Gateway or Cul de Sac?
Using Big Data to Assess Legal Recreational Marijuana and Changes in the Use of "Hard" Drugs

Robert Todd Perdue[1] and James Hawdon[2]

1.Elon University, Elon, NC, USA
rperdue@elon.edu

2.Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
hawdonj@vt.edu

Received April 26, 2019
Accepted for publication September 9, 2019
Published September 30, 2019

Abstract

Does legalizing marijuana result in decreased narcotic use as legalization proponents and some researchers claim? Or, conversely, does legalizing marijuana provide a gateway for experimentation and future "hard" drug use as critics of legalization and proponents of the gateway hypothesis attest? Now that several states in the U.S. have legalized marijuana for recreational use, it is possible to begin assessing the validity of these competing claims. Relying on a novel strategy for tracking trends in illegal drug use, we use Internet search queries, specifically Google Trends, to examine patterns of drug searches in four recreational marijuana states: Alaska, Colorado, Oregon and Washington. We find that search trend patterns for drugs are consistent following legalization, with increased mean searches for marijuana, cocaine and heroin, and decreased mean searches for methamphetamines, Oxycodone, and bath salts. While this finding seems to suggest some validity to the gateway hypothesis, we argue the opposite, for these general trends are also found at the national level. As such, the trends found in these recreational marijuana states generally do not differ from the nation as a whole, implying no significant gateway effect. We conclude that using big data to assess drug-using trends can inform the ongoing marijuana legalization debate.

Keywords: marijuana, big data, Google trends, legalization

Publication Type: Original research article

Preferred Citation: Perdue, Robert Todd and James Hawdon. 2019. "Gateway or Cul de Sac? Using Big Data to Assess Legal Recreational Marijuana and Changes in the use of 'Hard' Drugs." *Sociation*, 18(2), 20-28.

Introduction

On November 6, 2018, Michigan voters made their state the tenth in the union (along with Washington D.C.) to allow recreational marijuana. On the same day, the majority of Utah and Missouri voters chose to support initiatives legalizing marijuana for medical purposes, joining the thirty-one states where it was already legal. These developments are not surprising when one considers the Pew Research finding that 62% of Americans support marijuana legalization, up from 16% in 1990.[1] In short, societal acceptance of the drug has grown more rapidly than all but the most optimistic proponents could have ever hoped. Given this dramatic shift, it seems appropriate to revisit the debate concerning the extent to which marijuana serves as a gateway to harder drugs.[1]

The gateway hypothesis is well established, ingrained in official U.S. policy for decades. The Federal Bureau of Narcotics' 1965 pamphlet "Living Death" makes this perspective clear:

> …it cannot be too strongly emphasized that the smoking of the marihuana cigarette is a dangerous first step on the road which usually leads to enslavement by heroin. . . Ordinarily, a person is tempted first with marihuana cigarettes. He may not even know they are dope. Then, someone already addicted makes it easy to try some heroin. *Most* teenage addicts started by smoking marihuana cigarettes. *Never let anyone persuade you to smoke even one marihuana cigarette. It is pure poison.*[2] (emphasis in original)

The argument that marijuana is the first step to more dangerous drugs like heroin appears sensical in that it seems incongruous that a young person would begin drug experimentation by using heroin. One of the earliest explanations was offered by Erich Goode, who argued that marijuana is a "sociogenic" drug, one that is overwhelmingly used in group activities, which can initiate users into a drug subculture that then encourages further drug experimentation (1969).

Researchers began using the term "gateway drug" in the 1980s to describe substances, typically thought of as alcohol, tobacco and marijuana, that when consumed lead to consumption of harder, more dangerous drugs. The National Institute of Drug Abuse, however, points to the gaps in our knowledge of these gateways, noting that "young people who have used marijuana are at greater risk of using cocaine than those who have not," but go on to conclude:

> research has not fully explained any of these observations, which are complex and likely to involve a combination of biological, social, and psychological factors. In addition, most people who use marijuana do not go on to use "harder" drugs.[3]

Similarly, most marijuana advocates argue that legalization can actually *decrease* narcotic use because it can serve as a safer alternative to opioid painkillers that have greater addiction potential. Moreover, they argue legalizing marijuana can divert many young people from entering the drug sub-cultures they would likely encounter when attempting to procure a black-market product.

Now that several states have legalized marijuana for recreational use, it is possible to begin assessing the validity of these competing claims: Does legalizing marijuana for recreational purposes lead to increased "hard" drug use? Relying on a novel strategy for tracking trends in illegal drug use (Perdue, Hawdon, & Thames 2018), we use Internet search queries, specifically Google Trends, to examine patterns of searches for marijuana, cocaine, heroin, methamphetamine, Oxycodone, and bath salts in order to shed light on the marijuana gateway hypothesis.

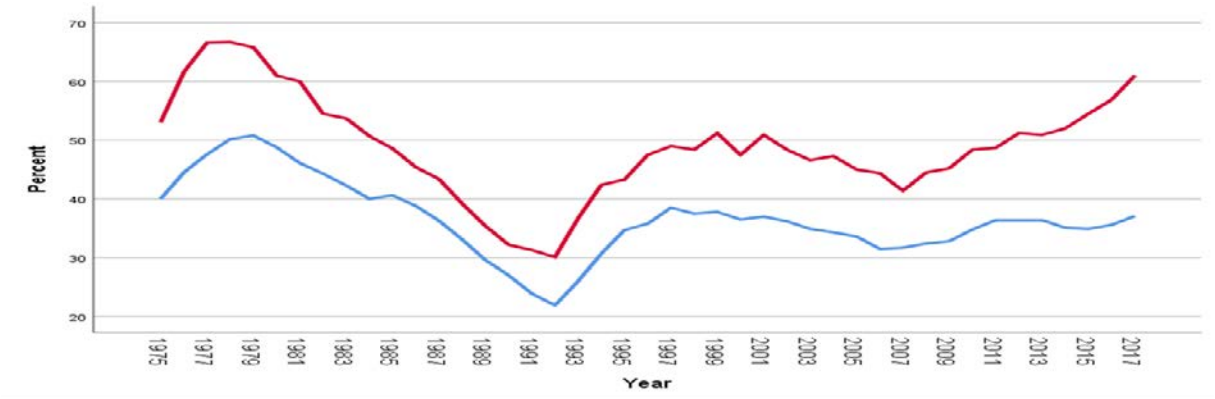
Gateways and Cul de Sacs

Most scholars agree that legalization increases marijuana use as the fear of legal sanctions undoubtedly deters some would be users, while legalization also conveys the message that use is socially acceptable (Hawdon 2005). For instance, Cerda et al. (2012) found that residents of states with legal medical marijuana used marijuana at higher rates than those from states without legal medical marijuana. They attribute this finding in part to "community norms supportive of the legalization of medical marijuana and of marijuana use" (2012, 22). This relationship between use and social acceptance is clearly shown in Monitoring the Future (MTF) data (see Figure 1). When we use MTF data from 1975-2018 to correlate the percentage of U.S. high school seniors who use marijuana annually with the percent of high school seniors in the preceding year who say using marijuana once or twice is acceptable, we find a striking Pearson's correlation coefficient of 0.908 ($p < .001$).

As the data in Figure 1 suggests, legalizing marijuana is likely to increase marijuana use. It is, however, less clear how legalization would affect hard drug use. In this debate, the two opposing arguments regarding marijuana can be thought of as gateways to future use or cul de sacs that do not take users further down the drug road:

1) **Gateway Argument:** Marijuana serves as a gateway to other, more dangerous drugs, especially for young people; legalization will only foster that process.
2) **Cul de sac Argument:** Legalization will decrease the use of other, harder drugs because marijuana users will not be lured into the drug subculture.

Figure 1. Trends in High School Senior Marijuana Use and Perceptions that Use Is Acceptable: 1975-2017



Percent HS Seniors saying using marijuana once or twice is acceptable **Percent HS Seniors using marijuana**

Most researchers addressing these arguments have reached few definitive conclusions. Indeed, findings tend to depend upon definitions, operationalization and data, while caveats abound. This is not to say the research has been unsound. Rather, it is because these are complex problems, with causality so difficult to pinpoint here. What actually constitutes a "gateway"? Take for instance two people:

> Person 1 uses marijuana, then tries heroin the next day. Person 1 never uses drugs again.
>
> Person 2 uses marijuana, then one year later tries heroin and becomes addicted.

The temporal gateway effect appears clear for Person 1, while the consequences are much more severe for Person 2. While these may not be the most common experiences for first time marijuana users, they are instructive in that they present the difficulty in attributing causality. Morral et al. (2002), for instance, argue drug users are more likely to have higher drug use propensities than non-users, and the ordering required for the gateway hypothesis is due to the fact that opportunities to use marijuana typically precede opportunities to use hard drugs, often by many years. As such, general liability to use drugs and opportunities to use are the primary factors in hard drug use.

Hall and Lynskey (2005, 39) contend the link between marijuana and harder drugs is in part explained by pre-existing traits, genetic and otherwise, but like Goode, they also highlight the importance of sub-culture: "the affiliation of cannabis users with drug using peers in settings that provide more opportunities to use other illicit drugs at an earlier age… supported by socialisation into an illicit drug subculture with favourable attitudes towards the use of other illicit drugs" are critical factors leading to eventual hard drug use. Others have argued that the *intensity* of marijuana use is indicative of future hard drug use. Fergusson et al. (2006) conducted a twenty-five year longitudinal study in New Zealand, finding a significant association between the frequency of cannabis use and other illicit drugs, but regarding the gateway hypothesis, they conclude "the extent to which these causal mechanisms are direct or indirect, remain unclear" (2006, 556). Similarly, Melberg et al. (2010) recognize the importance of nuance in their research utilizing a novel set of drug price data. These economists find that for most youths, cannabis use has little impact on later hard drug use. However, for a smaller group of "troubled youths," there is a sizable and significant gateway effect.

As this brief review suggests, the extent to which marijuana serves as a gateway drug is still unclear. To contribute to these discussions, we approach the debate from a different perspective by employing big data to shed light on this puzzling question.

Methods and Data

With the rise of commercial search engines, the Internet has increasingly helped shape public knowledge about topics ranging from cooking recipes and dog-grooming tricks to geo-politics and methamphetamine production. Indeed, the Internet is increasingly important in knowledge dissemination

and has become one of—if not the most—fundamental sources of news and information (Van Couvering 2008; Trevisan 2014). Because Internet search engines determine how an individual explores an issue, Internet searches are behavioral measures of an individual's interest in an issue (Granka 2009). As such, a given issue's salience is reflected by, and can be inferred based upon, the aggregated volume of Internet search queries related to that issue (Zheluk et al. 2014). Because search trend data (e.g. Google Trends) are free and provide dynamic, real-time observations concerning the public's attention to issues, Internet search trend analyses have benefits over traditional survey-based research methods (see Trevisan 2014). Such unobtrusive data that tracks individuals' real-world behaviors while avoiding problems that have traditionally plagued survey methods, such as telescoping, memory decay, social desirability, and selection bias, is clearly valuable (Fowler Jr. 2014).

Because of these advantages, Internet search analyses are an increasingly popular tool for scientific inquiry. Among the first to use Internet search data to track real-time trends, public health researchers have used Internet search data to track and forecast the spread of infectious diseases. These data have predicted outbreaks of influenza (Carneiro and Mylonakis, 2009; Eysenbach 2006; Ginsberg et al. 2009; Hulth, Rydevik, and Linde 2009; Pelat et al; 2009; Polgreen, Chen, Pennock, and Nelson 2008), listeriosis (Wilson and Brownstein 2009), dengue fever (Chan, Sahai, Conrad, and Brownstein 2009), and Lyme disease (Seiffer, Scharzwalder, Geis, and Aucott 2010). Search queries have also shed light on the health-related implications of governmental policies. For example, Ayers, Ribisl and Brownstein (2011) used searches to track the substitution of e-cigarettes and other electronic nicotine delivery systems after increases in the cigarette tax. Similarly, researchers have used these data to analyze international abortions rates relative to laws regulating abortion (Reis and Brownstein 2010).

While Internet searches have proved valuable in public health, the methodology is gaining popularity in other fields including economics and political science. For instance, Askitas and Zimmerman (2009) forecast trends in unemployment rates in Germany, Israel, Italy, and the United States Similarly, while Choi and Varian (2009; 2012) used Internet search analyses to predict automobile and real estate sales. Political scientists have used Internet search trends to measure public attention to numerous issues, including global warming, terrorism, and healthcare in the U.S. (Ripberger 2011). Similarly, the racial animus towards President Barak Obama's 2008 presidential campaign, which traditional survey methods were unable to detect, was uncovered using an Internet search query analysis (Stevens-Davidowitz 2012).

More recently, researchers have used Internet search query data to explore patterns in the manufacture and use of novel psychoactive drugs (NPDs). For example, this methodology was used to study illicitly made desomorphine--the injectable drug known as "krokodil" in Russia (Zheluk, Quinn, and Meylakhs 2014). The volume of krokodil-related search queries increased appreciably immediately following media and governmental reports declaring krokodil a significant social problem, but then these significantly decreased immediately following the implementation of legal restrictions against krokodil (ibid). Similarly, a study by Kapitany-Foveny and Demetrovics (2017) found that the volume of Internet entries—including search queries, online articles published, and online advertisements—increased after the NPD mephedrone (bath salts) was banned. Of note, the increase in Internet activity concerning mephedrone occurred while measures of Internet activity related to cocaine, heroin, and ecstasy remained consistent during the same period (Kapitany-Foveny and Demetrovics 2017). Finally, Perdue, Hawdon and Thames (2018) used Google Trends to track the emergence and use of five novel drugs (Adderall, salvia, snus, synthetic marijuana, and bath salts). Comparing the search results from Google Trends with the annual drug prevalence rates reported in the Monitoring the Future dataset, the researchers found that trends in NPD search queries were highly correlated with the MTF data. They conclude that Google Trends provide a useful proxy measure of NPD use prevalence (Perdue et al. 2018).

Taken together, these findings highlight both the utility of search query analysis in the study of drug use patterns and the influence of governmental regulations on these same patterns. As such we argue that such an approach will prove useful in assessing marijuana's role as a gateway. If we see rates of use for other drugs increasing after legalization, this would suggest that legal marijuana may very well be a gateway to other drugs. Conversely, if rates of use remain the same or decrease, legal marijuana may actual deter the use of other drugs.

The vast majority of Internet search trend studies discussed here have used Google's free, publicly available Google Trends data. Google Trends provides normalized population search popularity data that represents more than 73% of all desktop-based search queries and 81% of all mobile-based search queries in 2018 (Netmarketshare.com). Google Trends (https://www.google.com/trends/) offers global tracking of Google search queries that can be organized by country and various subnational units.

Google Trends explains their tracking methodology by stating:

> The numbers that appear show total searches for a term relative to the total number of searches done on Google over time. A line trending downward means that a search term's relative popularity is decreasing. But that doesn't necessarily mean the total number of searches for that term is decreasing.

We investigate trends in drug use in the first four states to legalize recreational marijuana in the U.S.: Alaska, Colorado, Oregon, and Washington. For each state and the nation as a whole, we collect monthly trends data from the time the state legalized recreational marijuana use through July 2019, and an equal number of months prior to legalization searching for marijuana, cocaine, heroin, methamphetamine, Oxycodone, and bath salts.[6] Thus, for Colorado, for example, we have Google Trends data for each drug for the 67 months prior to legalization (January 1, 2014) and 67 months post legalization. We then conduct T-tests comparing mean search trends pre-legalization to mean search trends post-legalization. We include national data to observe general drug-use trends to see if the patterns observed in the states that legalized recreational marijuana differed from the rest of the country in use patterns.

Table 1: T-Tests of Mean Google Trend Searches for Various Drugs Pre and Post Legalization

	Legalization Status (number of months)	Mean Google Search Marijuana (Standard Deviation)	Mean Google Search Cocaine (Standard Deviation)	Mean Google Search Meth (Standard Deviation)	Mean Google Search Oxycodone (Standard Deviation)	Mean Google Search Heroin (Standard Deviation)	Mean Google Search Bath Salts (Standard Deviation)
USA	Pre-Legal (67)	53.15 (7.78)	51.22 (5.56)	52.28 (14.75)	59.71 (11.17)	41.37 (6.82)	6.15 (12.69)
	Post-Legal (67)	58.70 (7.76)	53.48 (3.48)	36.31 (17.93)	35.82 (7.36)	62.58 (11.73)	2.48 (0.59)
	Pre/Post t-test	T=-4.14	T=-2.81	T=5.63	T=14.62	T=-12.80	T=2.37
Colorado	Pre-Legal (67)	36.78 (11.04)	45.52 (7.58)	21.31 (7.13)	34.21 (8.94)	36.82 (7.59)	5.85 (12.46)
	Post-Legal (67)	46.51 (9.53)	54.21 (6.12)	13.49 (7.77)	21.13 (6.16)	53.31 (9.88)	2.64 (0.92)
	Pre/Post t-test	T=-5.46	T=-7.30	T=6.07	T=9.86	T=-10.84	T=2.09
Alaska	Pre-Legal (55)	31.38 (12.53)	29.96 (8.05)	15.07 (6.43)	12.11 (8.84)	39.54 (13.42)	9.11 (13.55)
	Post-Legal (55)	49.25 (8.54)	30.51 (7.15)	9.80 (5.83)	6.74 (2.81)	52.05 (13.22)	3.16 (1.95)
	Pre/Post t-test	T=-8.74	T=-0.38	T=4.47	T=4.28	T=-4.93	T=3.22
Washington	Pre-Legal (61)	39.97 (9.79)	43.73 (6.01)	35.06 (7.85)	37.92 (16.06)	53.12 (11.14)	6.59 (12.90)
	Post-Legal (61)	43.72 (8.19)	46.66 (4.99)	20.54 (11.11)	18.14 (4.53)	68.46 (10.98)	2.96 (0.78)
	Pre/Post t-test	T=-2.30	T=-2.92	T=8.33	T=9.25	T=-7.62	T=2.19
Oregon	Pre-Legal	37.38 (17.83)	35.13 (10.68)	21.35 (8.39)	11.82 (4.24)	51.49 (10.91)	6.78 (14.99)
	Post-Legal	53.71 (11.63)	52.18 (6.40)	11.64 (4.23)	10.42 (2.97)	65.51 (7.90)	3.01 (0.96)
	Pre/Post t-test	T=-5.15	T=-9.18	T=6.93	T=1.81	T=-6.98	T=1.66

Note: All differences are significant at $p < .05$ level except Alaska cocaine, Oregon oxycodone, and Oregon bath salts.

Results

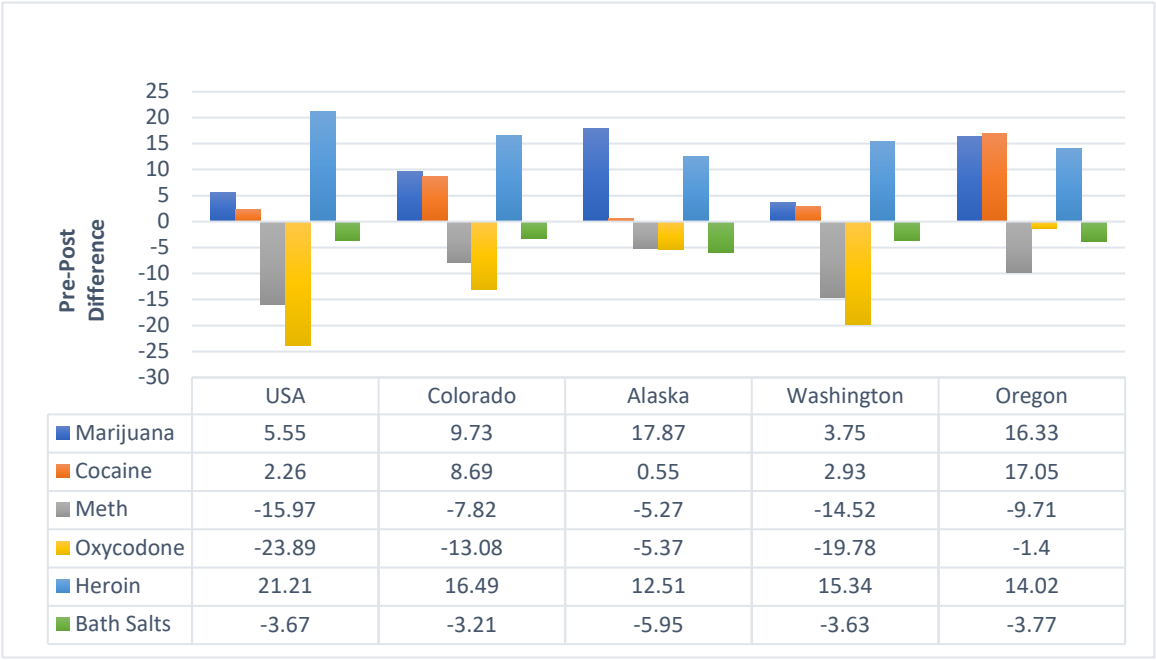
Table 1 presents the mean Google Trends searches for a variety of drugs pre- and post-legalization in the four states with legal recreational marijuana and the nation as a whole.[7] As can be seen in Table 1, searches for marijuana increased in all four states and the nation after legalization. Similarly, searches for heroin increased in all four states and the nation post-legalization. In contrast, searches for Oxycodone, Bath Salts, and methamphetamine decreased in all four states and the nation as a whole after recreational marijuana use was legalized in those four states. For cocaine, the pattern is less clear. Searches were slightly higher post-legalization for the nation (51.22 vs. 53.48) and Washington (43.73 vs. 46.66), and they were virtually unchanged in Alaska (29.96 vs 30.51); however, there was a clear increase in searches post legislation in Colorado (45.52 vs. 54.21) and Oregon (35.13 vs 52.18).

The differences in searches pre- vs post-legalization can be visualized in Table 2. While we do not mean to suggest that the magnitude of these differences can be compared across geographic units or are precise measures of rates of use of these drugs within these geographic areas, we do believe they reflect the general trends in use of these drugs within these areas. That is, while it would be misleading to suggest that the decrease in mean Google Trends searches for Oxycodone of 23.89 in the nation as a whole represents a significantly greater decrease in use than the decrease of 5.37 in Oxycodone searches in Alaska, we do believe that the decrease in searches for Oxycodone post-legalization reflects a general decrease of use in both the nation as a whole and the state of Alaska (Perdue et al. 2018).

Looking at the basic patterns above and assuming they reflect usage trends, we would argue that, unsurprisingly, legalizing recreational marijuana use leads to an increase in marijuana use. This is unsurprising since people are far more likely to engage in legal drug use than illegal drug use (see Hawdon 2005). These data also suggest that the legalization of recreational marijuana use in several states may lead to a liberalizing of attitudes and, in turn, greater use in the nation. Indeed, looking at the Monitoring the Future data, attitudes concerning the dangers of marijuana use have been liberalizing since about 2006, and rates of use among High School seniors have been trending upward since that time too (Johnston et al. 2019).

While it seems clear (and likely) that legalization of recreational marijuana use leads to greater marijuana use, there is no clear evidence that legalized recreational marijuana is a gateway to harder drugs. While it appears that heroin use and cocaine use may have increased in the states where recreational marijuana was legalized, this is also true for the nation as a whole. Moreover, the use of Oxycodone, methamphetamine, and Bath Salts appears to have

Table 2: Differences in Pre/Post Legalization Google Trend Searches for Various Drugs



	USA	Colorado	Alaska	Washington	Oregon
Marijuana	5.55	9.73	17.87	3.75	16.33
Cocaine	2.26	8.69	0.55	2.93	17.05
Meth	-15.97	-7.82	-5.27	-14.52	-9.71
Oxycodone	-23.89	-13.08	-5.37	-19.78	-1.4
Heroin	21.21	16.49	12.51	15.34	14.02
Bath Salts	-3.67	-3.21	-5.95	-3.63	-3.77

decreased in all of the states that legalized recreational marijuana use. That said, we also cannot argue that legalizing recreational marijuana use is a cul de sac, for Google Trends points to declines in these drugs at the national level as well. In short, what we find is that the trends in use in the four states that legalized recreational marijuana generally reflect trends for the nation as a whole.

Discussion and Conclusion

In a 2014 article in *Science*, Lazer and his colleagues pointed out two issues that can lead to mistakes when using big data: big data hubris and algorithm dynamics. By hubris, they mean overconfidence exhibited by substituting traditional research methods with big data collection and analysis, rather than merely supplementing them with big data (Lazer, Kennedy, King and Vespignani 2014). Indeed, with a topic as complex as human behavior, let alone drug use, it is indeed pure hubris to assume one tool is sufficient. Here, we argue that the use of big data is a useful adjunct to traditional methods that are limited by time and cost constraints, as well as the age-old challenges of researching illegality. This free and timely data should be used in conjunction with existent data such as overdose deaths, emergency room visits, drug seizures and drug arrests, before definitive statements are offered. The second limitation outlined by Lazer et al., algorithm dynamics, refers to the constant retooling of Google Trends algorithm. This is an important and often overlooked complication of Trends, for the data constantly change, making future comparisons difficult if not impossible (ibid). Moreover, and specifically related to drug research, big data can be distorted by moral panics and spikes in public interest. As such, researchers and the general public should be aware of the limitations of using Google Trends as scientific data.

Nevertheless, we are relatively confident in Google Trends ability to serve as an adequate proxy for estimating drug trends (see Perdue et al. 2018). A primary reason is that Google is the first place many Americans go for information, while the privacy of one's own computer eliminates the social stigma tied to curiosity about drugs and drug use. As noted in the review of the literature, definitive claims of causality are difficult regarding marijuana's role as a gateway drug, and we must conclude that this complexity is reflected in our findings. That said, it does appear that general U.S. drug trends are reflected in our analysis, as we find that the drug search trends found in the recreational states are similar to those found at the national level: following marijuana legalization we see searches for marijuana, cocaine and heroin increasing, while searches for bath salts, methamphetamines and Oxycodone decrease. These results echo general trends in American drug use; for instance, since 2010 we see heroin usage dramatically rise, while Oxycodone and other prescription opioid usage flattened (CDC 2018).

More specifically regarding the gateway hypothesis, we conclude that this data does not suggest legalizing recreational marijuana serves as a gateway for other drugs, as hard drug search trends generally are not significantly higher in legal states compared to the United States as a whole. Moreover, these data suggest that for three drugs (methamphetamine, Oxycodone and bath salts) the gateway hypothesis is not supported as searches actually decreased post-legalization. A closer examination in the future, however, may suggest alternative scenarios and possibilities. For instance, while we observe that cocaine searches have increased following legalization in all four states, we also see cocaine searches increase in the U.S. as a whole, lending support to the cul de sac argument. However, as traditional drug data collection methods "catch up" to these more current data, we will be able to explore the relationship between those data sources and these Google Trends data. More fined grained analyses will be possible and we will be able to ascertain whether the *mean search numbers* hold explanatory value. In other words, will we see, as these numbers suggest, that Colorado (8.69) and Oregon (17.05) witnessed dramatic increases in cocaine use following marijuana legalization? If so, then our understandings of the gateway hypotheses vis-à-vis these big data trends would have to be revised.

Indeed, this is the promise of using big data to analyze drug trends in general and the gateway hypothesis specifically. Going forward, and as more states legalize marijuana, we will have more traditional data at hand available for correlation with big data. This will allow us to more deeply understand the validity of the gateway hypothesis, and the extent to which legalization may reduce or eliminate contact with illegal drug subcultures, directing would-be users into the proverbial drug cul-de-sac. In turn, these data can better inform policymakers of the potential dangers (or lack thereof) of drug legalization, and (hopefully) help them strike an acceptable balance between individual freedom and public health.

Notes

1. http://www.pewresearch.org/fact-tank/2018/10/08/americans-support-marijuana-legalization/

2. We use the term "hard drugs" to refer to illicit drugs that have well documented negative health impacts, such as heroin and methamphetamine.

3. Federal Bureau of Narcotics, "Living Death: The Truth About Drug Addiction." Washington D.C., U.S. Government Printing Office, 1965.

4. National Institute on Drug Abuse: "Marijuana: Facts Parents Need to Know." (updated July 2018). https://www.drugabuse.gov/publications/marijuana-facts-parents-need-to-know/want-to-know-more-some-faqs-about-marijuana

5. When examining big data trends, choices must be made regarding the search terms to input. For instance, should researchers take into consideration misspellings or slang terms in their study? Here we argue that regardless of whether one searched "marijuana" or "marihuana", "heroin" or herion", the general thrust of the search is preserved. Stated differently, the noise that is undoubtedly created by these various searches will be flattened by the volume of searches. Arguably, the number of "wrong" searches is relatively consistent over time, thereby making general trends relatively unbiased. We believe our approach is analogous to analyzing Uniform Crime Reporting rates over time in that the actual numbers can be viewed as less important than general trends. Since we are not trying to estimate the rates of use and are instead interested solely in the trends in use, we therefore decided to use correct spellings and not to use misspelled words or slang terms for our study.

6. The authors are happy to share these data: rperdue@elon.edu

References

Anderson, M.D., Hansen, B., & Rees, D. I. (2015). Medical marijuana laws and teen marijuana use. *American Law and Economics Review*, *17*(2), 495-528.

Askitas, N. & Zimmerman, K. F. (2009). Google econometrics and unemployment forecasting. *Econometrics Quarterly*, 55(2), 107-120.

Ayers, J. W., Ribisl, K., & Brownstein, J. S. (2011). Using search query surveillance to monitor tax avoidance and smoking cessation following the United States' 2009 "SCHIP" cigarette tax increase. *PLoS One*, 6(3), e16777.

Carneiro, H. A., & Mylonakis, E. (2009). Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases*, *49*(10), 1557-1564.

Centers for Disease Control (CDC). 2018. "Opioid overdose: Understanding the epidemic" https://www.cdc.gov/drugoverdose/epidemic/index.html

Cerda, M., Wall, M., Keyes, K.M., Galea, S., & Hasin, D. (2012). "Medical marijuana laws in 50 states: Investigating the relationship between state legalization of medical marijuana and marijuana use, abuse and dependence." *Drug and Alcohol Dependence* 120 (1-3) 22-27.

Chan, E. H., Sahai, V., Conrad, C., & Brownstein, J. S. (2011). Using web search query data to monitor dengue epidemics: A new model for neglected tropical disease surveillance. *PLoS Neglected Tropical Diseases*, *5*(5), Article e1206.

Choi, H. & Varian, H. (2009). Predicting claims for unemployment benefits. *Google Inc*, 1-5.

Choi, H. & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record* 88(1), 2-9.

Eysenbach, G. (2006). Infodemiology: Tracking flu-related searches on the Web for syndromic surveillance. *AMIA Annual Symposium Proceedings*. American Medical Informatics Association, 2006, 244-248.

Fergusson, D. M., Boden, J.M., & Horwood, L.J. (2006). "Cannabis use and other illicit drug use: Testing the cannabis gateway hypothesis." *Addiction* 101, 556-569.

Fowler, F. J. (2014). *Survey research methods*. SAGE Publications.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature* 457(7232), 1012-1014.

Goode, E. (1969). "Multiple drug use among marijuana smokers." *Social Problems* 17, 48-64.

Granka, L. (2009). Inferring the public agenda from implicit query data. *Proceedings of the Workshop on Understanding the User—Logging and Interpreting User Interactions in Information Search and Retrieval*. SIGIR-2009, Boston, MA., July 19-23 2009.

Hall, W.D. & Lynskey, M. (2005). "Is Cannabis a gateway drug? Testing hypotheses about the relationship between cannabis use and the use of other illicit drugs." *Drug and Alcohol Review* 24, 39-48.

Hawdon, James. 2005. *Drugs and alcohol consumption as a function of social structure: A cross-cultural sociology*. Lewiston, N.Y.: Mellen Press.

Hulth, A., Rydevik, G., and Linde, A. (2009). Web queries as a source for syndromic surveillance. *PLoS One*, 4(2), e4378.

Kapitany-Foveny, M. and Demetrovics, Z. (2017). Utility of web search query data in testing theoretical assumptions about mephedrone. *Human Psychopharmacology*, 32.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science*, 343(6176), 1203-1205. doi:10.1126/science.1248506

Melberg, H.O. & Jones, A.M., Bretteville-Jensen, A.L. (2010). "Is cannabis a gateway to hard drugs?" *Empirical Economics* 38, 583-603.

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E. (2019). Monitoring the future national survey results on drug use 1975-2018: Overview, key findings on adolescent drug use. Ann Arbor: Institute for Social Research, University of Michigan.

Morral, A. R., McCaffrey, D.F., & Paddock, S.M. (2002). "Reassessing the marijuana gateway effect." *Addiction* 97, 1493-1504.

Netmarketshare. (2018). *Search Engine Market Share*. http://netmarketshare.com

Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., & Valleron, A. J. (2009). More diseases tracked by using Google Trends. *Emerging infectious diseases*, *15*(8), 1327.

Perdue, R.T., Hawdon, J. & Thames, K. (2018). "Can big data predict the rise of novel drugs?" *Journal of Drug Issues*. Vol. 48(4).

Polgreen, P. M., Cehn, Y., Pennock, D. M., Nelson, F. D., & Weinstein, R. A. (2008). Using internet searches for influenza surveillance. *Clinical Infectious Diseases*, 47(11), 1443-1448.

Reis, B.Y. & Brownstein, J.S. (2010). Measuring the impact of health policies using Internet search patterns: The case of abortion. *BMC Public Health*, *10*, 514.

Ripberger, J. T. (2011). Capturing curiosity: Using Internet search trends to measure public attentiveness. *Policy Studies Journal*, 39(2), 239-259.

Seiffer, A., Schwarzwalder, A., Geis, K., & Aucott J. (2010). The utility of Google Trends for epidemiological research: Lyme disease as an example. *Geospat Health*, 4(2), 135-137.

Stevens-Davidowitz, S. (2012). The effects of racial animus on a black presidential candidate: Using Google search data to find what surveys miss. *SSRN Journal*. 9-5-2014.

Suhoy, T. (2009). *Query indices and a 2008 downturn: Israeli data*. Bank of Israel.

Trevisan, F. (2014). Search engines: From social science objects to academic inquiry tools. *First Monday*, 19(11).

Van Couvering, E. (2008). The history of the Internet search engine: Navigational media and the traffic of commodity." *Web Search. Multidisciplinary Perspectives (Information Science and Knowledge Management)*, 177-206.

Wilson, K., & Brownstein, J. S. (2009). Early detection of disease outbreaks using the Internet. *Canadian Medical Association Journal*, *180*, 829-831.

Zheluk, A., Quinn, C., & Meylakhs, P. (2014). Internet search and krokodil in the Russian Federation: An infoveillance study. *Journal of Medical Internet Research*, 16(9), e112.