WHAT ARE THE ODDS: EVIDENCE FROM OHIO ON THE RELATIONSHIP BETWEEN ONLINE LOTTERY GAMES & ONLINE SPORTS BETTING

Heervansh M. Keswani

A THESIS

Presented to the Faculty of Miami University in
partial fulfillment of the requirements for the degree of

Master of Arts

Department of Economics

The Graduate School
Miami University
Oxford, Ohio

2024

Dr. Charles Moul, Advisor
Dr. Peter Nencka, Reader
Dr. Mark Tremblay, Reader

ABSTRACT

Lotteries are an established source of revenue for state governments across the US. Following the 2018 SCOTUS ruling that allowed states to legalize sports betting, several states seized the opportunity to create an additional source of revenue via its legalization and regulation. I exploit the staggered nature with which different states implemented their respective legislation to analyze associated impacts on online lottery activity in Ohio, where neighboring Michigan, Indiana, Pennsylvania and West Virginia all legalized sports betting before the Buckeye state. Leveraging a rich dataset of online lottery transactions and a variety of difference-in-difference specifications, I study the impact of having the option of legal online sports betting on online lottery activity in Ohio. I present cross-sectional and time-series models to analyze my predicted effects, finding mixed interactions between online lottery games and the availability of online sports betting. Overall, my results suggest that the relationship between the two gambling "products" is driven by several other factors, observed and unobserved, beyond simply proximity to a jurisdiction where an individual enjoys a choice between online sports betting and online lottery games.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I’d like to express my gratitude to Dr. Mark Tremblay and Dr. Charles Moul for the opportunity to work on this project and their continued guidance through its completion. My study would not have been possible without data from the Miami University Institute for Responsible Gaming, Lotteries, and Sport (MUIRGLS). Thank you to Dr. Peter Nencka for liaising with MUIRGLS to obtain updated data and providing valuable input on empirical technicalities, as well as Dr. Riley Acton for her suggestions on how to best convey complicated results.

1. Introduction

Idiosyncratic attitudes towards risk and the “ticket” to dream big for a nominal cost have attracted people to participate in low-probability lotteries since time immemorial. Governments have long sought to capitalize on this by using lotteries to raise funds for various undertakings. Early examples range from the Han Dynasty’s use of a keno-style game in 200 B.C.E to raise funds for the construction of the Great Wall of China[1], to the lottery organized by the Golden Ambrosian Republic to fund the fifteenth century Milanese War of Succession against the Republic of Venice [2]. By the early modern period, lotteries were a common method of raising funds for provision of public infrastructure[3], helping the poor[4], and funding military budgets[5] across most European powers.

In the nascent United States, the establishment of the first permanent colony in Jamestown, Virginia, was financed in part by a lottery organized by the Virginia Company of London. Similar lotteries quickly became a popular way to quickly and “fairly” raise money in a society that was extremely averse to taxation. Early American examples included the Province of Massachusetts Bay using a lottery to fund its “Expedition against Canada” and Benjamin Franklin financing the purchase of cannons for the defense of Philadelphia in the War for Independence. Taxes are now the main source of revenue for governments at the state and federal level in the United States[6], but most states still maintain lotteries to raise money for earmarked purposes such as public-school funding and college scholarships for low-income students[7].

Though they have stood the test of time, lotteries are rarely bettors’ preferred method of gambling. A key cause of this is that miniscule probability of winning leads to an expected payout lower than price of a lottery ticket, mechanically creating a relatively high and unattractive effective price. Secondly, the process that determines the winner is opaque, especially when compared to betting on the outcomes of some observable event like a sports match. Forrest et al. (2010) show empirical evidence that people only choose lottery tickets over sports bets when the jackpot is unusually high, thereby decreasing the effective price of a lottery ticket.

When Congress made gambling on sports effectively illegal across most of the United States by passing the Professional and Amateur Sports Protection Act (PAPSA) in 1992, state lotteries were largely protected from competition from sports betting. However, PAPSA was declared to

---

[1] The connection between keno and the Great Wall's construction is a popular anecdote, often repeated in gambling literature. The game's origins are indeed traced to ancient China, coinciding with periods of the Wall's construction, though definitive evidence directly linking keno to the Wall's financing remains elusive.
[2] The Ambrosian Republic: *A History of Milan under the Sforza*, 1907, Methuen & Co: London, Cecilia M. Ady, p. 49
[3] https://www.smithsonianmag.com/smart-news/queen-elizabeth-i-held-englands-first-official-lottery-nearly-450-years-180957804/
[4] R. Shelley (1989). *The Lottery Encyclopedia*. Austin, TX: Byron Pub. Services. p. 109
[5] Stigler, S. M. (2022). Casanova’s Lottery: The History of a Revolutionary Game of Chance, University of Chicago Press
[6] https://www.taxpolicycenter.org/briefing-book/what-are-sources-revenue-federal-government#:~:text=Over%20half%20of%20federal%20revenue,from%20a%20mix%20of%20sources.
[7] https://blog.jackpocket.com/where-lottery-money-goes-in-every-state/

be unconstitutional by the Supreme Court of the United States (SCOTUS) in 2018, with a 6-3 majority of Justices opining that it violated the Tenth Amendment.[8] Following this ruling, several states exploited the opportunity to monetize sports betting via legalization, which yields revenue from licensing fees for sports books, taxes from sports betting companies, and taxes levied on bettors' winnings. The staggered timing in which different states passed their respective sports betting legislation led to natural experiments in how legalization of sports betting in a neighboring state may impact lottery gambling in a state where sports betting is illegal.

My thesis leverages one such experiment as it arose in the case of Ohio. With the exception of Kentucky[9], all the states bordering Ohio legalized sports betting before Ohio itself did in 2023. From data that records transactions made by users with an Ohio Lottery account, I observe online lottery activity in different geographic areas of Ohio. This allows for analysis of online lottery activity in regions close to each of the four state borders relative to regions not within conveniently accessible distance of the legal option of online sports betting. My results show mixed interactions between demand for the two products, indicative of lottery activity being driven by factors other than proximity to an alternative gambling option in sports betting.

Previous studies have found evidence of lottery demand cannibalization by alternative gambling options. Cummings (2017) shows evidence of this by studying the effect of proximity to casinos on lottery sales in Maryland, as do Siegel and Anders (2001) who consider the impact of Indian casinos on state lotteries in Arizona. However, some research also suggests possible complementarity between gambling products. Walker and Jackson (2008) assert that while some gambling industries in the US cannibalize each other (like casinos and lottery), other forms of gambling, such as pari-mutuel wagering and charitable gaming showed mixed interactions, with some evidence of complementarity.

Leveraging the previously described natural experiment contributes to this literature an analysis of the relationship between two online gambling "products" – sports betting and lottery games. Improving technology and the draw of reaching wider audiences were already pushing society to shift more activities online, including gamble, before the circumstances of the COVID-19 pandemic rapidly accelerated that process. Today, around 90% of sports betting in most US markets happens online[10] and about 40% of the total Ohio Lottery revenue comes from online transactions.[11]

The next section provides additional context on the lottery in Ohio and advent of legal, online sports betting. Section 3 describes the data and empirical strategy, and a discussion of my main results follows in Section 4. Section 5 frames these results in the present-day, post-legalization context of "competing" revenue sources for the Ohio government and gives a brief overview of policy implications therein. Section 6 concludes.

---

[8] The Tenth Amendment to the US Constitution gives states the right to pass independent legislation on matters not delegated to the federal government.

[9] Kentucky legalized sports betting on September 7, 2024, with online options available beginning September 28, 2024.

[10] https://www.gamingtoday.com/revenue/

[11] https://www.ohiolottery.com/getattachment/1cd1a42b-660e-45f2-87e7-13a2b5b1501b/ACFR_FY2023_FINAL_12062023_reduced.pdf?lang=en-US

2. Background

2.1 Ohio Lottery

In 1973, voters in Ohio approved the creation of the Ohio Lottery Commission. So began the Ohio Lottery, with its first game, *Buckeye 300,* launching in August 1974. In 1983, the Ohio General Assembly began earmarking lottery profits for education, a designation made permanent by Ohio voters approving a constitutional amendment to this effect in 1987. As of 2020, the Ohio Lottery contributed over $26 billion to public education funding in the state.

In the fiscal year ending June 2020, total Ohio lottery sales totaled $4.3 billion, with approximately $1.27 billion going to the Lottery Profits Education Fund (LPEF)[12]. Of this total revenue, the biggest component was the face-to-face sale of lottery tickets with $1.87 billion, followed by online ticket sales of $1.6 billion and video lottery terminals (VLTs) contributing $820 million. The online transactions are the subject of my interest since each transaction records the user ID of the individual making the bet, which can be matched in the user database to find the address associated with the account which made the transaction. This allows me to geo-code the data and observe online lottery activity in Ohio at the zip code tabulation area (ZCTA) level.

The Ohio Lottery app offers a number of online games, with bettor activity highly fragmented among the various products available. *Pick 3* (11.5%), *Pick 4* (6.9%), *Mega Millions* (2.9%) and *Powerball* (2.5%) tickets represent only24% of online lottery sales with the other games individually accounting for smaller fractions of revenue. Instant scratcher-style games cumulatively comprise over 50% of online lottery activity, but this is fragmented into a multitude of different games. The top four online games hence contributed a total of roughly $384 million to revenue in the fiscal year 2020, compared to the $1.87 billion that offline games raked in the same year. Considering these relative shares, I am unable to divide my study to analyze individual online lottery games. I instead consider all online lottery transactions to be purchases of a single composite gambling commodity.

2.2 Sports Betting Legalization & Online Availability

The exceptions within the 1992 PAPSA legislation allowed for sports betting in states that already had existing regulatory framework in place, which were Nevada, Delaware, Oregon and Montana. New Jersey was racing to formalize its regulation as well but was not able to do so before PAPSA took effect, thus sowing the seed of events that transpire to create my treatment shock. In 2010, recognizing the revenue "lost" to these four states and illicit entities, New Jersey sought to pursue legal recourse. State Senators who brought the issue to a district court found that such a suit could only be filed by the State Governor, and so was born *Christie v. National Collegiate Athletic Association*. The Supreme Court accepted to hear the case in 2017, combining it with another lawsuit, *NJ Thoroughbred Horsemen v. NCAA*, which posed the same question: Was it unconstitutional to have a federal ban on sports betting? Phil Murphy assumed office as Governor of New Jersey in January 2018, earning him the privilege of being enshrined by the Supreme Court's landmark *Murphy v. NCAA* ruling in May that indeed, it was. Pennsylvania already had a law passed in 2017 that would authorize sports betting in the state if

[12] https://www.ohiolottery.com/cms/getattachment/042ab8ac-aa12-4bf5-b90b-d80de688ba3c/CAFR_FY20.pdf

federal law allowed states to regulate the activity[13]. Several other states (including West Virginia) had bills on the docket in state legislature and mechanisms in place to legalize sports betting but were waiting on a ruling from the Supreme Court[14]. Rhode Island even proactively included sports gambling revenues in the state's budget for the upcoming fiscal year, more than a month before the court's decision, indicating the readiness of states to capitalize on a potential source of revenue.

With the monies at stake and legal context of operating licenses involved, the territorial limits of legal online sports betting are stringently enforced. Every sportsbook that provides online betting services is required to ensure that its online users are within the geographic boundaries of a state in which it has a valid operating license. My field research confirmed the use of mandatory geo-verification protocol by all licensed online sports betting platforms in Michigan, Indiana, West Virginia, and Pennsylvania. My hypothesis of people in regions of Ohio near states with legal online sports betting exercising the option of betting on sports therefore requires that they physically cross the border to do so. While this introduces a cost to online sports betting that lacks a symmetric counterpart in the case of online lottery gaming, it proves crucial to my study since I am able to define a mechanical treatment in terms of proximity to relevant state borders.

2.3 Overview of Adjacent Literature

In addition to the papers referenced in other sections that are directly relevant to my thesis, this section presents a brief overview of other studies in the literature that are outside the direct scope of my study but nonetheless offer valuable context.

Theoretical foundations of gambling and lottery behavior were laid by the work of Friedman and Savage (1948) and Markowitz (1952) on utility functions and risk preferences. These papers sought to explain why individuals might participate in lotteries, suggesting that the curvature of the utility functions changes as income changes. Early studies on effective price (Vrooman 1976; Vasche 1985; Mikesell 1987) find no statistically significant of effective price on lottery ticket sales. Later studies by DeBoer (1986), Clotfelter and Cook (1989), and Miller and Morey (2003) show that while effective price may not be a significant determinant of lottery demand, there is a negative relationship between lottery sales and the takeout rate (percentage of winnings collected if opting for an immediate collection rather than annuity payout). Several studies also consider the own-price elasticity of demand for lottery tickets, finding elastic demand (Farrell and Walker, 1999; Farrell et al. 1999; Papachristou and Karamais 1998) in some markets and inelastic demand (Lee, Lin, and Lai, 2010) in others.

My study does not focus on the relationship between different online lottery games for reasons outlined previously, but several studies highlight underlying patterns and differences therein. Farrell and Forrest (2008) present evidence of electronic gaming machines and online Keno cannibalizing high jackpot lottery games in Australia, while Gulley and Scott (1993) find widely varying elasticities for two different jackpot games within the Massachusetts Lottery.

[13] https://www.cnn.com/2018/05/14/politics/sports-betting-ncaa-supreme-court/index.html
[14] https://www.usatoday.com/story/sports/2018/05/14/sports-gambling-status-every-state-after-supreme-court-ruling/607334002/

Several studies present evidence from casinos pertinent to how sports betting legalization affected revenues from slot machines. Abarbanel et al. (2011) analyzed daily revenues from slot machines and sports book wagers at a Las Vegas casino, finding no statistical relationship between the two. Suh and Tsai (2013) examined the relationship between daily slot machine revenues and poker players at two Las Vegas casinos, reporting no statistical relationship. Lucas (2014) studied the relationship between daily slot machine revenues and sports book wagers at three Las Vegas casinos, finding no relationship in two casinos and a positive relationship in the third. Several of these studies, as well as others are included in an extensive (if dated) survey of the literature on lotteries by Grote and Matheson (2011).

3. Data & Empirical Strategy

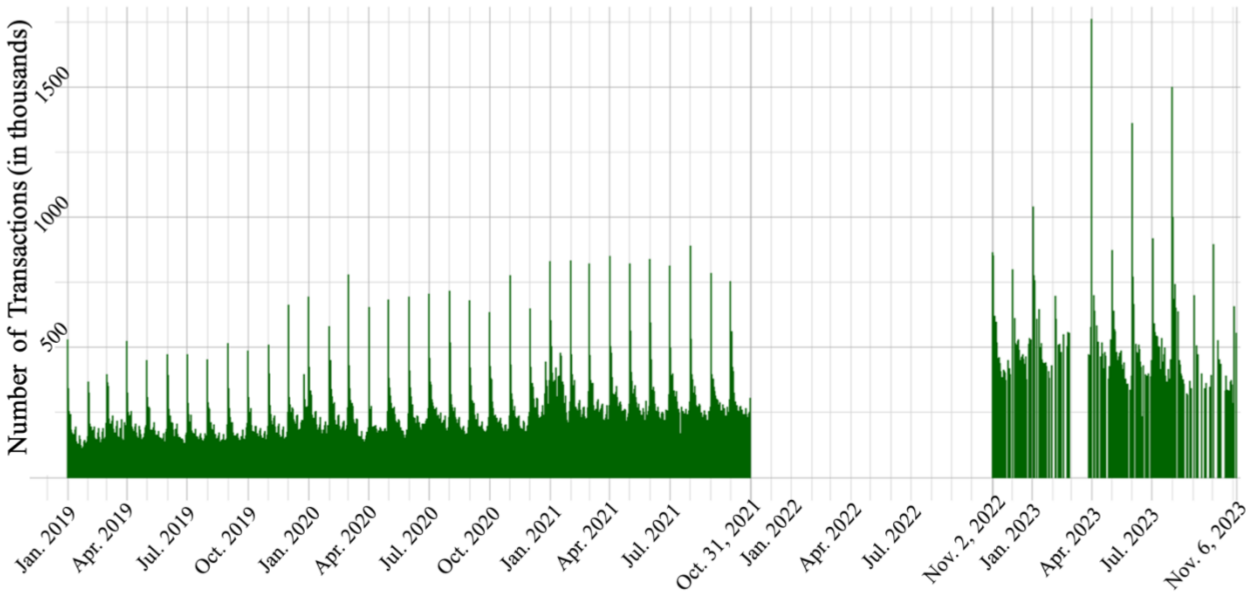
3.1 Lottery Data

Miami University Institute for Responsible Gaming Lotteries and Sports (MUIRGLS) collects data on online transactions made within the Ohio Lottery system of games, with each observation capturing the coupon cost (amount wagered), game played, user ID associated with the account that made the transaction, and a timestamp.

I use the user ID to match each transaction to the database of Ohio Lottery users, also provided by MUIRGLS. The user database captures the date on which an account was created, when it was most recently used, and, crucially, the zip code of the address associated with the account. Matching this information to the transaction data allows for each bet to be labelled with the zip code of the bettor’s home address. Given the spatial nature of the treatment, this information is critical, and observations with incorrect, non-Ohio zip codes (mistyped or otherwise) are hence dropped. The cleaned data thus represents online, paid Ohio Lottery transactions made by accounts associated with an Ohio zip code.

I aggregate this data to the zip-date level, giving me 1,420,525 observations across $i = 1{,}122$ zip code tabulation areas (ZCTA) on $t = 1{,}228$ dates. The transaction data is continuous from January 2019 through October 2021 but there are no data available for November 2021 through October 2022. The data resume in November 2022 but are only available for some dates at sporadic intervals, with random periods for which no transactions were recorded. Figure 1 illustrates this with a frequency distribution of transactions across dates.

Figure 1: Frequency Histogram of Observations by Date



I was unable to get explanations from MUIRGLS for the gap from October 2021 to November 2022 or for the sporadic data post-November 2022. However, MUIRGLS staff were able to confirm that the dates for which they did have data post-November 2022 are accurate and not in fact wrongly capturing bets from other dates that are missing. Consequently, upon aggregating to the zip-date level, I only have data for some of the days from November 2022 onwards, but the individual dates present in the dataset are accurate aggregations and comparable to dates that are in the continuous data up to October 2021.

3.2 Ohio Demographic Data

I compile ZCTA-level demographic data for Ohio from the 2019 American Community Survey (ACS) and the rural-urban commuting area (RUCA) classification for each ZCTA in Ohio from the United States Department of Agriculture (USDA) Economic Research Service. Table 1 presents the legend of RUCA codes, based on which I classify a ZCTA as metropolitan if it has a RUCA code of 1, 2 or 3. Table 2 contains summary statistics for selected demographic variables from the ACS data. Bolded variables are included directly in my analysis.

Table 1: Legend of RUCA Codes

Area type	RUCA	Area type	RUCA	Area type	RUCA
Primary flow within an urban area (UA)	1	Primary flow within large urban cluster (UC)	4	Primary flow within a small UC	7
Primary flow 30% or higher to UA	2	Primary flow 30% or higher to large UC	5	Primary flow 30%+ to small UC	8
Primary flow 10% to 30% to UA	3	Primary flow 10% to 30% to large UC	6	Primary flow 10% to 30% to small UC	9

$$metro_i = \begin{cases} 1, & if\ RUCA_i \in [1,2,3] \\ 0, & otherwise \end{cases} \tag{1}$$

Table 2: Summary Statistics of Ohio ZCTA Demographics from ACS 2019

	Min	Q1	Median	Mean	Q3	Max
Population	38	1521	4433	10362	15322	71189
Housing units	20.0	595.2	1715	4162	6346.8	26998
Median HHI	9583	46064	56217	57924	66551	173750
Poverty rate	0	0.0270	0.1110	0.1342	0.1730	0.6680
Unemployment rate	0	0.0270	0.0420	0.0506	0.0660	0.3050
Veterans	0	99.25	286	631.02	927	5586
Veterans Share of Population $(prop\ vets)_i$	0	0.0498	0.0629	0.0646	0.0767	0.2491
SNAP recipients	0	45.0	163	545.4	629.5	6964
Share of population receiving SNAP benefits $(prop\ snap)_i$	0	0.0204	0.0387	0.0515	0.0695	0.3810

I match the data for each ZCTA to the zip code associated with each transaction in my lottery data. Of the 1,137 zip codes in the lottery data, 1,122 had a matching ZCTA in the ACS data, with the remaining 15 being zip codes that only serve P.O. boxes and so do not have any demographic information. I filter out transactions associated with these zip codes, leaving me with 1,365,074 observations of online lottery activity at the zip-date level measured in amount wagered (lottery revenue) as well as the number of users and transactions. I compute per capita metrics of these measures based on the population of a ZCTA to create a standardized outcome variable that is comparable across ZCTAs of widely varying populations, densities, and demographics. Table 3 outlines summary statistics from my dataset of zip-date observations to build a general idea of online activity on the Ohio Lottery mobile application.

Table 3: Summary Statistics of Ohio Online Lottery Activity (zip-date level)

	Min	Q1	Median	Mean	Q3	Max
Amount wagered	1	67	349	1428	1506	563108
Transactions	1	14	70	258.1	290	18272
Users	1	3	10	30.92	33	937
Per Capita Amount Wagered $(per\ capita\ wager)_{it}$	0.00001	0.0269	0.0687	0.1371	0.1424	85.2338
Per capita Transactions	0.000014	0.0059	0.0139	0.0252	0.0275	22.6104
Per capita Users	0.000014	0.0012	0.0021	0.0031	0.0033	0.3450

3.3 Geographic Data

From the United States Department of Housing and Urban Development (HUD), I obtain the geographic coordinates of the population-weighted centroids of Ohio ZCTAs. Fortuitously, Ohio's borders with Michigan, Indiana and Pennsylvania are straight lines, which can be represented as equations of the lines connecting the borders' endpoints. The border with West Virginia is quite irregular as it is defined by the course of the Ohio river. Because the relevant aspect of these borders in my study is the ability to cross them, I compile a list of all motorable crossings of the Ohio River along the Ohio-West Virginia border. For each ZCTA, distance to the WV border is computed by selecting the minimum distance between the (population weighted) centroid and the coordinates of each of the border crossings. For the other three states it is computed as the shortest distance from the centroid to each of the straight lines that capture the MI, IN and PA borders. All distances are computed by finding the geodesic path (shortest distance between two points on the surface of an ellipsoid) using the World Geodetic System 1984 (WGS84) of mapping latitude/longitude coordinates within the "sf" package in R. This accounts for the curvature of the earth even though it has a negligible effect on calculations at my level of spatial proximity.

I create binary variables to denote whether a ZCTA is within a certain distance from each of these states' borders with Ohio. Table 4 outlines the structure of these variables.

Table 4: Construction of Distance Variables

Variable structure	Values
Distance Thresholds	
$State_s\ Distance_d$, where $s = \{MI, IN, WV, PA\}$ and d= $\{5, 10, 15, 20, 25\}$	$(State_s\ Distance_d)_i = \{ {1 \text{ if zip } i \text{ within } d \text{ miles of } s \text{ border} \atop 0\ otherwise}$ Ex: $(IN\ 20)_i$ captures zip codes with centroids within 20 miles of the Ohio-Indiana border
Distance Bins	
$State_s\ Dist_x, Dist_{x+5}$, where $s = \{MI, IN, WV, PA\}$, $x = \{5, 10, 15\}$	$(State_s\ Dist_x\ Dist_{x+5})_i = \{ {1 \text{ if zip } i\ between\ a \text{ and } b \text{ miles of } s \text{ border} \atop 0\ otherwise}$ Ex: $(MI\ 5, 10)_i$ captures zip codes with centroids over 5 miles but under 10 miles from the Michigan-Ohio border

The distribution of zip codes into different distance groups are shown in Table 5A. I also present select summary statistics for these distance bins as well as a 20-mile threshold of the relevant borders as that is how I define treatment in my main specification. Table 5B helps contextualize these with summary statistics for the control ZCTAs.

Table 5A: Group-wise Summary Statistics of Treated ZCTAs

	MI 0-5	MI 5-10	MI 10-15	MI 15-20	MI 0-20
# of ZCTAs	18	17	15	11	61
Avg. ZCTA population	11,121	13,560	8,322	4,406	9,902
Avg.daily users/ZCTA	41	48	21	12	33
Avg.daily revenue/ZCTA	$1,481	$1,986	$927	$579	$1,331
Avg.per capita wager	$0.13	$0.15	$0.13	$0.11	$0.13
Avg. % Metro commuting	72%	76%	73%	64%	72%
Avg.% vets/ZCTA	5.5%	6.5%	6.2%	7.5%	6.3%
Average % SNAP/ZCTA	7.6%	7.1%	2.9%	2.9%	6.5%
	IN 0-5	**IN 5-10**	**IN 10-15**	**IN 15-20**	**IN 0-20**
Number of ZCTAs	19	20	32	37	108
Avg.ZCTA population	4,844	7,643	12,959	9,961	9,520
Avg.daily users/ZCTA	7	24	46	42	34
Avg.daily revenue/ZCTA	$348	$873	$1,752	$1,392	$1,231
Avg.per capita wager	$0.08	$0.13	$0.14	$0.15	$0.13
Avg. % Metro commuting	16%	25%	50%	65%	44%
Avg.% vets/ZCTA	5.8%	5.9%	6%	4.9%	5.6%
Avg.% SNAP/ZCTA	3.4%	2.6%	6%	6.2%	5%
	WV 0-5	**WV 5-10**	**WV 10-15**	**WV 15-20**	**WV 0-20**
Number of ZCTAs	26	26	26	28	106
Average ZCTA population	7,574	2,519	3,259	2,422	3,915
Avg.daily users/ZCTA	14	6	7	7	9
Avg.daily revenue/ZCTA	$980	$330	$361	$309	$514
Average per capita wager	$0.13	$0.12	$0.09	$0.12	$0.12
Avg. % Metro commuting	62%	62%	31%	29%	45%
Avg.% vets/ZCTA	7.6%	7.2%	7.6%	6.4%	7.2%
Avg.% SNAP/ZCTA	7.2%	5.7%	6.9%	5.6%	6.4%
	PA 0-5	**PA 5-10**	**PA 10-15**	**PA 15-20**	**PA 0-20**
Number of ZCTAs	21	20	16	16	73
Avg.ZCTA population	6,482	7,792	11,528	7,848	8.246
Avg.daily users/ZCTA	21	28	38	23	27
Avg.daily revenue/ZCTA	$978	$1,304	$1,885	$1,208	$1,325
Average per capita wager	$0.14	$0.17	$0.16	$0.14	$0.15
Avg. % Metro commuting	62%	75%	56%	75%	67%
Avg.% vets/ZCTA	7.8%	7.1%	8.5%	7.6%	7.7%
Avg.% SNAP/ZCTA	5%	10.6%	6.1%	5.5%	6.9%

Table 5B: Summary Statistics for Control ZCTAs

Control ZCTAs: 801	Min	Q1	Median	Mean	Q3	Max
Population	38	1681	4930	11382	17810	65577
Users/ZCTA	1	3	11	33	37	937
Revenue/ZCTA	1	73	396	1552	1660	563108
Per capita wager	0.00002	0.02799	0.07071	0.13914	0.14492	85.23377
Metro % ZCTA	0	0	100%	53%	100%	100%
% vets/ZCTA	0	5%	6.2%	6.4%	7.5%	17.2%
% SNAP/ZCTA	0	1.8%	3.6%	4.9%	6.7%	38.1%

As evident from the tables above, there is considerable variation in demographic characteristics of ZCTAs near each of the four relevant borders. From the last column, I note that the parts of Ohio within 20 miles of Indiana or West Virginia are much less urban (44% and 45% metro respectively) than those within 20 miles of Pennsylvania (67% metro) or Michigan (72% metro). The region near Indiana also has lower levels of poverty (proxied by the proportion of population receiving SNAP benefits) and fewer veterans as a proportion of population than the 20-mile treatment zones near the other three state borders.

These characteristics also vary by distance, within the four different 20-mile thresholds. Interestingly, despite the evident variation in demographic characteristics, the group average per capita wager for all the distance groups described in Table 5A does not vary substantially. In fact, the average per capita wager for the control ZCTAs is also relatively similar at just under 14 cents. Given the variation in underlying demographic characteristics, the similarity in averages of my dependent variable, per capita wager, is reassuring to the credibility of my empirical design

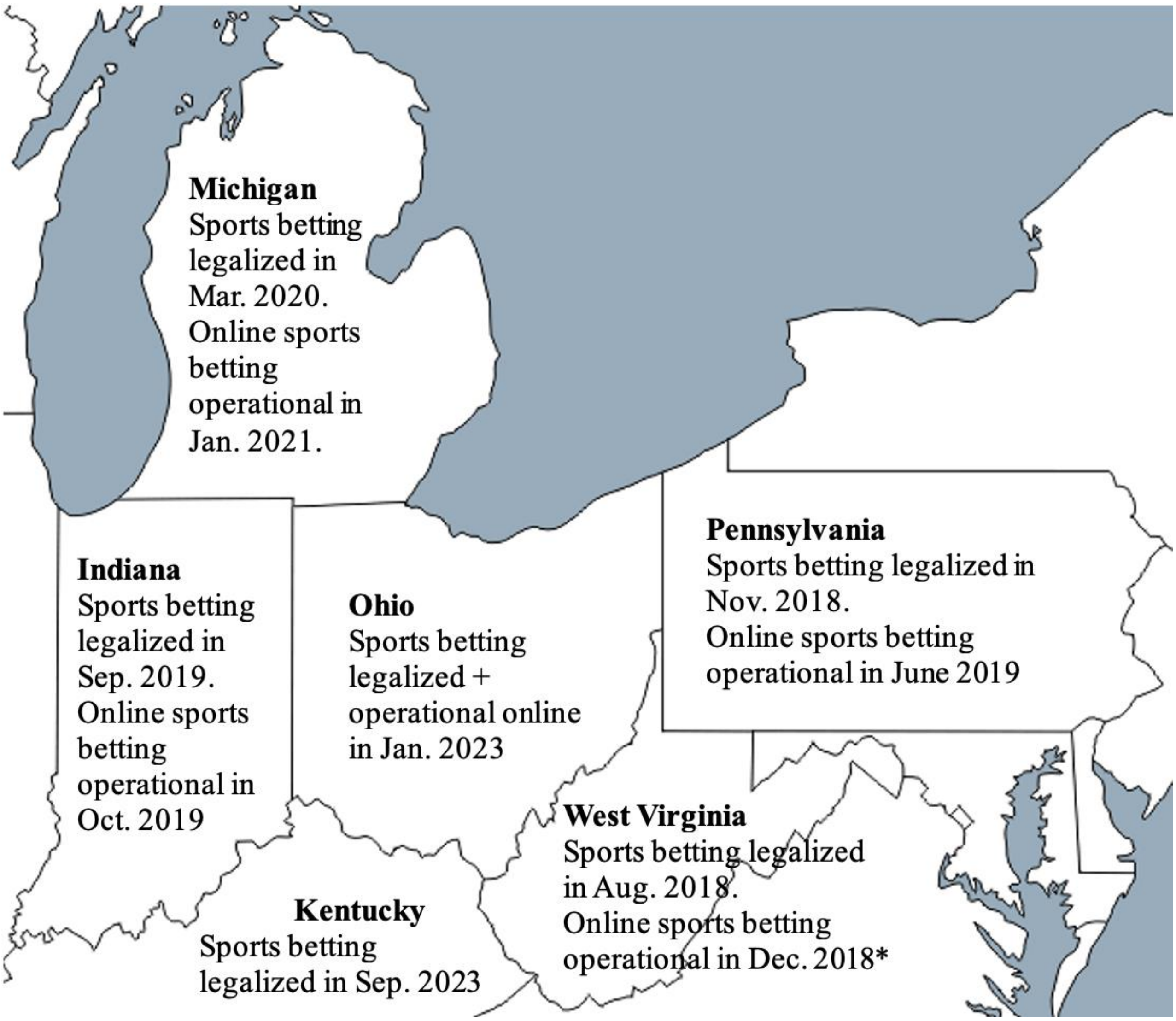
The binary variables indicating whether or not a ZCTA belongs to a border-proximity group that are the column headings for the above table are used to define the location component of treatment in my analysis. I project results from specifications on a ZCTA map of Ohio and the geo-data on ZCTA boundaries come from the TIGER/Line shapefiles available on the Census website.

3.4 Data on Sports Betting Legislation

I now define binary variables that indicate the legal status of sports betting in Michigan, Indiana, West Virginia, Pennsylvania, and Ohio on each date for which I have data. West Virginia and Pennsylvania had legal sports betting since before the start of my data, with the former even having online sports betting operational from December 2018. However, the single online sportsbook initially licensed in West Virginia faced solvency issues in 2019 as a result of which there was no online sports betting operational in the state from March 7, 2019, to August 26, 2019. Consequently, I do observe variation in the treatment indicator for West Virginia, despite its early adoption of online sports betting.

The distinction between face-to-face and online sports betting is important for my study since my hypothesis is predicated on the ability of people to drive across the border and bet on sports. With online sports betting, individuals can do so without having to go to a designated sports betting location like a casino. Furthermore, lottery games have many more face-to-face "venues" (gas stations, for example) than sports betting, but online versions of both games are comparable in accessibility. Hence, I use the dates on which online sports betting became operational in the states that legalized in defining the treatment in my analysis. Figure 2 denotes these dates and Equation 2 describes the binary variables hence defined for my data that spans 2019 to 2023.

Figure 2: Map of Ohio and Neighboring States Annotated with Sports Betting Legislation Information



*West Virginia halted online sports betting from March 7, 2019, to August 26, 2019

$$(online\ WV)_t = \begin{cases} 0\ , & March\ 7, 2019 \leq t\ \leq\ August\ 16, 2019 \\ 1\ , & otherwise \end{cases}$$

$$(online\ PA)_t = \begin{cases} 0\ , & t < May\ 31, 2019 \\ 1\ , & t \geq May\ 31, 2019 \end{cases} \qquad (online\ IN)_t = \begin{cases} 0\ , & t < October\ 3, 2019 \\ 1\ , & t \geq October\ 3, 2019 \end{cases}$$

$$(online\ MI)_t = \begin{cases} 0\ , & t < January\ 22, 2021 \\ 1\ , & t \geq January\ 22, 2021 \end{cases} \tag{2}$$

3.5 Defining Treatment

I am interested in the response of consumers of online lottery games to the option of online sports betting. To this end, I classify my zip-date level observations of $(per\ capita\ wager)_{it}$ as treated if ZCTA i is in proximity to Michigan, Indiana, West Virginia, or Pennsylvania and if date t is on or after online sports betting began in the ZCTA's nearby state. Equation 3 represents how this is achieved by interacting the binary variables for online sports betting availability (as explained in subsection 3.2) with the binary variables for proximity to state borders (as explained in subsection 3.4). I select 20 miles as the threshold for proximity to the border with a state that legalizes online sports betting, though I notice similar effects when using 15 miles and 25 miles as the threshold as well. The chosen 20-mile distance corresponds to roughly 25 minutes of driving time in the largely rural areas that comprise the treated group. Estimates from the main results that I present later using alternate proximity thresholds in defining treatment were also computed and are presented in the appendix.

$$\begin{aligned} (online\ MI\ 20)_{it} &= (MI\ 20)_i \times (online\ MI)_t \\ (online\ IN\ 20)_{it} &= (IN\ 20)_i \times (online\ IN)_t \\ (online\ WV\ 20)_{it} &= (WV\ 20)_i \times (online\ WV)_t \\ (online\ PA\ 20)_{it} &= (MI\ 20)_i \times (online\ PA)_t \end{aligned} \tag{3}$$

I initially model sports betting legalization as a homogeneous shock to all ZCTAs within 20 miles of a border with a state that legalizes online sports betting. Equation 4 explains how the observations in each of the four groups defined above are "collected" into a single treatment group by the following construction:

$$(online\ 20)_{it} = \begin{cases} 1\ , & if\ \begin{matrix} (online\ MI\ 20)_{it} = 1 \quad or \quad (online\ IN\ 20)_{it} = 1 \quad or \\ (online\ WV\ 20)_{it} = 1\ \ or \quad (online\ PA\ 20)_{it} = 1 \end{matrix} \\ 0\ , & \begin{matrix} otherwise \\ (control\ group) \end{matrix} \end{cases} \tag{4}$$

In later specifications, the location component of treatment is defined in terms of the distance bin variables described in Table 3 as opposed to the distance thresholds used in Equation 4. The series of treatment variables hence created is described in Equation 5:

$$(online\ State_s\ x, x+5)_{it} = \sum_{s=\left\{\substack{MI,IN,\\ WV,PA}\right\}} \sum_{x=0}^{3} [(State_s\ Dist_{5x}\ Dist_{5x+5})_i \times (online\ State_s)_t] \tag{5}$$

3.6 Empirical Strategy

I rely on a variety of difference-in-difference specifications to study the effect of the treatment – access to legal online sports betting at the cost of a 25-minute (or shorter) drive – on online lottery activity in Ohio. My data includes several measures of lottery activity. I choose the total amount wagered in a zip code on a date divided by the population of the ZCTA, logged for better interpretability, as the outcome variable for my regressions. Results from specifications with alternate measures of lottery activity as the outcome variable are included in the appendix. Two-way fixed effects at the zip code level and the date level capture the idiosyncratic heterogeneity in lottery preferences across ZCTAs and the variation in lottery activity due to temporal factors respectively. Equation 6 represents my baseline specification from which I extract zip fixed effects and date fixed effects. Equation 7 decomposes the zip fixed effects on the demographic variables previously summarized in Table 2.

$$log(per\ capita\ wager)_{it} = \beta(online\ 20)_{it} + \gamma_i + \delta_t + \varepsilon_{it} \tag{6}$$

$$\gamma_i = \gamma_0 + \gamma_1(metro)_i + \gamma_2(prop\ vets)_i + \gamma_3(prop\ snap)_i + \nu_i \tag{7}$$

I also decompose the date fixed effects on relevant temporal factors as shown in Equation 8, where $(time)_t$ captures the daily growth in lottery activity over time and dummy variables for days of the week control for weekday/weekend cyclical trends. The cumulative advertised jackpot value (in 100 millions) for the Powerball and Mega Millions lotteries is defined as $(jackpot)_t$ to capture the social perception of potential lottery winnings. I choose the advertised jackpot instead of effective price of lottery tickets for the two games in light of several previous studies, including Walker (1998), Gully et al. (2000) and Forrest et al. (2002), that find lottery demand is more influenced by overall jackpot size than the expected value of the game. I also include the square of this cumulative jackpot, because previous studies such as Forrest et al. (2010) and Forrest and Pérez (2011) find that lotteries cannibalize other forms of gambling only when the lottery jackpot is extraordinarily large, suggesting a non-linear relationship between jackpot value and online lottery demand. Plotting the date fixed effects also reveals large spikes on the first two days of the month which I hence include in the fixed effects' decomposition. I conclude with day-of-the-week fixed effects and $(OH\ legal)_t$, an indicator variable for when Ohio legalized (and launched online) sports betting starting in 2023.

$$\begin{aligned}\delta_t = \delta_0 + \delta_1(time)_t + \delta_2(OH\ legal)_t + \delta_3(jackpot)_t + \delta_4(jackpot^2)_t + \\ \delta_5(first\ of\ month)_t + \delta_6(second\ of\ month)_t + \delta_7(day\ of\ week)_t + \upsilon_t\end{aligned} \tag{8}$$

The remainder of my study builds on this design by including as regressors demographic variables interacted with the variable indicating treatment to allow for heterogeneity analysis of my estimates. I complement this with a longitudinal perspective by modeling lottery activity as a function of the treatment and its interactions with temporal variables.

I then relax the assumption of the treatment effect being homogeneous by state and split the treatment variable into separate components for each state (as defined in equation 4). Again, interactions with demographic variables allow for cross-sectional analysis and interactions with temporal variables reveal the effect of treatment given the underlying time trends in the data. Finally, to allow for added granularity in my heterogeneity analysis, I define treatment using the distance bins explained in Table 3, as opposed to a single distance threshold. As with the other specifications, I interact these treatment variables with the demographic variables to showcase the heterogenous effects of access to legal online sports betting on online lottery activity.

4. Results

4.1 Baseline Model & Decomposition of Fixed Effects

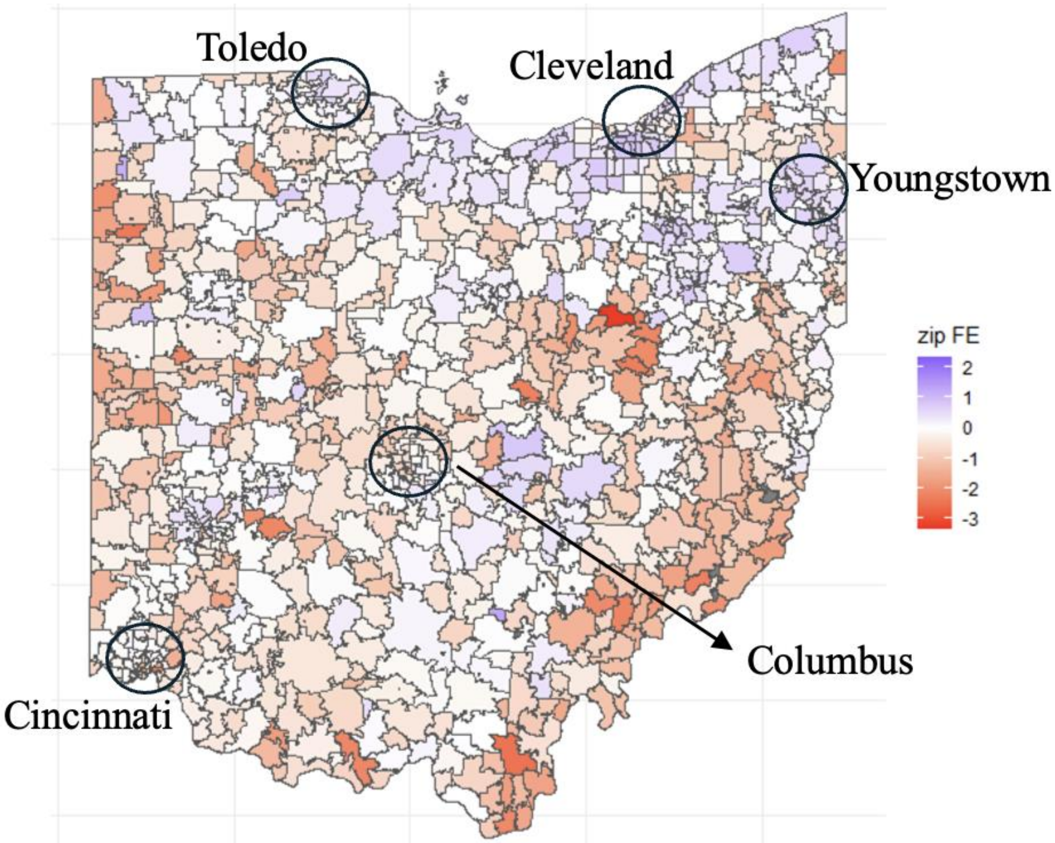
In Table 6, I present my baseline specification which corresponds to equation 6 applied to my full sample of dates from January 1, 2019, to November 4, 2023. The treatment indicator is the only regressor in addition to zip and date fixed effects.

Table 6: Baseline Specification with
Combinations of Fixed Effects

Intercept	-3.0289*** (0.0017)	---	---	---
$(online\ 20)_{it}$	0.0246*** (0.0041)	0.0047 (0.0041)	0.0459*** (0.0059)	-0.0085 (0.0064)
Zip Fixed Effects	N	N	Y	Y
Date Fixed Effects	N	Y	N	Y
Observations	1,365,074			
Adjusted R^2	0.00002	0.0463	0.20	0.25
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

I extract the zip fixed effects from this regression and project them on a ZCTA map of Ohio to present a cross-sectional view of the expected logged per capita lottery wager in each zip code. Blue (red) denotes larger (smaller) per capita wagers.

Figure 3: Estimated Zip Fixed Effects
For Ohio ZCTAs



To contextualize these results, note that the dependent variable in the regression is the log of the daily per capita wager in each ZCTA. The fixed effects around zero hence correspond to a roughly 1:1 ratio of a ZCTA's population to the dollar amount wagered on online lottery games. Most ZCTAs have a population-wager ratio of less than 1 as captured by the negative zip fixed effects on the map, and the ZCTAs with positive zip fixed effects represent areas of exceptionally high lottery activity with population-wager ratios greater than 1.

As evident from the map, there is considerable cross-sectional heterogeneity across ZCTAs in the daily level of online lottery activity. To understand the relevant drivers of this variation, I decompose the zip code fixed effects computed in the Table 5 specification on ZCTA demographic variables from the ACS 2019 data. These estimates are presented in Table 7 and correspond to Equation 7 described earlier

Table 7: Zip Fixed Effects Decomposed on Demographic Variables

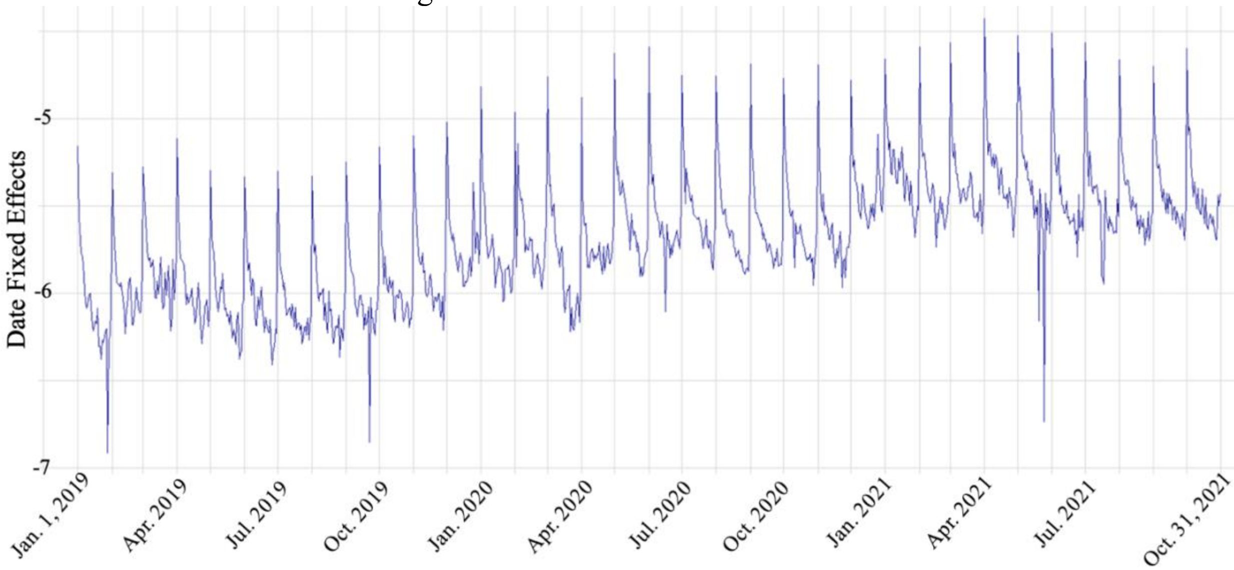
	$intercept$	2.2748***
		(0.0719)
	$(metro)_i$	-0.0341
		(0.0442)
	$(prop\ vets)_i$	3.452***
		(0.8825)
	$(prop\ snap)_i$	3.2368***
		(0.4944)
Observations		1122
Adjusted R^2		0.12
Significance Codes:	'*' 0.05 '**' 0.01	'***' 0.001

These coefficients highlight some of the observable factors driving the heterogeneity in online lottery activity across ZCTAs. They suggest higher participation in lotteries among veterans and among low-income groups, with the latter result in line with past studies in the literature[15]. Being a metropolitan area does not show statistical significance but is associated with a slightly lower level of online lottery activity.

In addition to the cross-sectional variation in data, I am also interested in understanding the variation in data over time. Figure 3 presents a time series plot of the date fixed effects from the regression reported in Table 5. Given the missing observations in my data, the figure plots these effects for the period during which MUIRGLS records continuous lottery data viz. January 1, 2019, to October 31, 2021.

[15] Haisley, Mostafa, and Loewenstein (2008) find evidence that low-income individuals are particularly drawn to purchasing lottery tickets because they provide an opportunity to correct for low-income status. Results from Clotfelter and Cook (1987), Clotfelter et al. (1999), Hansen (1995), Hansen et al. (2000), Lang and Omori (2009), Rubenstein and Scafidi (2002), and Welte et al. (2002) concur.

Figure 4: Plot of Date Fixed Effects



The most striking feature from the plot above is the spike at the beginning of every month. I tested a variety of hypotheses to explain this trend but was not able to conclusively identify the factors responsible. Among the explanations considered was the disbursement of paychecks on the first of the month, but pay schedules vary widely across industries and job profiles with many individuals receiving their wages bi-weekly or even weekly. Another theory was that this might be driven by the disbursement of Veterans' Association (VA) benefits on the first of the month, but that too shifts when the first of the month falls on a non-business day for which I see no corresponding evidence in the online lottery data. Some locals who play the lottery in Oxford, Ohio, which belongs to a ZCTA bordering Indiana, opine that bulk purchases of lottery tickets by one individual on behalf of their work colleagues may be responsible for the spikes. However, given the ease of individual access to *online* sports betting, this too seems unlikely. To better understand the drivers of this time-variation in the data, I decompose the date fixed effects on temporal factors that are relevant to online lottery activity. These estimates are presented in Table 8 and correspond to Equation 8 described earlier.

Table 8: Date Fixed Effects Decomposed on Temporal Variables

$intercept$	-6.0980 (0.0207)
$(time)_t$	0.0714*** (0.0018)
$(OH\ legal)_t$	-0.1810*** (0.0250)
$(jackpot)_t$	-0.0010 (0.0051)
$(jackpot^2)_t$	0.0007 (0.0004)
$(first\ of\ month)_t$	0.8804*** (0.0325)
$(second\ of\ month)_t$	0.4688*** (0.0305)
Observations	1,228
Adjusted R^2	0.76
Jackpot value in 100 millions, time in 100s of days	

The positive coefficient on $(time)_t$ highlights the growth in online lottery activity over the period for which I observe data. Day-of-the-week fixed effects are also included in this specification[16]. Though $(jackpot)_t$ doesn't show statistical significance in determining date fixed effects from the specification in Table 6, its quadratic term $(jackpot^2)_t$ does at the 90% level. This suggests that only when jackpot values grow very large do they induce increased online lottery activity. This is in line with studies in the literature cited earlier that suggest jackpot value only induces lottery activity when it is unusually high. The statistically significant negative coefficient on $(OH\ legal)_t$ indicates a negative association between in-state access to online sports betting and the date fixed effects influencing online lottery activity. A cautionary note on interpreting this coefficient is to contextualize it with the linear positive time trend. It hence represents a decline in growth rather than absolute reduction in online lottery activity. Furthermore, I cannot rule out the effects of the COVID-19 pandemic in slowing the growth trajectory. While "COVID-friendly" given the mobile platform, the draw of making risky "investments" perhaps waned during a time of such uncertainty.

The map and time series plot collectively highlight the need to allow for heterogeneous impacts of the introduction of nearby online sports betting. This is important at the cross-sectional level (across ZCTAs) as well as to capture variation over time. At the cross sectional-level I interact

[16] Not shown in Table 8, coefficients on the binary variables for day-of-the-week highlight higher lottery activity over the weekends that ebbs to its lowest around Monday-Wednesday.

my treatment indicator with ZCTA demographics to permit heterogeneity analysis. Likewise, for a longitudinal perspective, I interact my treatment indicator with temporal variables.

4.2 Main Specifications

4.2.1 Assuming legalization shocks from different states have a homogeneous treatment effect

I begin by modelling sports betting legalization as a homogeneous shock to all ZCTAs within 20 miles of a border with a state that legalizes online sports betting. Equation 9 represents the cross-sectional version of this model where I interact the demographic variables from Table 6 with the treatment variable $(online\ 20)_{it}$ defined in Equation 5. Regression estimates corresponding to Equation 9 are reported in Table 9.

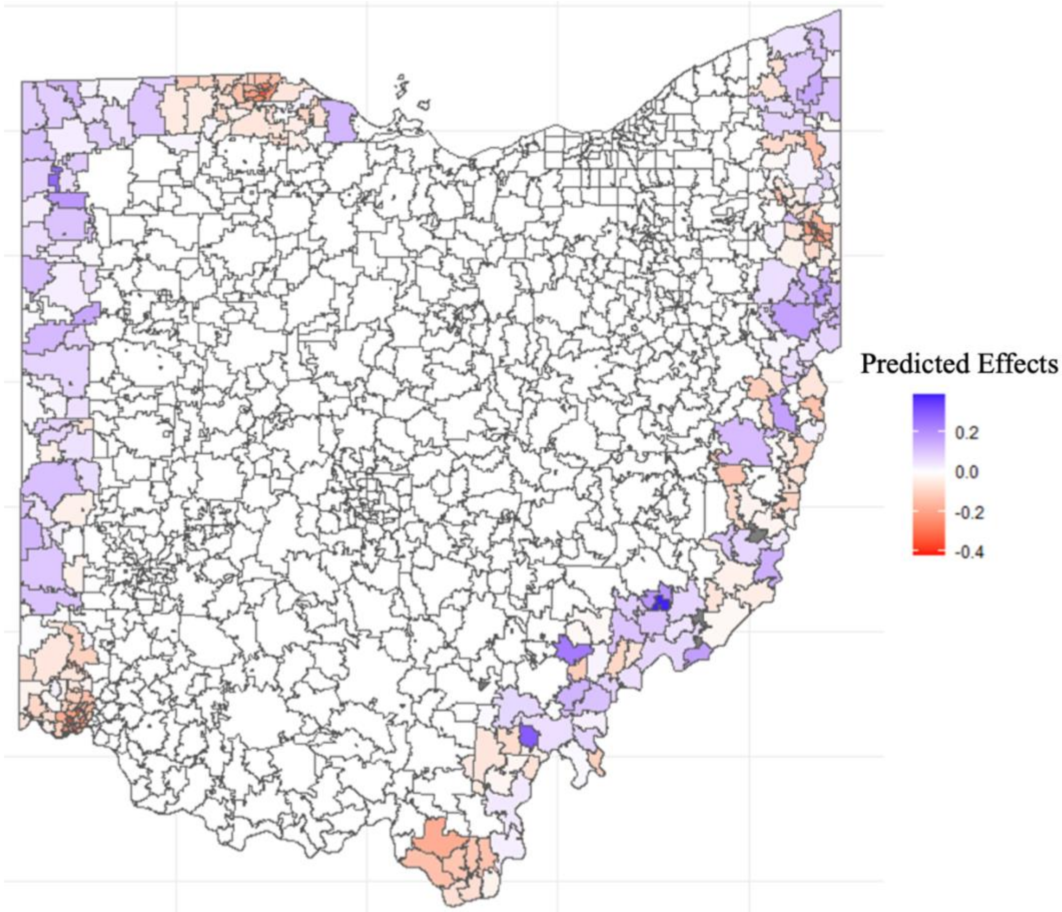
$$log(per\ capita\ wager)_{it} = \beta(online\ 20)_{it} + (online\ 20)_{it} \times (metro)_i + (online\ 20)_{it} \times (prop\ vets)_i + (online\ 20)_{it} \times (prop\ snap)_i + \eta_{it} \quad (9)$$

Table 9: Treatment Interacted with Demographic Variables

$(online\ 20)_{it}$	-0.03547* (0.0211)
$(online\ 20)_{it} \times (metro)_i$	-0.1122*** (0.0118)
$(online\ 20)_{it} \times (prop\ vets)_i$	2.5913*** (0.2529)
$(online\ 20)_{it} \times (prop\ snap)_i$	-1.2387*** (0.1281)
Observations	1,365,074
Adjusted R^2	0.24
Zip FE + Date FE Included	
Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001	

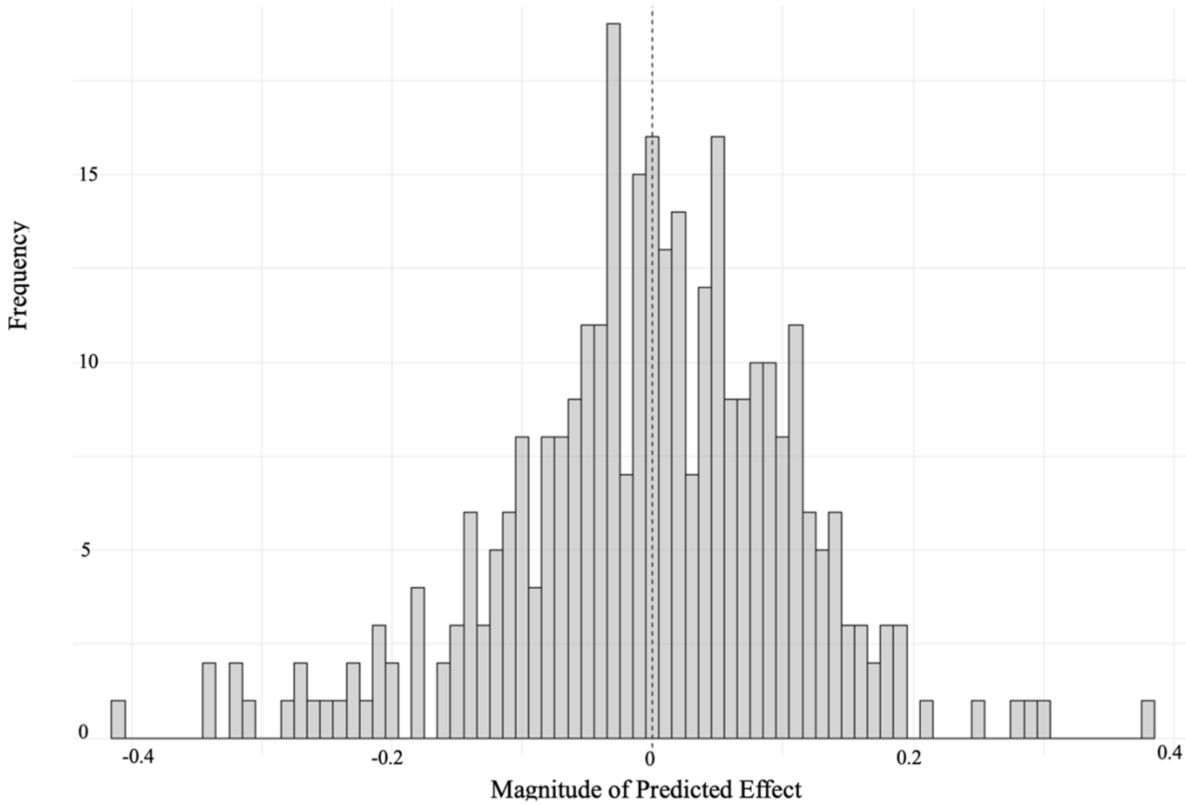
The negative coefficient (significant at the 90% level) on $(online\ 20)_{it}$ indicates a 3.5% decline in online lottery per capita wager in (hypothetical) non-metro ZCTAs with no veterans and no SNAP recipients. Given the interaction terms, actual predicted effects depend on each ZCTA’s demographic characteristics. Higher levels of poverty and urban-ness are estimated to increase the magnitude of the substitution effect, while populations with a high proportion of veterans are estimated to in fact increase their online lottery “consumption” in response to treatment. Hence, some ZCTAs may have demographics that yield a positive treatment effect based on the above specification, which would be indicative of a complementary relationship between online sports betting and online lottery games. To better visualize the insight from the estimates in Table 9, I compute the predicted effect for each ZCTA based on its individual demographic characteristics and project these on a map in Figure 5.

Figure 5: Cross-Sectional Predicted Effects from Homogeneous-by-border 20-mile Treatment



The map highlights that, with the exception of the red in the south, the bulk of the negative (substitution) effects are observed in the relatively metropolitan areas around Toledo (near the Michigan border), Youngstown (near the Pennsylvania border) and Cincinnati (near the Indiana border). The swathes of blue indicate ZCTAs in which the treatment actually led to an increase in online lottery activity, suggesting complementarity with online sports betting. Because non-metro ZCTAs tend to be larger in size than metropolitan, the above map may visually overstate their importance. Figure 6 plots the distribution of effects to highlight a near-zero net effect in line with the estimates reported in Table 6.

Figure 6: Frequency Distribution of Predicted Effects
from Homogeneous-by-border 20-mile Treatment



These results showcase that ZCTAs may differ fundamentally in whether residents who play online lottery games view online sports betting as a substitute or complement to said lottery games. In addition to this cross-sectional analysis that illustrates *where* we might see an effect, I am also interested in the time-component of treatment, i.e., *when* we see an effect relative to the legalization shocks. To this end, I build a time-series version of this model where I interact the temporal variables from Table 8 with the treatment. However, given the interaction terms, predicted effects from such a specification depend on values of the temporal variables on the date t associated with each observation of $(per\ capita\ wager)_{it}$. Note that, because the treatment indicator $(online\ 20)_{it}$ compiles observations that were treated at staggered times into a single treatment group, there exist dates t for which the value of $(online\ 20)_{it}$ may equal 1 or 0 depending on which state i is nearby. Put differently, for a given date t, $(online\ 20)_{it}$ is not necessarily equal for all i, which makes this specification unsuitable for predicting online lottery activity as a function of temporal variables.

To overcome this, I de-bundle the treatment into the individual legalization shocks coming from each of Michigan, Indiana, West Virginia, and Pennsylvania (described in Equation 3). The models that follow hence no longer assume a homogeneous treatment effect in each of the four border regions that receive staggered treatment.

4.2.2 Assuming legalization shocks from different states have heterogeneous treatment effects

I begin with a baseline model for heterogeneous treatment shocks with two-way fixed effects and the only regressors the treatment variables described in Equation 3. Separate treatment variables allow for each shock to have a distinct impact on its respective proximity-area of Ohio ZCTAs. Date fixed effects control for time-variation and ZCTA fixed effects control for underlying heterogeneity in online lottery activity. The estimates from this specification are presented in Table 10.

Table 10: Heterogeneous-by-border 20-mile Treatment Variables

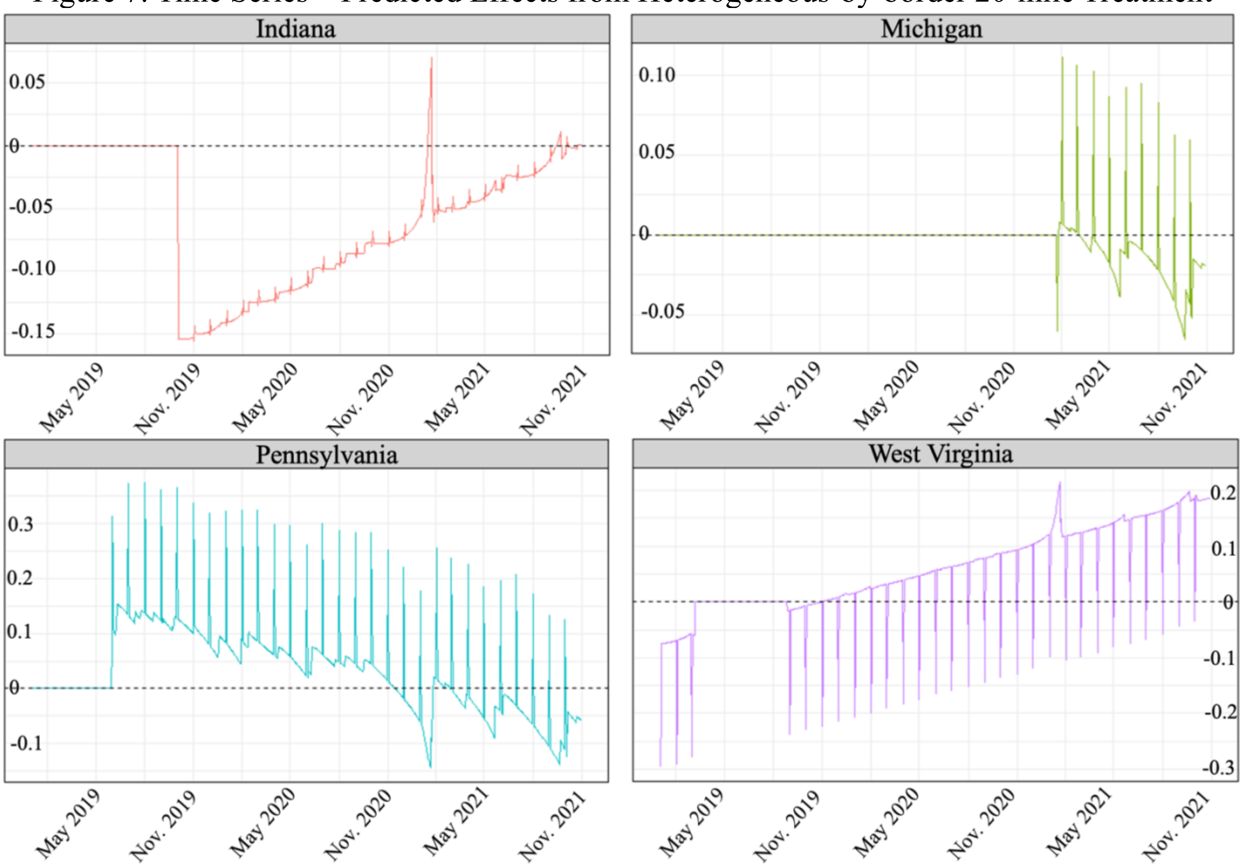
	$(online\ MI\ 20)_{it}$	-0.0218 (0.0132)
	$(online\ IN\ 20)_{it}$	-0.0665*** (0.0099)
	$(online\ WV\ 20)_{it}$	0.0736*** (0.0110)
	$(online\ PA\ 20)_{it}$	0.0328** (0.0124)
Observations		1,365,074
Adjusted R^2		0.25
Zip FE + Date FE Included		
Significance Codes:	'*' 0.05 '**' 0.01	'***' 0.001

Interacting temporal variables with each of these treatment indicators now yields results that can be used to predict the average logged per capita wager on a given date for the separate treatment groups and the control group. Table 11 reports the regression estimates, and Figure 7 plots the associated predicted lottery activity.

Table 11: Interaction of Treatment with Temporal Variables

		$\times (time)_t$	$\times (jackpot)_t$	$\times \binom{square}{jackpot}_t$	$\times \binom{1st\ of}{month}_t$	$\times \binom{2nd\ of}{month}_t$
$\binom{online}{MI\ 20}_{it}$	0.0753 (0.0597)	-0.009 (0.006)	-0.002 (0.009)	-0.0004 (0.0006)	0.1047 (0.0687)	0.0296 (0.0623)
$\binom{online}{IN\ 20}_{it}$	-0.2048*** (0.0200)	0.02*** (0.002)	-0.006 (0.005)	0.0008* (0.00037)	-0.0028 (0.0340)	0.0095 (0.0327)
$\binom{online}{WV\ 20}_{it}$	-0.0741*** (0.0188)	0.03*** (0.002)	-0.003 (0.005)	0.0005 (0.0004)	-0.2223*** (0.0322)	-0.1182 (0.0309)
$\binom{online}{PA\ 20}_{it}$	0.2056*** (0.0205)	-0.02*** (0.002)	-0.01* (0.006)	0.0002 (0.0004)	0.2384*** (0.0355)	0.1091** (0.0342)
Observations	1,365,074					
Adjusted R^2	0.25					
Zip FE + Date FE Included				Significance Codes:		
Jackpot value in 100 millions, time in 100s of days				'*' 0.05	'**' 0.01	'***' 0.001

Figure 7: Time Series – Predicted Effects from Heterogeneous-by-border 20-mile Treatment



With the treatment broken up into separate shocks corresponding to each state's online sports betting legalization, my time-series model starts to paint a clearer picture of substitution from online lottery games in some treated regions and a complementary relationship between the two in others. In the ZCTAs near Indiana, I observe that access to online sports betting is associated with a statistically significant immediate decrease in online lottery activity. However, the coefficient on $(online\ IN\ 20)_{it}$ cannot be interpreted in isolation because it captures the hypothetical effect of the Indiana-shock occurring on the first day of my sample. Online sports betting began in Indiana on October 3, 2019, the 276th day in my sample and the positive coefficient on $(time)_t$ interacted with the treatment implies a net negative impact of -14.96% at the actual time of Indiana's treatment. The positive coefficient on $(online\ IN\ 20)_{it} \times (time)_t$ also indicates that this cannibalization effect wanes with time, which can also be seen in the plot of predicted effects. These results are in line with findings from the literature cited earlier that show a substitution relationship between lotteries and other gambling products. Siegel and Anders (2001) studied the expansion of casino gaming on Native land in Arizona, finding evidence that it cannibalized Arizona's lottery revenues while Cummings (2017) finds similar results for casinos in Maryland and the state's lottery revenues. Cummings' study utilizes a similar proximity framework as mine, using distance to the nearest casino to define treatment. My results show evidence that online lottery games and online sports betting have a similar relationship in the regions of Ohio that are within 20 miles of Indiana.

Pennsylvania's legalization of online sports betting seems to have the opposite effect on online lottery activity in nearby Ohio ZCTAs with a coefficient of similar magnitude as the Indiana shock on the un-interacted treatment variable $(online\ PA)_{it}$. As in the previous example, I consider the coefficient on the $(time)_t$ variable when interacted with the treatment, which is negative in this case. Given online sports betting began in Pennsylvania beginning on the 151st day of my sample (May 31, 2019), I estimate a net effect corresponding to 17.54% higher per capita wagers in ZCTAs within 20 miles of Pennsylvania at the time of the Pennsylvania shock. As cited earlier, Walker and Jackson (2008) find evidence of mixed interaction, and even complementarity, between some forms of gambling. Their work identifies this with parimutuel racing and charity gaming, and my results suggest a similar relationship between online lottery games and online sports betting in the regions of Ohio that are within 20 miles of Pennsylvania. My results suggest a similar relationship between online sports betting and online lottery games in the parts of Ohio in proximity to Pennsylvania. Like the substitution observed near Indiana, the complementarity observed near Pennsylvania as indicated by the negative coefficient on $(online\ PA\ 20)_{it} \times (time)_t$ and evident from Figure 7. However, while the substitution effect estimated near Indiana did not have significant interactions with the pre-existing spikes in online lottery activity, the complementarity between online lottery games and online sports betting near Pennsylvania is even more pronounced on dates that are predisposed to having higher lottery activity. This is indicated by the statistically significant positive coefficients on $(online\ PA\ 20)_{it} \times (1^{st} of\ month)_t$ and $(online\ PA\ 20)_{it} \times (2^{nd}\ of\ month)_t$. It is also apparent from the persistence of monthly spikes in the plot for Pennsylvania which disappear in the plot for Indiana in Figure 7.

The results for West Virginia demand a more cautious interpretation given that it is treated from the start of my sample but experiences a brief period where online sports betting is not operational. Given my regression estimates and the corresponding plot in Figure 7, I interpret my results as evidence of increasing complementarity between online lottery games and online sports betting. While the coefficient on $(online\ WV)_{it}$ is negative, recall that this is the estimated effect on the first day of my sample, January 1, 2019. I am limited by my lack of pre-treatment observations for West Virginia, where online sports betting began in December 2018, but given the positive coefficient on $(online\ WV\ 20)_{it} \times (time)_t$ and as seen in the plot, I estimate net positive impacts near West Virginia soon after their online sports betting resumes in August 2019. My hypothesis here is that exposure to online sports betting in the initial treatment period led people near West Virginia to consider online lottery games as an alternative when they lost the ability to bet on online sports in the neighboring state. This complementarity continues when online sports betting resumed. In contrast to the complementary relationship observed in Pennsylvania, the coefficients on $(online\ WV\ 20)_{it} \times (1^{st}\ of\ month)_t$ and $(online\ WV\ 20)_{it} \times (2^{nd}\ of\ month)_t$ are negative. This is suggestive of the fact that despite the overall complementarity, there is a decrease in the type of intensive online lottery activity observed at the beginning of every month. This is in line with the theory that exposure to online sports betting may have led more people to discover the availability of Ohio Lottery games online, while also providing a desirable alternative to those trying to "maximize" their chances of winning. My argument for a complementary relationship between online sports betting and online lottery games near West Virginia is also supported by evidence from Humphreys (2021), who finds a net positive impact of legalized sports betting on revenue collected from video lottery terminals (VLTs) in West Virginia casinos. Within a casino, sports betting and playing VLT games are cost-identical, whereas access to online sports betting does have a travel-cost compared to online lottery gaming in my empirical design. However, finding a relationship between these online products that is similar to the relationship estimated between their offline analogues in West Virginia lends credibility to my results for parts of Ohio within 20 miles of West Virginia.

My results for Michigan lack statistical significance, which is also true in my next few specifications. My primary explanation for this is that Michigan's treatment shock occurred in 2021, much after the other shocks which had already taken place by October 2019. Given the limitations of my data in terms of the gap in observations from October 2021 to November 2022, the ZCTAs treated by Michigan's legalization shock are observed for a considerably shorter period than their counterparts bordering Indiana, West Virginia or Pennsylvania. Secondly, of Ohio's borders with the four states in question, the shortest one is with Michigan. There are 61 ZCTAs in the Michigan-treatment group $(MI\ 20)_i$, of which 9 are also within 20 miles of the Indiana border, effectively leaving 52 ZCTAs treated by Michigan's legalization of online sports betting. For comparison, there are 73 ZCTAs within 20 miles of the Pennsylvania border, 106 within 20 miles of the West Virginia border and 108 within 20 miles of the Indiana border. As a result, for ZCTAs near Michigan, I lack the power that the data has for estimating the treatment effect in ZCTAs near the other three states.

The wide variation in the effect of offering online sports betting on online lottery activity in different geographic regions motivates the introduction of even further granularity in defining treatment. This is achieved by way of the distance bins defined in Table 4 for which summary statistics were presented in Table 5.

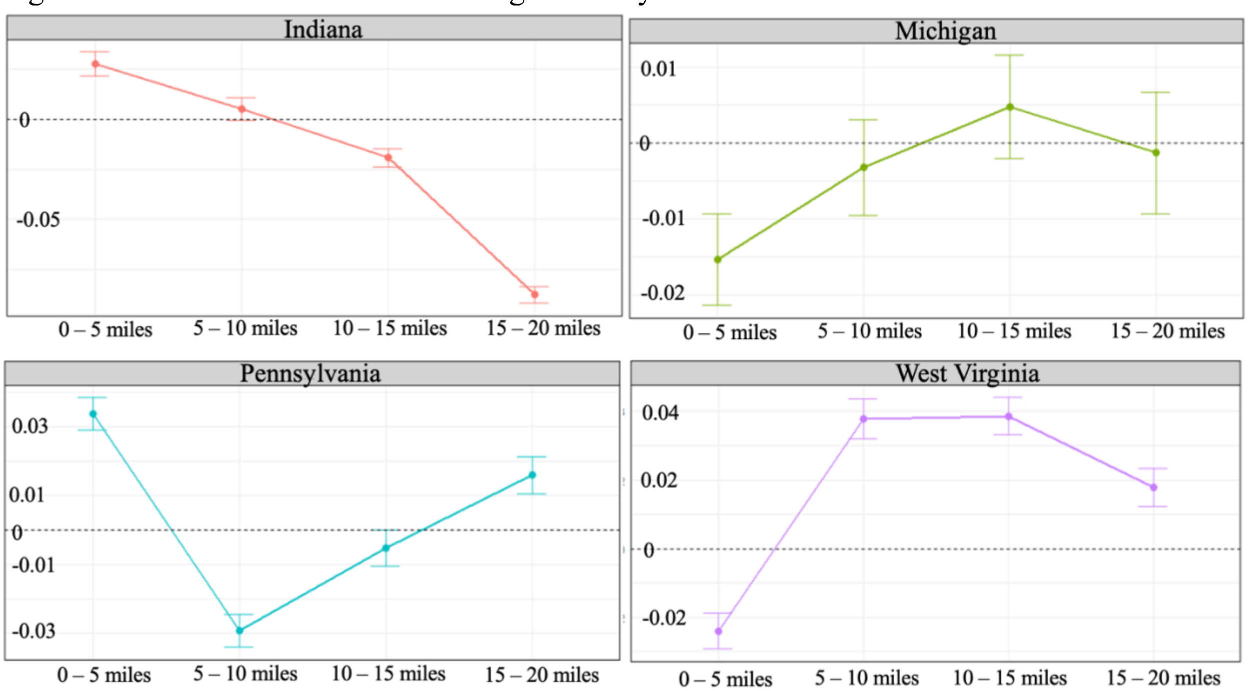
4.2.3 Assuming treatment is heterogeneous by state and distance bins

Estimating heterogeneous-by-distance (bin) impacts of each individual shock generates 16 different treatment variables. My initial specification only includes these treatment variables as regressors in addition to zip and date fixed effects to estimate (logged) per capita wager in ZCTA i on date t. Results are presented in Table 12 and plotted in Figure 8.

Table 12: Heterogeneous-by-border Distance Bins Treatment Variables

	Within 5 mi	5 – 10 mi	10 – 15 mi	15 – 20 mi
Michigan	-0.0154* (0.0060)	-0.0032 (0.0063)	0.0048 (0.0068)	-0.0013 (0.0080)
Indiana	0.0273*** (0.0061)	0.0049 (0.0057)	-0.0193*** (0.0045)	-0.0875*** (0.0042)
West Virginia	-0.0240*** (0.0052)	0.0379*** (0.0057)	0.0386*** (0.0055)	0.0178** (0.0056)
Pennsylvania	0.0420*** (0.0059)	-0.0365*** (0.0060)	-0.0065 (0.0065)	0.0198** (0.0067)
Observations	1,365,074			
Adjusted R^2	0.14			
Zip FE + Date FE Included				
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Figure 8: Predicted Effects from Heterogeneous by Border + Distance Bin Treatment



The plots in Figure 8 project the estimated effect of treatment for ZCTAs in each of the distance bins away from a state's border. I am now able to discern that the substitution near Indiana highlighted by the previous time-series model is driven by ZCTAs between 10 and 20 miles away from the Indiana border. This corresponds to parts of the Cincinnati metro area, which could be a possible driver of online sports betting demand. The Pennsylvania story grows fuzzy since I only notice the complementary relationship discovered by the time-series model in ZCTAs less than 5 miles or between 15 and 20 miles of the border.

The effect in West Virginia, obscured by the negative coefficient in the time-series specification, is clearer in this model. While ZCTAs within 5 miles of the border show evidence of substitution between online sports betting and online lottery games, those between 5 and 20 miles of the border show strong evidence of a complementary relationship between the two. As explained earlier, my data is not as powered to detect an effect in the region of Ohio that borders Michigan, and this is seen again with the confidence intervals being significantly larger than those of the estimates for the other three states.

Aware of the significant underlying variation in the data, I decide to add demographic characteristics to this specification before reading into the heterogeneity therein. To this end I interact demographic variables with the 16 treatment variables from Table 12. Unsurprisingly, this creates a rather crowded regression with coefficients that are only insightful when contextualized with values of the interacted variables. For this reason, I report my estimates from this specification in Table 13 but focus my discussion on the cross-sectional and frequency distribution of predicted effects, presented in Figure 9 and Figure 10 respectively.

Table 13: Heterogeneous Treatment Variables Interacted with Demographic Variables

	Within 5 mi	$5-10$ mi	$10-15$ mi	$15-20$ mi
Michigan	-0.1429*** (0.0326)	-0.2059*** (0.0446)	-0.0287 (0.0270)	0.0106 (0.0218)
Indiana	-0.0669*** (0.0134)	-0.1144*** (0.0122)	0.2852*** (0.0149)	-0.0401*** (0.0095)
West Virginia	0.0349 (0.0178)	-0.1482*** (0.0144)	-0.0291** (0.0099)	0.1691*** (0.0122)
Pennsylvania	0.0709*** (0.0144)	0.0155 (0.0223)	0.0103 (0.0248)	0.0513* (0.0222)
$\times\ (metro)_i$	Within 5 mi	$5-10$ mi	$10-15$ mi	$15-20$ mi
Michigan	0.0204 (0.0143)	0.0079 (0.0153)	-0.0625*** (0.0154)	-0.0246 (0.0149)
Indiana	-0.0021 (0.0095)	-0.0205** (0.0095)	-0.0986*** (0.0061)	-0.0230*** (0.0061)
West Virginia	0.0630*** (0.0063)	0.0956*** (0.0071)	-0.0181** (0.0065)	0.0275*** (0.0072)
Pennsylvania	0.0333*** (0.0069)	-0.0152 (0.0095)	0.0100 (0.0087)	0.0067 (0.0093)
$\times\ (prop\ vets)_i$	Within 5 mi	$5-10$ mi	$10-15$ mi	$15-20$ mi
Michigan	1.9792*** (0.5183)	3.1444*** (0.0526)	1.2607** (0.4300)	-0.2361 (0.3198)
Indiana	0.6450** (0.2246)	3.1600*** (0.1736)	-2.8662*** (0.1865)	0.9079*** (0.1498)
West Virginia	-1.5214*** (0.1506)	0.3610* (0.1192)	0.0121 (0.1064)	-1.736*** (0.1218)
Pennsylvania	-0.8249*** (0.1615)	0.6012** (0.2174	0.2106 (0.2341)	0.1633 (0.2842)
$\times\ (prop\ snap)_i$	Within 5 mi	$5-10$ mi	$10-15$ mi	$15-20$ mi
Michigan	-0.2017 (0.1076)	0.1689 (0.1311)	0.4761 (0.5751)	0.0081 (0.5923)
Indiana	0.1209 (0.1884)	-2.023*** (0.1886)	-1.0958*** (0.0701)	-0.3176*** (0.0550)
West Virginia	0.1203 (0.1134)	0.6896*** (0.1289)	-0.2101** (0.0744)	-1.4352*** (0.1070)
Pennsylvania	-0.1505 (0.1019)	-0.6533*** (0.0562)	-0.3589* (0.1416)	-0.7926*** (0.1093)
Observations	1,365,074			
Adjusted R^2	0.04			

Zip FE + Date FE Included

Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001

Figure 9: Cross-Sectional Predicted Effects from Heterogeneous-by-border Distance Bins Treatment

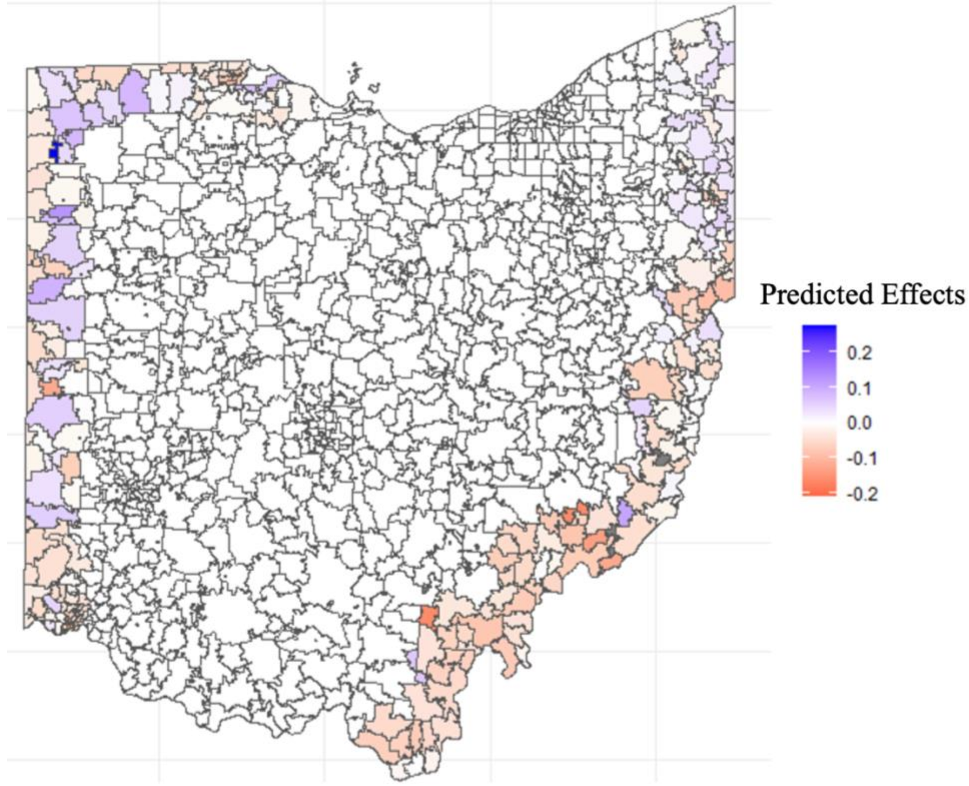
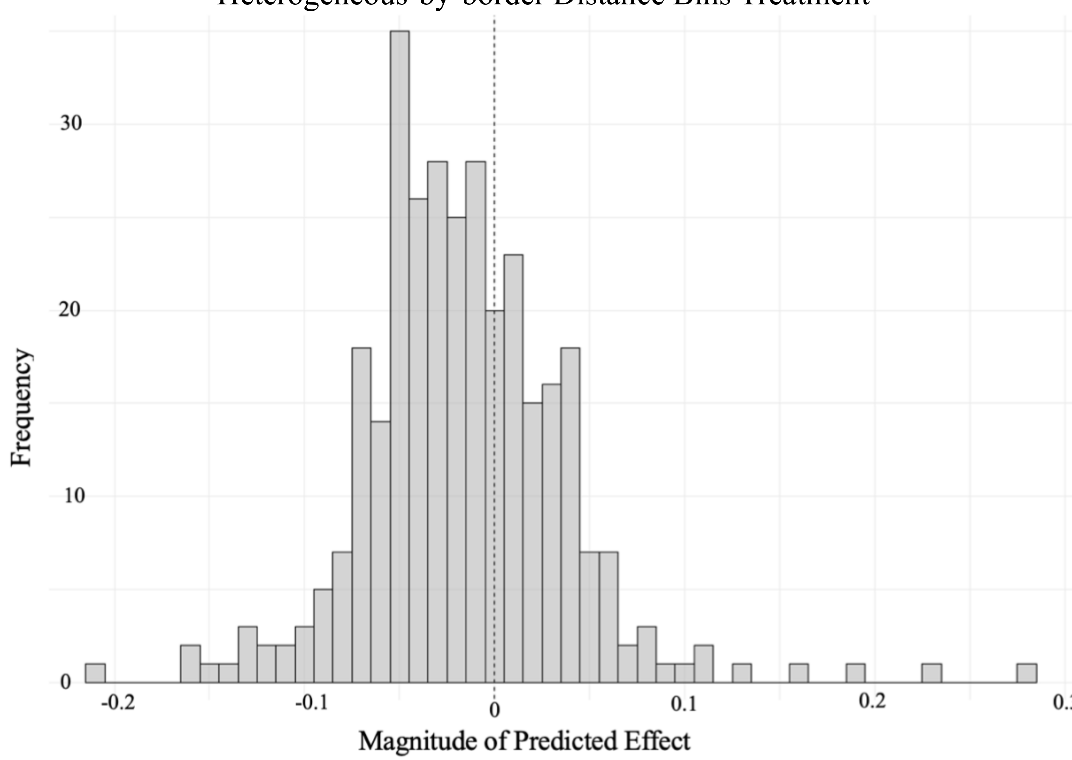


Figure 10: Frequency Distribution of Predicted Effects from Heterogeneous-by-border Distance Bins Treatment



As evident from the coefficients in Table 13, the ZCTAs in each of the distance bins vary widely, not just in underlying demographics as previously established in Table 5A, but even in how these factors influence online lottery activity. As in an earlier specification that interacted these demographic variables with a homogeneous 20-mile treatment, the first panel of un-interacted treatment terms in Table 13 denotes the "intercept" of predicted effects in each of the 16 groups. Given the interaction terms, actual predicted effects depend on each ZCTA's demographic characteristics.

Because my hypothesis rests on the ability to legally access online sports betting, I would hope to see predicted effects wane as distance to the pertinent border increases. However, Figure 9 shows the lack of such a pattern, even when accounting for zip fixed effects and demographic-driven heterogeneity in treatment effects. This is because the demographics explain some, but not all, of the underlying heterogeneity and the residual (unexplained) variation then obscures any distance-effect I may have been able to discern. Given my data and model, my findings therefore suggest that that underlying idiosyncratic preferences outweigh proximity to online sports betting in determining online lottery activity in Ohio.

The clearest evidence for this apparent from my map comes from the predicted effects near the Indiana border. I observe a mix of substitution and complementary behavior along the border, varying based on the relative levels of poverty and veteran population in each ZCTA. Though underrepresented on the map due to their geographic size, the metropolitan ZCTAs near Cincinnati drive a significant chunk of this substitution, while the rural, low-poverty ZCTAs along the Indiana border exhibit complementarity between online lottery games and online sports betting.

To visualize predicted effects from this specification sans distortion caused by the variation in geographic size, I plot a histogram of predicted effects as I did for the earlier cross-sectional model. While that one showed ambiguous effects (refer Figure 6), Figure 10 suggests it may be a net substitution effect that prevails overall. This is driven in part by the region near West Virginia which previously showed evidence for complementarity between online lottery games and online sports betting now displaying significant evidence for substitution between the two.

5. Policy Implications

I consider the impact of sports betting legalization through the lens of a government that offers a portfolio of gambling products and seeks to maximize the revenue it generates from consumption of said products. Recent estimates[17] show that for every dollar wagered on an Ohio Lottery game, 24 cents make their way to the Lottery Profit Education Fund (LPEF). Sportsbooks are required to pay a combination of license fees in order to operate in different forms (online, casino, kiosk, etc.) and are taxed on their gross profits, i.e., wagers collected minus winnings paid. For simplicity and comparability, I consider the licensing fees a one-time fixed cost of doing business and focus my analysis on comparative statics associated with the tax rate on sports betting. The government chooses the optimal rate τ_S to maximize total "gambling

[17] https://fox8.com/news/does-my-lottery-ticket-purchase-help-ohio-schools/

revenue" as shown in Equation 10, where W_L and W_s represent the amount wagered on lottery games and sports betting respectively. Profits from the lottery go directly to state revenue with π_L roughly equal to 0.24, and sportsbooks' gross profits are given by $\pi_S = \frac{W_S - P_S}{W_S}$ where P_S captures the winnings paid out by sportsbooks.

$$\max_{\tau_S} \; \Pi = \pi_L W_L + \tau_S(\pi_S W_S) \tag{10}$$

The tax rate on sports betting was initially 10% in Ohio but was raised to 20% just seven months later in July 2023. The country's first ever increase in sports betting tax rate was motivated by what Ohio Governor Mike DeWine considered reckless advertising that "crossed the line" in trying to lure in bettors with excessive promotions and "free" bets. The higher tax is intended to dampen advertising and promotional spending which is presumably positively correlated with W_S. Although bettors are not directly impacted by the government's policy because sportsbooks are taxed on gross profits, sports betting activity as captured by W_S is negatively correlated with τ_S.

My results point to mixed interactions between online lottery activity and online sports betting availability in different regions. While my model may not fully identify all the drivers of these effects, it does provide robust evidence that sports betting legalization in Ohio is likely to have a non-uniform impact on online lottery activity across the state. The government's choice of sports betting tax rate varies based on the relationship between demand for online lottery games and online sports betting. In particular, given that online sports betting is "available" across Ohio since January 2023, I loosely proxy for the degree of availability using the extent of advertising and promotional spending by sportsbooks. I denote this as A_S and assume $\frac{\partial W_S}{\partial A_S} > 0$. In regions that show evidence for substitution (complementarity) between online lottery games and online sports betting, I therefore expect $\frac{\partial W_L}{\partial A_S} < 0 \left(\frac{\partial W_L}{\partial A_S} > 0\right)$. Assuming the government's policy has the intended effect, i.e. $\frac{\partial A_S}{\partial \tau_S} < 0$, I present a simplified overview of policy optimization given different types of interaction between demand for online lottery games and online sports betting.

For the baseline scenario, assume the two products are completely independent: $\frac{\partial W_L}{\partial A_S} = 0$. Because lottery sales W_S are unaffected by the availability of sports betting and the degree to which it is advertised, the government's optimization problem reduces to maximizing the second term in Equation 10, bearing in mind that $\frac{\partial W_S}{\partial \tau_S} < 0$. Let τ_0 denote the state's revenue-maximizing tax rate on sports betting in this scenario. If there is a complementary relationship between lottery games and sports betting as some of my results suggest, one would expect $\frac{\partial W_L}{\partial A_S} > 0$. In this case, the government would not want to inhibit sportsbooks' promotional spending as much because it drives demand for sports betting, which in turn induces increased lottery consumption that contributes to government revenues. Hence $\tau_{Comp} < \tau_0$. On the other hand, should online sports betting and online lottery games be net substitutes, one would expect $\frac{\partial W_L}{\partial A_S} <$

0. In this case, the government may tax sportsbooks more aggressively to strongly discourage promotional spending so $\tau_0 < \tau_{Sub}$.

If online lottery games and online sports betting are complements, the government faces no trade-off in terms of losing lottery revenue at the cost of earning revenue from taxes on sports betting. On the other hand, if people substitute away from online lottery games in favor of online sports betting, a state with the option of banning sports betting must *at least* recoup the lost lottery profits at the optimal τ_S. Recall that lottery profits directly constitute government revenue but revenue from sports betting comes from taxing sportsbooks' profits. It would require an unrealistically high tax rate to equalize the government's "take" per unit wager from each activity. According to data from the Ohio Casino Control Commission, over \$7.65 billion was wagered on sports bets in Ohio in 2023, of which the taxable amount (gross profits) was \$937 million. In terms of the notation used in Equation 10, this equates to $\pi_S \approx 12.25\%$. This means a dollar wagered on sports betting in Ohio yields the government just 2.4 cents, compared to 24 cents per dollar wagered on the Ohio Lottery. The required tax rate to equalize these yields would be 200%, which would drive any sports betting company to exit. A more realistic perspective is that with a 20% tax on sportsbooks, the government nets as much revenue from \$1 wagered on lottery games as it does from \$10 wagered in sports bets. So, substitution from lottery games to sports betting does not harm state revenue so long as the substituting actors spend 10 times the amount sports betting that they would have on lottery games. My field research found that all the online sports betting providers in Ohio require a minimum of \$5 to be deposited in the betting account, which is also the minimum wager on most of the odds offered. This is five times the dollar-minimum on online lottery games, so substituting actors who purchase the "minimum" wager do spend at least five times what they would have on a lottery game on sports betting instead.

A high volume of sports betting, while favorable to state coffers, may be in conflict with the government's desired social outcomes given the risk of addiction associated with gambling. A sizable portion of the sports audience, and hence sports betting, includes younger members of the population, a segment that is particularly vulnerable to developing addictive behaviors. Room et al. (1999) finds that the opening of a casino in Niagara led to an increase in reported gambling problems among locals, presenting a clean example of how a popular gambling product may pose society-wide addiction concerns. Furthermore, given my mixed results, sports betting legalization may also cause an increase in lottery demand in some markets. With the associated increase in government revenue also comes the increase in negative social outcomes from lottery spending largely funded by low incomes. Lockwood et al. (2024) describe how lotteries act as a regressive tax by finding that measures of behavioral biases (e.g., financial illiteracy, statistical mistakes) that are associated with lower income and education are strongly associated with higher lottery spending.

6. Conclusion

The lottery is an established source of revenue across several states, Ohio included. Balancing the social cost of gambling, state governments seek to maximize this revenue. To this end, the Ohio Lottery has, over the course of its existence, introduced new games, modified odds for existing games, and most importantly for my study, offered online lottery gaming since December 2012. Legalizing sports betting in January 2023 created an additional source of revenue for the government, but, as my results show, there is significant interaction in the demand for online sports betting and online lottery games. Understanding these interactions is crucial in optimizing regulation across the portfolio of gambling products now offered in Ohio to maximize government revenue.

I model the impact of online sports betting being legalized in neighboring states using a variety of specifications. My initial results highlight the significance of certain demographic and temporal variables in explaining cross-sectional heterogeneity and variation over time respectively. I leverage this to compute predicted effects based on interacting the treatment with each, the demographic variables and the temporal variables. I consider alternate definitions of treatment, beginning with a homogeneous treatment applied 20 miles from each of the four pertinent borders, then heterogenous treatments from each state's shock applied 20 miles from each border, and finally heterogeneous treatments from each shock that are allowed to vary for distance bins in increments of 5 miles up to 20 miles from each border.

I have a considerable quantity of high-quality data, and many of my results are highly significant – but they point to an ambiguous aggregate relationship between online lottery games and online sports betting. I find strong evidence of a negative relationship between the two products in the metropolitan areas around Toledo, Cincinnati, and Youngstown. Some of my results suggest there might be a complementary relationship between online lottery games and online sports betting in the parts of Ohio that border West Virginia, but these dissipate in the specification with the most granular treatment definition. The border region with Indiana shows mixed interactions with significant evidence for complementarity in rural ZCTAs with low levels of poverty and a high veteran proportion of population.

The variation in predicted effects across ZCTAs makes it difficult to pin down an overall net effect, but any cannibalization is expensive due to the fundamental difference in how the state earns revenue from the lottery versus from sports betting. Lottery profits directly go towards government revenue, whereas sports betting is taxed based on the gross profits of a sportsbook. As a result, the government gets a much smaller fraction of the wager from sports betting than it does from the lottery. To recover cannibalized lottery profits therefore requires a disproportionately high amount of sports betting, which poses the risk of incurring social costs associated with gambling addiction. The government must balance the optimal tax rate so as to maintain a level of sports betting activity that provides the basis for tax revenue while limiting consumption to a socially optimal level.

References

1. Abarbanel, B. L., Bernhard, B. J., & Singh, A. K. (2011). Daily revenue patterns and the impact of race and sports book wagering at Las Vegas casinos. UNLV Gaming Research & Review Journal, 15(2), 1-13.
2. American Gaming Association (AGA). (2023). "State of the states: The AGA survey of the casino industry."
3. Clotfelter, C. T., & Cook, P. J. (1987). Implicit taxation in lottery finance. National Tax Journal, 40(4), 533-546.
4. Clotfelter, C. T., & Cook, P. J. (1989). Selling hope: State lotteries in America. Harvard University Press.
5. Clotfelter, C. T., Cook, P. J., Edell, J. A., & Moore, M. (1999). State lotteries at the turn of the century: Report to the National Gambling Impact Study Commission. Duke University.
6. Cummings, W. L. (2017). The cannibalization effect of casinos on lottery revenue: A spatial analysis of Maryland. Economics Bulletin, 37(2), 950-960.
7. DeBoer, L. (1986). Lotto sales stagnation: Product maturity or small jackpots? Growth and Change, 17(3), 33-41.
8. Farrell, L., & Forrest, D. (2008). The demand for lotto: The role of prize structure. Journal of the Royal Statistical Society: Series A (Statistics in Society), 171(1), 151-171.
9. Farrell, L., & Walker, I. (1999). The welfare effects of lotto: Evidence from the UK. Journal of Public Economics, 72(1), 99-120.
10. Farrell, L., Morgenroth, E., & Walker, I. (1999). A time series analysis of UK lottery sales: Long and short run price elasticities. Oxford Bulletin of Economics and Statistics, 61(4), 513-526.
11. Foden-Vencil, K. (2022, August 9). New research shows lotteries rely on low-income players. OPB. Retrieved from https://www.opb.org/article/2022/08/09/new-research-shows-lotteries-rely-on-low-income-players/
12. Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risk. Journal of Political Economy, 56(4), 279-304.
13. Grote, K. R., & Matheson, V. A. (2011). A survey of the economics of lottery games. Handbook on the Economics of Gambling, Edward Elgar Publishing.
14. Gulley, O. D., & Scott, F. A. (1993). The demand for wagering on state-operated lotto games. National Tax Journal, 46(1), 13-22.
15. Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Subjective relative income and lottery ticket purchases. Journal of Behavioral Decision Making, 21(3), 283-295.
16. Hansen, A. (1995). The tax incidence of the Colorado state instant lottery game. Public Finance Quarterly, 23(3), 385-398.
17. Hansen, A., Miyazaki, A. D., & Sprott, D. E. (2000). The tax incidence of lotteries: Evidence from five states. Journal of Consumer Affairs, 34(2), 182-203.
18. Lang, K. B., & Omori, M. (2009). Can demographic variables predict lottery and pari-mutuel losses? An empirical investigation. Journal of Gambling Studies, 25(2), 171-183.
19. Lee, Y. H., Lin, Y. C., & Lai, C. Y. (2010). Demand for lottery: The case of Taiwan. Applied Economics, 42(6), 717-725.
20. Lockwood, Brian. (2024). The relationship between state lotteries and sports betting. Journal of Public Economics.

21. Lucas, A. F. (2014). Estimating the effect of sports book revenues on slot machine performance at Las Vegas casinos. UNLV Gaming Research & Review Journal, 18(1), 1-18.
22. Markowitz, H. (1952). The utility of wealth. Journal of Political Economy, 60(2), 151-158.
23. Mikesell, J. L. (1987). The effect of aggregate economic variables on state lottery revenue. Policy Studies Journal, 15(3), 423-431.
24. Miller, J. D., & Morey, E. R. (2003). The demand for lottery products and their distributional consequences. Review of Economics and Statistics, 85(4), 1249-1259.
25. National Conference of State Legislatures (NCSL). (2021). "State sports betting policies."
26. Ohio Legislative Service Commission (2021). "Overview of gambling legislation in Ohio." Ohio Legislature. Retrieved from https://www.legislature.ohio.gov.
27. Ohio Lottery Commission (2020). "Annual Report 2020." Ohio Lottery Commission. Retrieved from https://www.ohiolottery.com.
28. Papachristou, G., & Karamais, C. (1998). Price and income elasticity of demand for lottery tickets in Greece. Applied Economics Letters, 5(11), 733-735.
29. Rubenstein, R., & Scafidi, B. (2002). Who pays and who benefits? Examining the distributional consequences of the Georgia Lottery for Education. National Tax Journal, 55(2), 223-238.
30. Siegel, D., & Anders, G. C. (2001). The impact of Indian casinos on state lotteries: A case study of Arizona. Public Finance Review, 29(2), 139-147.
31. Smith, J. (2022, June 5). "Ohio sports betting revenue projections." Cleveland Plain Dealer. Retrieved from https://www.cleveland.com/sports-betting.
32. Suh, E. E., & Tsai, M. C. (2013). The relationship between poker gaming and slot machine revenues at Las Vegas casinos. Journal of Gambling Business and Economics, 7(1), 1-15.
33. Thompson, D. (2023, March 14). "The economic impact of sports betting in Ohio." Cincinnati Business Courier. Retrieved from https://www.bizjournals.com/cincinnati.
34. U.S. Census Bureau (2019). "American Community Survey 2019 Data." U.S. Government Printing Office.
35. Vasche, J. D. (1985). The incidence of state lotteries: A reexamination. National Tax Journal, 38(3), 389-399.
36. Vrooman, D. (1976). The economic impact of legalized gambling. Journal of Legal Studies, 5(2), 187-210.
37. Walker, D. M., & Jackson, J. D. (2008). Do U.S. gambling industries cannibalize each other? Public Finance Review, 36(3), 308-333.
38. Welte, J. W., Barnes, G. M., Wieczorek, W. F., Tidwell, M. C., & Parker, J. C. (2002). Gambling participation in the U.S.—Results from a national survey. Journal of Gambling Studies, 18(4), 313-337.

Appendix

A. *Regression results using alternate outcome variables (measure of online lottery activity)*

Table A1: Baseline Specification with Outcome Variable Logged Per Capita Users

Intercept	-6.4005*** (0.0012)	---	---	---
$(online\ 20)_{it}$	0.0565*** (0.0030)	0.0820*** (0.0030)	-0.0567*** (0.0040)	0.0003 (0.0043)
Zip Fixed Effects	N	N	Y	Y
Date Fixed Effects	N	Y	N	Y
Observations				1,365,074
Adjusted R^2	0.0003	0.0638	0.2904	0.353
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Table A2: Baseline Specification with Outcome Variable Logged Per Capita Transactions

Intercept	4.6053*** (0.0016)	---	---	---
$(online\ 20)_{it}$	0.0949*** (0.0038)	0.0878*** (0.0038)	0.0337*** (0.0054)	-0.0025 (0.0058)
Zip Fixed Effects	N	N	Y	Y
Date Fixed Effects	N	Y	N	Y
Observations				1,365,074
Adjusted R^2	0.0005	0.0353	0.2204	0.2568
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Table A3: Impact of Treatment Interacted with Demographic Variables on Measures of Online Lottery Activity

	Per capita users	Per capita transactions
$(online\ 20)_{it}$	-0.0135	-0.0327
	(0.0142)	(0.0194)
$(online\ 20)_{it} \times (metro)_i$	-0.1536***	-0.0973***
	(0.0080)	(0.0109)
$(online\ 20)_{it} \times (prop\ vets)_i$	3.1051***	2.1492***
	(0.1706)	(0.2327)
$(online\ 20)_{it} \times (prop\ snap)_i$	-1.5983***	-0.8446***
	(0.0864)	(0.1179)
Observations	1,365,074	1,365,074
Adjusted R^2	0.35	0.26
Zip FE + Date FE Included		
Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001		

Table A4: Impact of Heterogeneous-by-state 20-mile Treatment on Measures of Online Lottery Activity

	Per capita users	Per capita transactions
$(online\ MI\ 20)_{it}$	0.0044	-0.0082
	(0.0089)	(0.0122)
$(online\ IN\ 20)_{it}$	-0.0933***	-0.0487***
	(0.0067)	(0.0091)
$(online\ WV\ 20)_{it}$	0.1149***	0.0739***
	(0.0074)	(0.0101)
$(online\ PA\ 20)_{it}$	0.0289***	0.0204
	(0.0084)	(0.0114)
Observations	1,365,074	1,365,074
Adjusted R^2	0.35	0.26
Zip FE + Date FE Included		
Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001		

Table A5: Impact of Heterogeneous-by-state 20-mile Treatment
Interacted with Temporal Variables on Logged Per Capita Users

		$\times (time)_t$	$\times (jackpot)_t$	$\times \binom{square}{jackpot}_t$	$\times \binom{1st\ of}{month}_t$	$\times \binom{2nd\ of}{month}_t$
$\binom{online}{MI\ 20}_{it}$	-0.0197 (0.0403)	0.002 (0.004)	-0.002 (0.006)	-0.0003 (0.0004)	-0.0033 (0.0464)	0.0198 (0.0421)
$\binom{online}{IN\ 20}_{it}$	-0.2919*** (0.0135)	0.03*** (0.002)	-0.004 (0.004)	0.0005 (0.0003)	0.0168 (0.0229)	0.0269 (0.0221)
$\binom{online}{WV\ 20}_{it}$	0.0817*** (0.0127)	0.01*** (0.001)	-0.003 (0.004)	0.0003 (0.0003)	-0.0058 (0.0218)	0.0024 (0.0209)
$\binom{online}{PA\ 20}_{it}$	0.1087*** (0.0138)	-0.01*** (0.001)	-0.006 (0.004)	-0.000004 (0.0003)	0.0402 (0.0239)	-0.0068 (0.0231)
Observations	1,365,074					
Adjusted R^2	0.35					
Zip FE + Date FE Included				Significance Codes:		
Jackpot value in 100 millions, time in 100s of days				'*' 0.05	'**' 0.01	'***' 0.001

Table A6: Impact of Heterogeneous-by-state 20-mile Treatment
Interacted with Temporal Variables on Logged Per Capita Transactions

		$\times (time)_t$	$\times (jackpot)_t$	$\times \binom{square}{jackpot}_t$	$\times \binom{1st\ of}{month}_t$	$\times \binom{2nd\ of}{month}_t$
$\binom{online}{MI\ 20}_{it}$		-0.005 (0.01)	-0.00 (0.0001)	-0.0000 (0.0000)	-0.0071 (0.0632)	0.0241 (0.0574)
$\binom{online}{IN\ 20}_{it}$		0.02*** (0.00)	-0.01 (0.0000)	0.0000* (0.0000)	0.0448 (0.0313)	0.0514. (0.0311)
$\binom{online}{WV\ 20}_{it}$		0.02*** (0.00)	-0.0000 (0.0000)	0.0000 (0.0000)	-0.1901*** (0.0297)	-0.1206*** (0.0284)
$\binom{online}{PA\ 20}_{it}$		-0.02*** (0.00)	-0.01* (0.0000)	0.0000 (0.0000)	0.2113*** (0.0326)	0.1176*** (0.0315)
Observations	1,365,074					
Adjusted R^2	0.26					
Zip FE + Date FE Included				Significance Codes:		
Jackpot value in 100 millions, time in 100s of days				'*' 0.05	'**' 0.01	'***' 0.001

Table A7: Impact of Heterogeneous-by-state
Distance Bins Treatment on Logged Per Capita Users

	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	-0.0003 (0.0001)	0.0002 (0.0001)	0.0002 (0.0001)	0.0002 (0.0001)
Indiana	0.0009 (0.0001)	0.0040 (0.0001)	-0.0014 (0.0001)	-0.0045 (0.0001)
West Virginia	0.0013 (0.0001)	0.0015 (0.0001)	0.0014 (0.0001)	0.0014 (0.0001)
Pennsylvania	0.0012 (0.0001)	-0.0028 (0.0001)	0.0010 (0.0001)	0.0009 (0.0001)
Observations	1,365,074			
Adjusted R^2	0.33			
Zip FE + Date FE Included				
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Table A8: Impact of Heterogeneous-by-state
Distance Bins Treatment on Logged Per Capita Transactions

	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	-0.0015 (0.0012)	0.0013 (0.0012)	0.0013 (0.0013)	-0.0004 (0.0015)
Indiana	0.0040*** (0.0012)	-0.0004 (0.0011)	-0.0031*** (0.0009)	-0.0158*** (0.0008)
West Virginia	0.0044*** (0.0010)	0.0082*** (0.0011)	0.0065*** (0.0010)	0.0075*** (0.0011)
Pennsylvania	0.0057*** (0.0011)	-0.0088*** (0.0011)	0.0050*** (0.0013)	0.0028* (0.0013)
Observations	1,365,074			
Adjusted R^2	0.15			
Zip FE + Date FE Included				
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Table A9: Impact of Heterogeneous-by-state Distance Bins Treatment Interacted with Demographic Variables on Logged Per Capita Users

	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	-0.0061*** (0.0004)	-0.0062*** (0.0006)	-0.0015*** (0.0003)	-0.0002 (0.0003)
Indiana	-0.0026*** (0.0002)	-0.0045*** (0.0002)	0.0030*** (0.0002)	-0.0002 (0.0001)
West Virginia	0.0002 (0.0002)	-0.0020*** (0.0002)	-0.0008*** (0.0002)	0.0051*** (0.0002)
Pennsylvania	0.0005** (0.0002)	-0.0013*** (0.0003)	-0.0008** (0.0003)	-0.0002 (0.0003)
$\times\ (metro)_i$	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	0.0022*** (0.0002)	-0.0001 (0.0002)	-0.0002 (0.0002)	0.0001 (0.0002)
Indiana	-0.0004*** (0.0001)	-0.0007*** (0.0001)	-0.0017*** (0.0001)	-0.0003*** (0.0001)
West Virginia	-0.0002 (0.0001)	0.0010*** (0.0001)	-0.0003** (0.0001)	0.0008*** (0.0001)
Pennsylvania	0.0005*** (0.0001)	-0.0004** (0.0001)	0.0003** (0.0001)	0.00002 (0.0001)
$\times\ (prop\ vets)_i$	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	0.0955*** (0.0067)	0.1018*** (0.0068)	0.0032*** (0.0056)	0.0069 (0.0041)
Indiana	0.0387*** (0.0030)	0.1247*** (0.0022)	-0.0140*** (0.0024)	0.0100*** (0.0019)
West Virginia	-0.0124*** (0.0019)	0.0133*** (0.0015)	0.0074*** (0.0014)	-0.0545*** (0.0016)
Pennsylvania	0.0017 (0.0021)	0.0306*** (0.0028)	0.0191*** (0.0030)	0.0113** (0.0037)
$\times\ (prop\ snap)_i$	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	-0.0103*** (0.0014)	-0.0028. (0.0017)	0.0065 (0.0074)	-0.0124 (0.0074)
Indiana	-0.0034 (0.0024)	-0.0761*** (0.0024)	-0.0173*** (0.0009)	-0.0076*** (0.0007)
West Virginia	-0.0053*** (0.0015)	0.0060*** (0.0017)	-0.0093*** (0.0010)	-0.0268*** (0.0014)
Pennsylvania	-0.0125*** (0.0013)	-0.0131*** (0.0007)	-0.0109*** (0.0018)	-0.0135*** (0.0014)
Observations	1,365,074			
Adjusted R^2	0.12			

Zip FE + Date FE Included

Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001

Table A10: Impact of Heterogeneous-by-state Distance Bins Treatment Interacted with Demographic Variables on Logged Per Capita Transactions

	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	-0.0439*** (0.0063)	-0.0384*** (0.0086)	0.0023 (0.0052)	0.0008 (0.0048)
Indiana	-0.0128*** (0.0026)	-0.0245*** (0.0024)	0.0649*** (0.0029)	-0.0063*** (0.0018)
West Virginia	-0.0002 (0.0018)	-0.0284*** (0.0034)	-0.0082*** (0.0019)	0.0496*** (0.0024)
Pennsylvania	0.0121*** (0.0028)	0.0003 (0.0043)	-0.0060 (0.0048)	0.0037 (0.0043)
$\times (metro)_i$	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	0.0127*** (0.0028)	0.0023 (0.0030)	-0.01112*** (0.0030)	-0.0027 (0.0035)
Indiana	-0.0018 (0.0018)	-0.0032* (0.0015)	-0.0180*** (0.0012)	-0.006*** (0.0018)
West Virginia	0.0038** (0.0012)	0.0174*** (0.0014)	-0.0017 (0.0013)	0.0083*** (0.0014)
Pennsylvania	0.0068*** (0.0013)	0.0010 (0.0018)	0.0070*** (0.0017)	0.0037*** (0.0018)
$\times (prop\ vets)_i$	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	0.6055*** (0.1003)	0.6092*** (0.1018)	0.2266** (0.0832)	0.0127 (0.0619)
Indiana	0.1343** (0.0435)	0.6368*** (0.0336)	-0.6710*** (0.0361)	0.1910*** (0.0290)
West Virginia	-0.1992*** (0.0291)	0.1152*** (0.0231)	0.0481* (0.0206)	-0.4277*** (0.0236)
Pennsylvania	-0.1262*** (0.0313)	0.1169** (0.0421)	0.1108* (0.0453)	0.0038 (0.0550)
$\times (prop\ snap)_i$	Within 5 mi	5 − 10 mi	10 − 15 mi	15 − 20 mi
Michigan	-0.0330 (0.0208)	-0.0191 (0.0254)	-0.2376* (0.1132)	-0.1236 (0.1147)
Indiana	-0.0488 (0.0365)	-0.3917*** (0.0365)	-0.2457*** (0.0136)	-0.0551*** (0.0106)
West Virginia	0.0454* (0.0220)	0.1451*** (0.0250)	-0.0490*** (0.0207)	-0.3955*** (0.0207)
Pennsylvania	-0.0234 (0.0197)	-0.1156*** (0.0109)	-0.0639* (0.0274)	-0.0937*** (0.0212)
Observations	1,365,074			
Adjusted R^2	0.03			
Zip FE + Date FE Included				
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

B. Regression results using alternate distance thresholds in defining treatment and logged per capita wager as the outcome variable

Table A11: Baseline Specification with 15-mile Treatment

Intercept	-2.7272*** (0.0013)	---	---	---
$(online\ 15)_{it}$	-0.0503*** (0.0036)	-0.1841*** (0.0036)	0.2331*** (0.0053)	-0.0981*** (0.0054)
Zip Fixed Effects	N	N	Y	Y
Date Fixed Effects	N	Y	N	Y
Observations	1,365,074			
Adjusted R^2	0.0001	0.1003	0.1890	0.2899
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Table A12: Baseline Specification with 25-mile Treatment

Intercept	-2.7084*** (0.0014)	---	---	---
$(online\ 25)_{it}$	-0.1195*** (0.0030)	-0.2831*** (0.0030)	0.1887*** (0.0042)	-0.1702*** (0.0046)
Zip Fixed Effects	N	N	Y	Y
Date Fixed Effects	N	Y	N	Y
Observations	1,365,074			
Adjusted R^2	0.0012	0.1047	0.1898	0.2905
Significance Codes:	'*' 0.05	'**' 0.01	'***' 0.001	

Table A13: 15-mile Treatment and 25-mile Treatment
Interacted with Demographic Variables

	$(online\ 15)_{it}$	$(online\ 25)_{it}$
	0.0583** (0.0190)	-0.2894*** (0.0142)
$\times\ (metro)_i$	0.1053*** (0.0103)	-0.0171* (0.0081)
$\times\ (prop\ vets)_i$	-0.4835* (0.2289)	3.0437*** (0.1787)
$\times\ (prop\ snap)_i$	-1.2075*** (0.1118)	-1.1709*** (0.0909)
Observations	1,365,074	1,365,074
Adjusted R^2	0.29	0.29
Zip FE + Date FE Included		
Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001		

Table A14: Heterogeneous-by-state
15-mile Treatment and 25-mile Treatment

	$15\ miles$	$25\ miles$
$Michigan$	-0.0709*** (0.0113)	-0.0759*** (0.0097)
$Indiana$	-0.1105*** (0.0092)	-0.2151*** (0.0066)
$West\ Virginia$	-0.1225*** (0.0094)	-0.1866*** (0.0078)
$Pennsylvania$	-0.0269* (0.0105)	-0.0262** (0.0087)
Observations	1,365,074	1,365,074
Adjusted R^2	0.29	0.29
Zip FE + Date FE Included		
Significance Codes: ‘*’ 0.05 ‘**’ 0.01 ‘***’ 0.001		

Table A15: Temporal Variables Interacted with 15-mile Treatment

		$\times (time)_t$	$\times (jackpot)_t$	$\times \binom{square}{jackpot}_t$	$\times \binom{1st\ of}{month}_t$	$\times \binom{2nd\ of}{month}_t$
$\binom{online}{MI\ 15}_{it}$	-0.0573 (0.0513)	-0.002 (0.01)	0.001 (0.01)	-0.0001 (0.0005)	0.0964 (0.0595)	-0.0050 (0.0537)
$\binom{online}{IN\ 15}_{it}$	-0.0768*** (0.0188)	-0.003 (0.002)	-0.004 (0.005)	0.0003 (0.0004)	-0.1089*** (0.0322)	-0.0497 (0.0309)
$\binom{online}{WV\ 15}_{it}$	-0.1986*** (0.0162)	0.02*** (0.002)	0.001 (0.005)	-0.0006 (0.0003)	-0.2550*** (0.0028)	-0.1496*** (0.0268)
$\binom{online}{PA\ 15}_{it}$	0.0483** (0.0176)	-0.009*** (0.002)	-0.01** (0.005)	0.0007. (0.0004)	0.1889*** (0.0306)	0.1292*** (0.0297)
Observations	1,365,074					
Adjusted R^2	0.29					
Zip FE + Date FE Included				Significance Codes:		
Jackpot value in 100 millions, time in 100s of days				'*' 0.05	'**' 0.01	'***' 0.001

Table A16: Temporal Variables Interacted with 25-mile Treatment

		$\times (time)_t$	$\times (jackpot)_t$	$\times \binom{square}{jackpot}_t$	$\times \binom{1st\ of}{month}_t$	$\times \binom{2nd\ of}{month}_t$
$\binom{online}{MI\ 25}_{it}$	0.1241** (0.0044)	-0.02*** (0.005)	0.002 (0.007)	-0.0001 (0.0004)	0.1273* (0.0511)	-0.0092 (0.0459)
$\binom{online}{IN\ 25}_{it}$	-0.3280*** (0.0135)	0.02*** (0.002)	-0.006 (0.004)	0.0006* (0.0003)	-0.1108*** (0.0231)	-0.0423. (0.0225)
$\binom{online}{WV\ 25}_{it}$	-0.2078*** (0.0134)	0.009*** (0.001)	-0.004 (0.004)	-0.00008 (0.0003)	-0.3282*** (0.0234)	-0.1854*** (0.0025)
$\binom{online}{PA\ 25}_{it}$	0.0231 (0.0146)	-0.008*** (0.002)	-0.007 (0.004)	0.0005 (0.0003)	0.2350*** (0.0256)	0.1217*** (0.0248)
Observations	1,365,074					
Adjusted R^2	0.29					
Zip FE + Date FE Included				Significance Codes:		
Jackpot value in 100 millions, time in 100s of days				'*' 0.05	'**' 0.01	'***' 0.001