The value of cannabis characteristics: A hedonic price approach

Robin S. Goldstein[1], Ian Xu[1], Jarrett Hart[1], and Daniel A. Sumner[1]

Abstract

The U.S. legal cannabis market is expanding rapidly, but to date there is no research that examines the individual product characteristics that contribute to product prices. This is in large part due to the unavailability of data. We assemble a large data set of millions of cannabis prices and product characteristics in 19 U.S. states. We use hedonic analysis to estimate implicit prices of characteristics such as potency, product type, and plant species and strain.

[1] Department of Agricultural and Resource Economics, University of California, Davis

Introduction

Cannabis has been widely consumed for centuries and spent almost a century of being mostly illegal in the United States and many other countries. Now, despite remaining a schedule 1 narcotic under federal law and regulation, cannabis possession and sale is again legal under state law in most of the United States (by population). A abundance of economic issues surround legal cannabis markets. This paper investigates the central economic relationships that combines concerns of demand, supply and regulation.

Methodology and Data

The hedonic price method, pioneered by Rosen (1974), aims to estimate implicit prices for the individual "utility-bearing attributes or characteristics" of products. Typically, implicit prices are estimated from a regression of price on a vector of objective characteristics, using data that include a variety of different combinations of characteristics that correspond to different prices.

The hedonic method has been used to estimate the value of wine characteristics (e.g. Gustafson et al., 2016; Nerlove et al., 1995), but has never been applied to cannabis data. We use the hedonic price method by regressing price on cannabis product characteristics to estimate implicit price values of each characteristic in a variety of retail products and package types.

We construct separate hedonic price models for flower, concentrates, and several sub-categories of concentrates. Prices and log prices are the dependent variables, and a vector of product characteristics are used as the right-hand-side variables. We test several model specifications, including store- or state-level fixed effects to control for spatial variation, and weekly fixed effects or time trends of varying degrees to control for temporal variation. One of our primary variables of interest is THC content, akin to alcohol content in beer, wine, and spirits. We include versions with THC and product characteristic interaction variables to examine heterogeneity in the THC-price relationship across distinct product segments.

We look at the relationship between product characteristics and prices for two types of cannabis products separately: one-eighth ounce packages of flower, referred to as eighths and used for

smoking, and one-gram packages of concentrate used for vaping. For cannabis flower, price is measured in dollars per eighth-ounce. For cannabis concentrate, price is defined as dollars per gram.

The functional form of the hedonic price model determines the relationship of the marginal prices and product characteristics. We take several steps to identify the appropriate model specification. First, using an untransformed dependent price variable, we use AIC and BIC to test fit of the model when including various fixed effects, product characteristics, or linear, quadratic, and cubic potency covariates. Then using the preferred specification, we test multiple transformations of the dependent variable. This is a common practice in hedonic price modelling—Landon and Smith (1997) and Costanigro et al. (2007) find that the inverse square root transformation of price performs best when examining hedonic price attributes of wine. We test quadratic, linear, log-linear, reciprocal, reciprocal square root, and reciprocal quadratic Box-Cox transformations to determine the appropriate functional forms of the cannabis hedonic models. For all model specifications, we use cluster-robust standard errors.

After determining the appropriate model specification, we compare estimates using the full sample to subgroups of each product category. We test specifications in which cannabis flower is partitioned by price. We examine price distributions to identify multimodality in prices for obvious price segments, and rely on Weedmaps' online shopping portal to identify price segments as they are presented to consumers. Cannabis concentrate is segmented by qualitative characteristics: whole plant and distillate. Whole plant refers to products in which all cannabinoids (such as CBD and terpenes) from the plant are retained in the concentrate. Distillate refers to products in which only the primary cannabinoid, THC, is retained in the concentrate. In general, distillates have higher THC potency; there are also several qualitative characteristics that exclusively apply to whole plant or distillate concentrates, which we control for in our hedonic estimations.

Hedonic estimation

We test numerous specifications for cannabis flower and cannabis concentrate, estimating additional sets of regressions for California alone and all states with Weedmaps prices; regressions in which prices either include or exclude taxes and fees; regressions with various

transformations of price or THC; regressions with different time trends and fixed effects; and regressions on distinct product segments. Here we define the linear version of the hedonic price models, and we describe the variables in detail.

The hedonic price model for eighth-ounce packages of cannabis flower is defined in equation (1), with price serving as the dependent variable.

$$
\begin{aligned}
EighthPrice_{i,j,k,t} \\
= \beta_1 THC_{i,j,k,t} + \beta_2 Hybrid_{i,j,k,t} + \beta_3 Indica_{i,j,k,t} + \beta_4 Sativa_{i,j,k,t} \\
+ \beta_5 Recreational_{i,j,k,t} + \eta_i + \mu_j + \delta_k + \tau_t + \epsilon_{i,j,k,t}.
\end{aligned}
\tag{1}
$$

Price per eight-ounce of flower, $EighthPrice_{i,j,k,t}$, of product i at store j in state k during week t is regressed on a vector of product attributes. A summary of the variables used in various model specifications can be found in Table 1. Product characteristics include THC content, $THC_{i,j,k,t}$; plant type indicator variables $Hybrid_{i,j,k,t}$, $Sativa_{i,j,k,t}$, and $Indica_{i,j,k,t}$; recreational product type indicator variable, $Recreational_{i,j,k,t}$.

Brand-specific fixed effects, η_i, are included in some specifications to compare the THC-price relationship within-brands versus across brands. Store-specific fixed effects, μ_j, are used to control for price differences across markets that may arise from factors such as local regulations, per capita income, or unobservable store characteristics. Retailer state fixed effects, δ_k, are included to control for spatial variation in prices. Time-specific fixed effects, τ_t, correspond to the date of the price listing, and are included to control for seasonality and inflation. The error term is denoted by $\epsilon_{i,j,k,t}$, and is clustered at the store-level.

The hedonic price model for one-gram packages of concentrate is defined in equation (2).

$$
\begin{aligned}
ConcentratePrice_{i,j,k,t} \\
= \beta_1 THC_{i,j,k,t} + \beta_2 Hybrid_{i,j,k,t} + \beta_3 Indica_{i,j,k,t} + \beta_4 Sativa_{i,j,k,t} \\
+ \beta_5 Recreational_{i,j,k,t} + \beta_6 Disposable_{i,j,k,t} + \beta_7 Pod_{i,j,k,t} \\
+ \beta_8 Cartridge_{i,j,k,t} + \beta_9 Cured_{i,j,k,t} + \beta_{10} Resin_{i,j,k,t} + \beta_{11} Rosin_{i,j,k,t} \\
+ \beta_{12} CO2_{i,j,k,t} + \eta_i + \mu_j + \delta_k + \tau_t + \epsilon_{i,j,k,t}.
\end{aligned}
\tag{2}
$$

As in equation (1), price per gram of concentrate, $OilPrice_{i,j,k,t}$, of product i at store j in state k during week t is regressed on a vector of product attributes. Summaries of the product characteristics for concentrate can be found along with the flower characteristic summaries in Table 1. Product characteristics include THC content, $THC_{i,j,k,t}$; plant type indicator variables $Hybrid_{i,j,k,t}$, $Sativa_{i,j,k,t}$, and $Indica_{i,j,k,t}$; recreational product type indicator variable, $Recreational_{i,j,k,t}$; indicator variables for pod or cartridge packaging types, $Pod_{i,j,k,t}$ and $Cartridge_{i,j,k,t}$; and indicator variables for cured, resin, rosin, or CO2 advertised in product the product description, $Cured_{i,j,k,t}$, $Resin_{i,j,k,t}$, $Rosin_{i,j,k,t}$, and $CO2_{i,j,k,t}$.

Once again, the following fixed effects are used in various specifications: brand-specific fixed effects, η_i; store-specific fixed effects, μ_j; retailer state fixed effects, δ_k; and time-specific fixed effects, τ_t. The error term is denoted by $\epsilon_{i,j,k,t}$, and is clustered at the store-level.

In some estimations we omit store fixed effects in favor of state fixed effects and include an indicator variable for physical storefronts. We do so to estimate the difference in prices for eighths or concentrate from physical storefronts versus delivery only stores. In future analysis we will use state fixed effects in conjunction with zip-code level demographic characteristics.

Variable descriptions

In this section we describe the variables used in our hedonic estimations. Variable means and definitions are in Table 1.

The primary variable of interest is THC, which denotes the potency of cannabis. THC quantity is in percentage terms, with a mean of 26.0% and standard deviation of 5.34% for eighths, and a mean of 84.6% and standard deviation of 5.34% for concentrate.

Plant type indicator variables indicate if the product is Indica, Sativa, or a Hybrid. Hybrid plants are a mix of Indica and Sativa, but could predominately be either type. The counterfactual plant type is having no labeled classification. Indica and Sativa are generally associated with different effects of consuming cannabis, but the terms also refer to plant types with distinct physical characteristics. Weedmaps (2023) describes the difference in effects as, "indica plants are often characterized as having sedative effects, which typically lead to an intense body high, while

sativa strains are thought to be uplifting and produce more of a head high." As for the physical characteristics, Indica plants are shorter, have quicker growing cycles, and are well-suited for cool climates; Sativa plants are taller, have longer growing cycles, and are well-suited for warmer climates (Weedmaps, 2023). Because of these differences, Sativa strains are generally more suitable for outdoor growing whereas Indica strains for indoor or mixed-light growing. However, environmental factors, costs of production, and strain-specific considerations also affect producers growing decisions.

The recreational variable indicates if the product is listed for sale as a recreational product as opposed to being strictly available for purchase by medical patients. Although recreational and medical products are in fact identical, their prices can vary substantially (some of this price variation can be attributed to lower taxes for medical products).

Package types for concentrate products include 'pod,' 'disposable,' and 'cartridge.' If any of these terms are in the product title or description, the corresponding indicator variable takes a value of 1. The counterfactual is that none of these terms are included in the product title or description. Pods and cartridges are attachable concentrate vessels that typically require separate devices to vaporize the concentrate. The pod system of vaporizers are considered a high-end cartridge system, characterized as "usually compact, slick, and flat like a USB stick," whereas cartridges are typically, "cylinder designs similar to a cigarette..." (Cilicon, 2021). Pods are generally geared towards frequent users of vaporizers. "The design of pens is geared more towards convenience than longevity. Pods use larger cartridges that are easy to replace so you can keep vaping with the same device. As well, pod vaporizer systems use larger, more powerful batteries which are longer lasting." (Champoux, 2018).

Disposable products require no separate device, they include both a vape cartridge and vape battery—these are generally more expensive than cartridges. Disposable pens are ideal for infrequent concentrate users, as it is cheaper than purchasing a non-disposable pen.

Concentrates are categorized as whole plants or distillates. Whole plant concentrates contain all additional cannabinoids present in the plant such as CBD and terpenes; distillates may also contain these cannabinoids, but if so, they are infused after the distillation process. Terpenes are said to provide aromatic characteristics, and their interaction with THC and other cannabinoids

supposedly alters or enhances the effects of cannabis. In general, products may be high in both THC and terpenes, but terpene quantities are not listed in product descriptions. However, terpenes are a separate compound from the THC and CBD cannabinoids. Thus, if a concentrate cartridge is advertised as having 100% TH, it cannot contain any CBD or terpenes (unless the THC content is inaccurate, which is commonly the case). For this reason, we do not necessarily expect the price of concentrate to be strictly increasing with THC content—higher THC could be indicative of lower levels of other cannabinoids.

Results

Flower

Results for the hedonic price attributes of eights are provided in Table 2. The dependent variable is the total price per eighth, including taxes and delivery fees. In (1), variables are not transformed; therefore, coefficients can be interpreted as dollar amounts. In (2), price and THC are log-transformed. Product characteristic variables and weekly and store fixed effects are included in both specifications.

The coefficient for THC is consistently positive and statistically significant, 0.881 in (1) and 0.478 in (2). That is, increasing THC content by 1% increases price on by $0.881 or 0.478% on average.

Hybrid, Indica, and Sativa coefficients are negative and statistically significant across specifications, suggesting these products are sold at a discount relative to products without a plant type label. On average, Sativa products are discounted the most, followed by Indica, then Hybrid (coefficients for Hybrid and Indica are equivalent in (2)). This effect may stem from supply-side conditions—Sativa plants are more commonly grown outdoors (lower production costs), whereas Indica plants are more commonly grown indoors (higher production costs).

Thee coefficient for Recreational is positive and statistically significant at 3.147 and 0.106 in (1) and (2), respectively. This result suggests that the price difference between recreational and

medical products is \$3.147, or 11.2%.[1] Much of this difference is caused by taxes on recreational cannabis. However, there is still a premium for recreational products, which could be in part due to factors such as cultivation taxes or licensing fees (we find a small positive effect in our estimates excluding taxes in the appendix).

Concentrate

Results for the hedonic price attributes of 1-gram concentrate products are provided in Table 3. The dependent variable is the total price per gram, including taxes and delivery fees. We use a linear specification in (3) and (5), and log-transformation of THC and price in (4) and (6). Concentrates are stratified by product segment: whole plant in (3) and (4) and distillates in (5) and (6). Product characteristic variables and weekly and store fixed effects are included in each specification. In addition to the variables included in the eighths specifications, product characteristic variables for Disposable, Pod, Cartridge, Cured, Resin, Rosin, and CO2 are included. In terms of R-squared, the linear models fit better than the log-transformed models.

The coefficient for THC is statistically insignificant for whole plants (3) and (4), and negative and statistically significant for distillates, -0.042 in (5) and -0.050 in (6). In contrast to eighths, we do not find that price has either no significant relationship, or a small, negative relationship to THC content. One plausible explanation is that as THC increases, there may be a lower quantity of other cannabinoids such as CBD or terpenes. There may also be a nonlinear relationship between price and THC (which we examine in the appendix).

For both whole plants and distillates, Hybrid, Indica, and Sativa either have statistically significant premium or not significantly different than the baseline of no strain label (contrary to the results for eighths). Across all four specifications, the implicit price of Indica is the highest among strains at \$0.676 for whole plant and \$0.837 for distillates. Again, this could be partially driven by supply-side characteristics—Indica plants are better suited for indoor growing, which has a higher cost of production.

---

[1] The following transformation is used to interpret a coefficient of an indicator variable in percentage terms in a log model: $\exp(\beta) - 1$.

The coefficient for Recreational is positive and statistically significant, 1.952 in (3) and 1.615 in (5). As with flower, further testing suggests that the majority of the price difference between recreational and medical products is reflected in taxes.

Concentrate products that are sold in Pods are sold at the largest premium, ranging from $7.279 for whole plant to $8.558 for distillates relative to no package type indicated. Disposable products also attract a statistically significant premium, ranging from $1.081 for whole plant to $2.593 for distillates. Cartridges have statistically significant negative implicit price of $0.852 for whole plant and $2.938 for distillates. For these two product segments, the implicit prices associated with packaging are substantially larger for distillates than whole plant concentrates.

Rosin is a quality attribute typically associated with higher-end products. The coefficient for Resin is positive and statistically significant, at $12.24 for whole plants. Interestingly, there are large negative implicit prices for Cured and Resin whole plant concentrates. There is a positive implicit price for CO_2 distillates, which coincides with expectations of a premium for products considered to be safer to consumers.

Potential for Discussion

To our knowledge, we are the first to estimate a hedonic model of the value of cannabis product characteristics. The size and depth of our dataset enables us to observe meaningful patterns in hedonic values and stimulate discussion for new insights and research directions in the rapidly expanding North American cannabis markets.

References

Gustafson, C. R., Lybbert, T. J., & Sumner, D. A. (2016). Consumer sorting and hedonic valuation of wine attributes: Exploiting data from a field experiment. *Agricultural Economics (United Kingdom)*, *47*(1), 91-103.

Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers. *European Economic Review*, *39*(9), 1697–1716.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55.

Cilicon. (2021). Differences between cartridges, pods, and disposable pens. Available from: https://www.ciliconplus.com/differences-between-cartridges-pods-and-disposable-pens.aspx, last accessed May 12, 2023.

Champoux, D. (2018). Vape pens & vaporizer pods, what's the difference? *Greentank Technologies.* Available from https://www.greentanktech.com/education/vape-pens-vs-pod/, last accessed May 12, 2023.

Weedmaps. (2023). Indica. Available from: https://weedmaps.com/learn/dictionary/indica, last accessed May 12, 2023.

Table 1. Variable descriptions and means

	Description	Mean	Std. Dev.
EighthPrice	Price per eighth ounce of flower, 1 gram of concentrate, including taxes and fees		
ConcentratePrice	Price per 1 gram of concentrate, including taxes and fees	*Whole Plant* *Distillate*	
THC	Percent THC content	*Flower* *Whole Plant* *Distillate*	
Hybrid	Indicator variable with a value of 1 if product is an Indica-Sativa hybrid; 0 otherwise	*Flower* *Whole Plant* *Distillate*	
Indica	Indicator variable with a value of 1 if product is Indica; 0 otherwise	*Flower* *Whole Plant* *Distillate*	
Sativa	Indicator variable with a value of 1 if product is Sativa; 0 otherwise	*Flower* *Whole Plant* *Distillate*	
Recreational	Indicator variable with a value of 1 if product is sold to recreational users; 0 if sold to medical patients only	*Flower* *Whole Plant* *Distillate*	
Disposable	Indicator variable with a value of 1 if product title or description includes 'disposable'; 0 otherwise	*Whole Plant* *Distillate*	
Pod	Indicator variable with a value of 1 if product title or description includes 'pod'; 0 otherwise	*Whole Plant* *Distillate*	
Cartridge	Indicator variable with a value of 1 if product title or description includes 'cart' or 'cartridge'; 0 otherwise	*Whole Plant* *Distillate*	
Cured	Indicator variable with a value of 1 if product title or description includes 'cured'; 0 otherwise		

Table 1. Variable descriptions and means			
	Description	Mean	Std. Dev.
Resin	Indicator variable with a value of 1 if product title or description includes 'resin'; 0 otherwise		
Rosin	Indicator variable with a value of 1 if product title or description includes 'rosin'; 0 otherwise		
CO2	Indicator variable with a value of 1 if product title or description includes 'CO2'; 0 otherwise		

Table 2. Results for hedonic price attributes of eighths

Variable	(1)	(2)
THC	0.881***	
	(0.018)	
log(THC)		0.478***
		(0.026)
Hybrid	–0.795***	–0.0243***
	(0.151)	(0.0044)
Indica	–1.038***	–0.0241***
	(0.157)	(0.0044)
Sativa	–1.292***	–0.0384***
	(0.163)	(0.0046)
Recreational	3.147***	0.106***
	(0.346)	(0.012)
Price Includes Tax	Yes	Yes
Week FE	Yes	Yes
Store FE	Yes	Yes
State FE	No	No
Observations	2,950,617	2,950,617
Adj. R^2	0.473	0.466

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Numbers in parentheses denote standard errors. Store-level cluster robust standard errors are used.

Table 3. Results for hedonic price attributes of concentrates

Variable	(3)	(4)	(5)	(6)
THC	–0.015		–0.042***	
	(0.013)		(0.011)	
log(THC)		–0.032		–0.050*
		(0.025)		(0.028)
Hybrid	0.464**	0.0094	0.213	0.0083*
	(0.235)	(0.0063)	(0.161)	(0.0048)
Indica	0.676***	0.018***	0.837***	0.022***
	(0.239)	(0.0062)	(0.175)	(0.0052)
Sativa	0.112	0.0040	0.504***	0.010*
	(0.245)	(0.0063)	(0.190)	(0.0056)
Recreational	1.952***	0.049***	1.615***	0.043***
	(0.571)	(0.018)	(0.323)	(0.013)
Disposable	1.081***	0.042***	2.593***	0.117***
	(0.366)	(0.0089)	(0.278)	(0.0099)
Pod	7.279***	0.118***	8.558***	0.2371***
	(0.507)	(0.010)	(0.379)	(0.011)
Cartridge	–0.852**	–0.032***	–2.938***	–0.127***
	(0.391)	(0.0099)	(0.323)	(0.012)
Cured	–6.301***	–0.152***		
	(0.353)	(0.0088)		
Resin	–4.073***	–0.093***		
	(0.394)	(0.009)		
Rosin	12.24***	0.214***		
	(0.642)	(0.013)		
CO2			4.899***	0.173***
			(0.666)	(0.021)
Price Includes Tax	Yes	Yes	Yes	No
Week FE	Yes	Yes	Yes	Yes
Store FE	Yes	Yes	Yes	Yes
State FE	No	No	No	No
Observations	451,965	451,965	2,487,154	2,487,154
Adj. R^2	0.860	0.705	0.737	0.689

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. Numbers in parentheses denote standard errors. Store-level cluster robust standard errors are used. Specifications (3) and (4) are for the whole plants concentrate segment; specifications (5) and (6) are for the distillate product segment.