washington State.

U.S. Census Bureau. *American Fact Finder;* 2017.

Western States Information Network (WSIN). 2008. Illegal Drug Price & Purity Guide.

Washington State Liquor and Cannabis Board. *Medical Marijuana FAQ's*; 2016.

Washington State Department of Revenue. *Marijuana retailers - medical endorsement*; 2016.

Washington State Department of Revenue. *Taxes Due on Marijuana*; 2016.

Washington State Liquor and Cannabis Board. *Weekly Marijuana Report, December 12, 2016*; 2016.

Washington State University. *Housing and Resident Life;* 2016

Washington State University. *Quick Facts;* 2016

Washington State University. *Academic Calendar;* 2016