Clouded Thinking:

Behavioral Impacts of Air Pollution on Lottery Purchases

Clayson Shumway

Abstract

Environmental conditions, such as temperature and pollution, have been shown to influence human behavior, including risk-taking, criminal activity, and cognitive performance. I investigate how temperature and air pollution impact lottery sales in North Carolina. Using data on daily lottery ticket sales at the zip code level from July 2006 to June 2023, I analyze the relationship between environmental conditions and gambling behavior. I find results that suggest an additional 10 days of PM2.5 above 20 μ/m^3 would result in $1 million of increased revenue from the state lottery. Using lagged precipitation as an instrument for air pollution yields consistent results. Extreme temperatures, whether hot or cold, reduce lottery sales relative to a 65°F day. My research contributes to the literature by providing an example of the effect of pollution on cognitive behavior across a large geographic area.

1 Introduction

The effects of pollution and high temperatures are diverse. A large literature has examined the total cost of pollution and determined the costs of increased temperature from climate change. Environmental factors often have heterogeneous effects across space. Pollution and heat have been shown to increase risk-taking behavior, exacerbate crime rates, and diminish cognitive performance. Gambling represents another risk-loving behavior that can be impacted by these environmental factors.

In North Carolina, the lottery is used as a source of additional revenue to fund schools. While the state disburses more lottery revenue per capita to counties that are classified as economically distressed, these counties tend to expend more money per capita on lottery tickets to such an extent that "poorer" counties receive less than they contribute. The lottery therefore functions in practice as an elective tax that results in a reallocation of

resources from relatively poor counties to wealthier counties. Increases in pollution or heat—or changes to the spatial distributions of these factors from climate change—could exacerbate these disparities. Furthermore, lottery funds are intended to be extra funding for schools that is added to the state's contribution, but it is unclear to what extent the legislature adjusts its decisions about education funds due to expected lottery revenue.

Using data on sales of lottery tickets, I assess the relationship between temperature, pollution, and lottery sales throughout the lifetime of the North Carolina Lottery (2006 - present). Days with higher concentrations of particulate matter experience higher per capita lottery sales, which would result in an additional $100,000 per day if every day experiences elevated particulate matter concentration. I find that hot temperatures do not increase lottery ticket sales, but that sales peak around 65 degrees Fahrenheit, which would be consistent with sales increasing on days when consumers are more likely to go out. I apply an instrumental variables strategy to address concerns about endogeneity in the ambient concentration of particulate matter, and find consistent results. Furthermore, I evaluate how the relationship is impacted by economic conditions in each of the counties.

Whether lottery tickets are necessarily a "bad" in their own right is unclear. While lottery games have a negative expected value, one could argue that a rational consumer purchases lottery tickets because the utility from the thrill of scratching off a potential winner is greater than the expected loss. However, lottery tickets can be considered bad if they induce purchasers to deviate from rational behavior, or if they become habit-forming. Lab experiments that randomize participants into engaging with the lottery demonstrate that exposure can diminish self control (Kim 2013) and just under 5% of the US population can be considered problem gamblers (Clotfelter 2024). Across the country, lottery sales tend to be correlated with higher crime rates and poorer economic outcomes. In North Carolina, I present summary statistics to describe the relationship between counties' economic distress status and per capita lottery ticket sales; however, I do not find evidence that this relationship is exacerbated by pollution or temperature. I also implement the methodology of Dynan (2000) to assess the degree to which lottery sales appear to be habit forming at the county level. Last, I document that per capita sales are weakly correlated with per capita

car crashes in North Carolina in 2022.

2 Literature Review

I contribute to two strands of literature. The first examines the behavioral and social effects of environmental factors, including pollution, temperature, and other ambient conditions. The second explores the determinants of risky financial behavior, such as gambling and lottery play.

The detrimental effects of particulate matter pollution are well documented, ranging from respiratory illness to cognitive performance. The paper most closely related to mine is Herrnstadt et al. (2021), who exploit variation in wind direction to demonstrate the relationship between pollution and a societal bad—crime rates—in Chicago. Other research connects short-term particulate matter exposure to test scores (Zhang et al. 2018) and suicide risk among vulnerable populations (Molitor et al. 2023).

Many of the same undesirable outcomes have also been linked with temperature exposure. For example, Mullins and White (2019) find that emergency visits for mental health-related causes increase with abnormally hot temperatures and give suggestive evidence that a lack of sleep is the mechanism behind the association. Heilmann et al. (2021) demonstrate that crime rates in Los Angeles are associated with high maximum temperatures. Burke et al. (2015) provide a review of the climate and conflict literature.

Two theories exist regarding the association between temperature and crime. First, increased average temperatures could be related to aggressive behavior at the upper end of the temperature range. This effect is known as the Temperature-Aggression hypothesis. Second, moderately higher temperatures during colder months may increase social interaction, leading to higher crime rates. Both of these factors could also influence lottery sales, as increased movement may lead to higher foot traffic at gas stations and other lottery ticket vendors.

I contribute to the literature in three ways. First, I document another setting in which the social cost of pollution is evident. Second, the lottery data allow me to make a conclusion

at a larger spatial scale than many previous studies who use daily variation in cities. Last, I conclude that high temperatures do not drive lottery sales.

3 Background

The North Carolina legislature passed the act to create the lottery on August 30, 2005. Proponents of the lottery strategically timed the vote for a day when several key opponents were out of town, allowing the Lieutenant Governor to break a tie. Under normal circumstances, the bill would likely not have passed. Since then, the North Carolina Lottery has grown in sales, even after adjusting for inflation. The lottery initially offered some instant-win games as well as the multi-state jackpot game Powerball. In 2010, the state began selling tickets for Mega Millions as well.

Lottery games have a negative expected return, or the expected winnings of a ticket are less than the price. The expected loss has been decreasing (less negative) and careful timing in jackpot games can improve chances of winning, but the expected return remains negative. Polls of lottery players reveal monetary interest outweigh the "fun" aspect as motivation to play, although this trend reverses at the upward tail of the income distribution. Kahneman and Tversky 1979 reports that a majority prefer a sure $400 to a 50-50 chance at $1000, although they do not explain where these data are from. However, I recreated this survey and it holds true for respondents on my Instagram story.

Lottery participants could be rational, albeit misinformed actors. The probability of winning the jackpot is low, but advertising that emphasizes winners can distort beliefs into a higher expected chance of winning.

4 Data

I have data on lottery sales at the daily level for zip codes from July 2006 through June 2023. I convert sales totals to 2023 dollars using the Consumer Price Index (CPI). I aggregate zip code-level sales to the county level to take advantage of more accurate and

frequent measures of population. This aggregation is provided using the crosswalk from UnitedStatesZipCodes.org. I utilize the 1km by 1km gridded PM2.5 dataset compiled by Di et al. (2019). These data are available from 2006 to 2016, and thus the main sample is restricted to these years.

Figure 1: Trends in Sales and Pollution

Figure depicts trends in lottery sales and PM2.5 concentration. The dark line denotes the mean, while the shaded region depicts the inter-quartile range.

PM2.5 concentrations exhibit seasonal and time trends. Lottery sales exhibit trends across time, as well as across days of the week. The trends are displayed in Figure 1. Even after accounting for inflation, lottery sales increase over time. As I discuss below, in per capita terms, lottery sales are highest in counties that are considered economically distressed by the state Department of Commerce (see 10). Lottery sales vary by day of the week, with Sunday having the lowest sales but are relatively constant throughout the year. Pollution levels improve across the period of the study, with substantial variation across months of the the year, but levels are steady between days of the week.

To account for economic trends, I control for the monthly unemployment rate in each county using data from the North Carolina Department of Commerce and the yearly GDP of each county from the Bureau of Economic Analysis. I also include the size of the jackpot using historical data from the Powerball website, where jackpot size is updated every 3 days

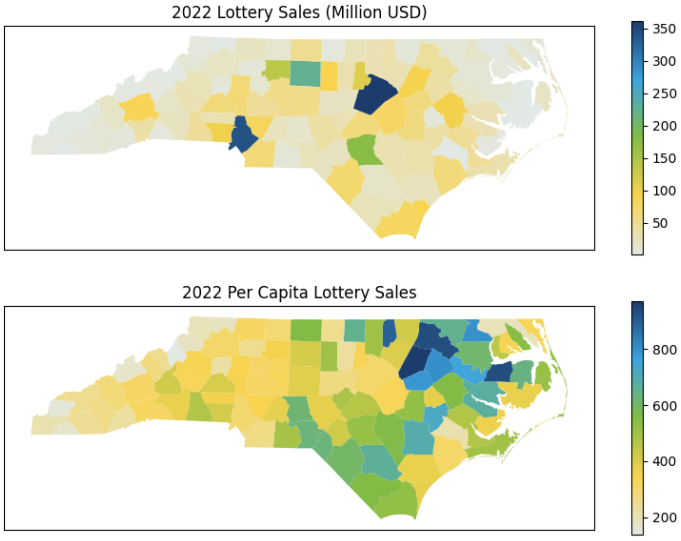

Figure 2: Lottery Sales by County

coinciding with draw dates and interpolate linearly for missing days. As depicted in Figure 2, the majority of lottery sales occur in the population centers in the state, but the eastern counties have higher lottery sales by per capita standards.

5 Main Specification

I estimate the relationship between pollution, lottery sales, and temperature non-parametrically through a binning approach. My outcome variable is per capita sales by county and the explanatory variables are indicators for the mean daily pollution and maximum temperature falling into predefined pollution and temperature bins. The monthly unemployment rate for the county, yearly gross county product, jackpot size, precipitation, and air pressure are included as controls. I estimate the following equation:

$$y_{cd} = \beta_0 + \Sigma_{b=0}^{50}\gamma_b poll_{bcd} + \Sigma_{b=40}^{90}\delta_b temp_{bcd} + \beta_u X_{cd} + \phi_k + \phi_c + \phi_t + \epsilon_{cd} \quad (1)$$

Temperature bins are 5 degrees Fahrenheit wide, and pollution bins are $3\mu g/m^3$ wide. Extreme values are grouped into the nearest bin. Counties are indexed by c, dates are indexed by d, days of the week denoted by k, and t represents years. Fixed effects are included for each county, day of week, and year. The coefficients from this equation are displayed in Figure 3. The blue bars represent a count of how many county-days fall into each pollution bin.

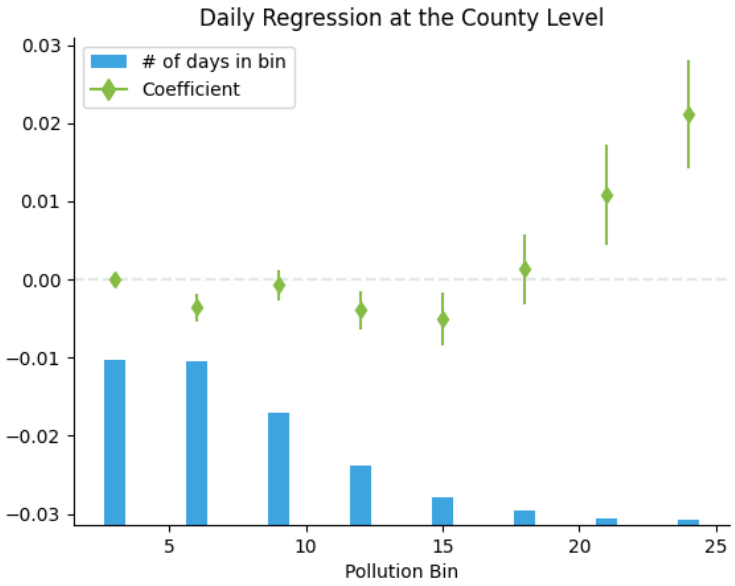

Figure 3: Coefficients from Bin Regression - Pollution

Regression coefficients for the effect of PM2.5 on lottery sales at the county-day level. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, precipitation, and temperature are included as controls. The histogram depicts the frequency of each pollution bin.

While most days fall into the lower bins, the 2% of days with pollution above 20 $\mu g/m^3$ are associated with higher spending on lottery tickets. The outcome variable is in per capita terms, so increased levels of pollution resulting in one more day in the range of 20-25 $\mu g/m^3$ PM2.5 would result in an increase of roughly $100,000 from the state lottery.

The binned results suggest that the relationship between lottery sales and air quality are not linear. I estimate a second specification in which PM2.5 concentration and temperature

enter as third-degree polynominals. The controls and fixed effects are the same as in the binned specification, but temperature is also included as a third-degree polynominal. The predicted means and estimated marginal effect of PM2.5 from this regression are presented in Figure 4. The results are similar, with the marginal effect of a day with PM2.5 above 20 $\mu g/m^3$ being a 10 cent increase in per capita sales.

$$y_{cd} = \beta_0 + f(poll_{bcd}, \beta_{PM}) + f(temp_{bcd}, \beta_{temp}) + \beta_u X_{cd} + \phi_k + \phi_c + \phi_t + \epsilon_{cd} \quad (2)$$

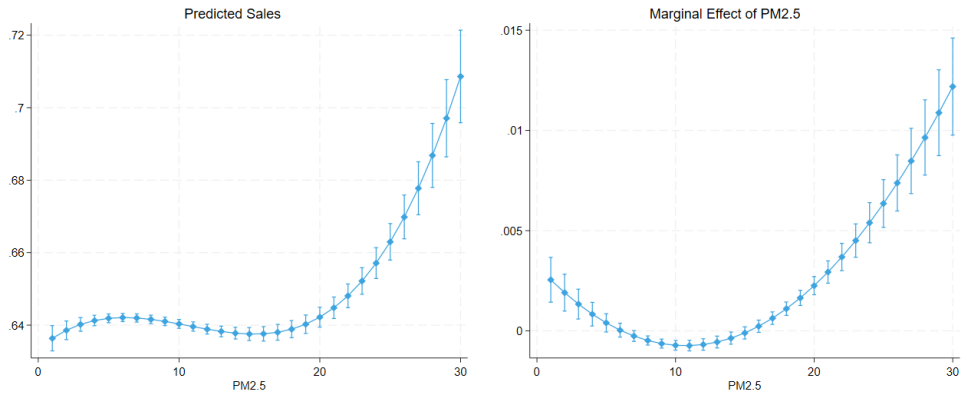

Figure 4: Estimated Effects of PM2.5 on Per Capita Sales

Regression coefficients for the marginal effect of PM2.5 on lottery sales at the county-day level using a 3rd degree polynomial. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, precipitation, and temperature are included as controls.

To assess to what extent pollution exacerbates the existing inequity in lottery expenditures, I interact each pollution bin with a county's assigned economic distress tier from the North Carolina Department of Commerce. The results are presented in Figure 5. Higher particulate matter concentrations have no statistically meaningful effect on sales in distressed counties. Based on this heterogeneity analysis, the effect in the main specification is concentrated in counties with lower economic distress.

The results for the main binned specification with temperature are displayed in Figure 6. Extreme temperatures, both hot and cold, are associated with lower per capita sales. These

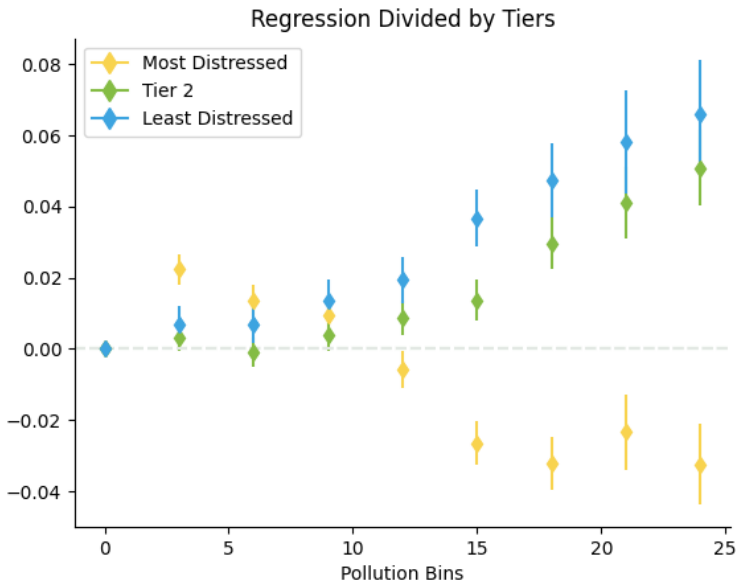

Figure 5: Coefficients from Bin Regression - By Tiers

Regression coefficients for the marginal effect of PM2.5 on lottery sales at the county-day level in each of 3 economic distress tiers. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, precipitation, and temperature are included as controls.

results could be consistent with people buying fewer lottery tickets when extreme weather causes them to venture outside less frequently.

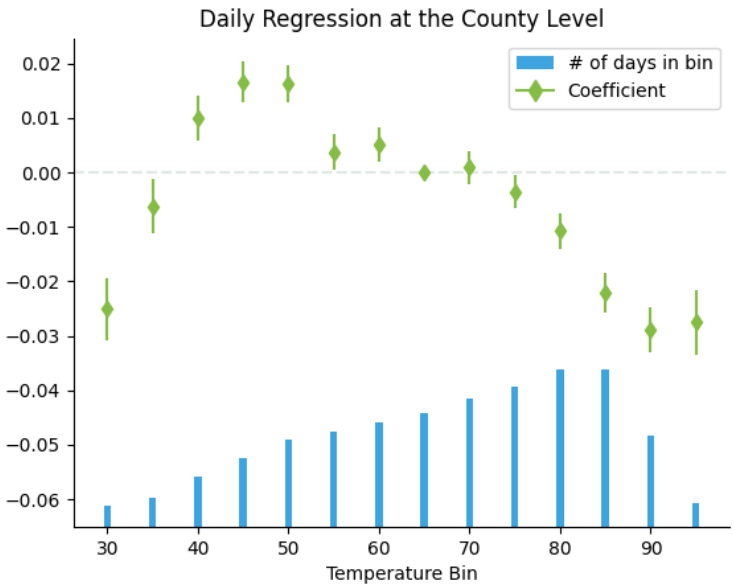

Figure 6: Coefficients from Bin Regression - Temperature

Regression coefficients for the marginal effect of temperature on lottery sales at the county-day level. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, precipitation, and PM2.5 are included as controls. The blue bars depict a histogram of each temperature bin's frequency.

6 Instrumental Variables

The main specification could suffer from endogeneity because lottery sales and air pollution may both be driven by unobserved economic activity. I include gross county product and the monthly unemployment rate as controls, but these measures vary only annually or monthly, limiting their ability to capture daily economic fluctuations. I use one day lagged precipitation and wind speed as instruments for daily air pollution, as these variables impact air pollution but are plausibly uncorrelated with lottery sales.

$$f(poll)_{cd} = \beta_0 + f(precip_{bc,d-1}, \beta_{precip}) + f(wind_{bcd}, \beta_{wind}) + f(temp_{bcd}, \beta_{temp}) + \beta_u X_{cd} + \phi_k + \phi_c + \phi_t + \epsilon_{cd} \quad (3)$$

$$y_{cd} = \beta_0 + f(\hat{poll_{bcd1}}, \beta_{poll}) + f(temp_{bcd}, \beta_{temp}) + \beta_u X_{cd} + \phi_k + \phi_c + \phi_t + \epsilon_{cd} \quad (4)$$

For the instrumental variables specification, I include PM2.5 as a 3rd degree polynomial, and use 3rd degree polynomials of wind speed and lagged precipitation as my instruments (see first stage Equation 3 and second stage Equation 4). I control for temperature with a 3rd degree polynomial. Other controls are present precipitation, unemployment rate, jackpot size, month, year, and day of week fixed effects. I plot the marginal effect of PM2.5 in Figure 7. The first stage F-statistics are 6786, 2689, and 422 and the R^2 values are 0.32, 0.2, and 0.04. My estimates are imprecise, but the point estimates for the effect of PM2.5 are positive and in a similar range to the point estimates in my main specification.

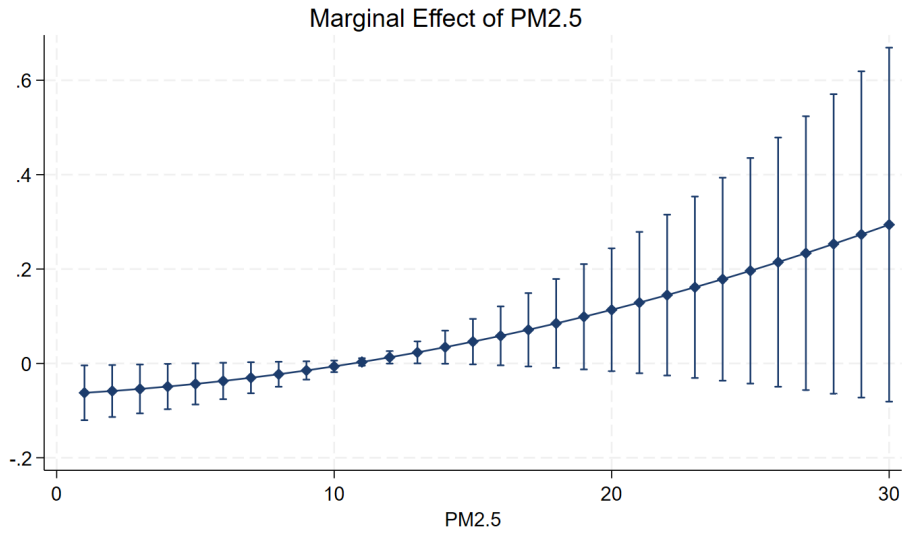

Figure 7: IV Results

Regression coefficients for the marginal effect of PM2.5 on lottery sales at the county-day level using an instrumental variables method. A 3rd degree polynomial of PM2.5 is instrumented for by a 3rd degree polynomial of both one day lagged precipitation and wind speed. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, current precipitation, and temperature are included as controls.

The pollution data are modeled from remote sensing and current monitor readings. Since precipitation and wind speed are inputs to this model, I also present IV results using the subset of counties which possess an EPA monitoring station for PM2.5. This sample is spatially smaller, but allows inclusion of daily sales through June 2023. First, I demonstrate that the results from Figure 3 also hold in this subsample. Running the same specification, the results are presented in Figure 8. The results of the marginal effect produced by the IV specification are presented in Figure 9. The first stage F-statistics for this specification are 1195, 566, and 137. The R^2 is 0.28, 0.14, and 0.03. These results are similar to those from my main specification and the IV using modeled pollution data. The marginal effect is positive and statistically significant for moderate to high levels of PM2.5.

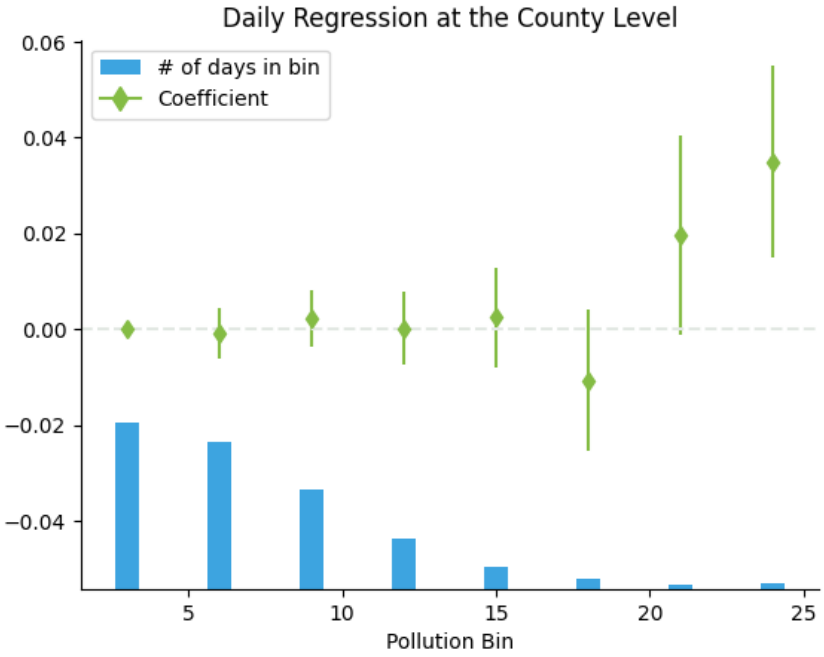

Figure 8: Binned Results on Selected Sample

Regression coefficients for the marginal effect of PM2.5 on lottery sales at the county-day level. The sample is restricted to counties with a PM2.5 monitor. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, precipitation, and temperature are included as controls.

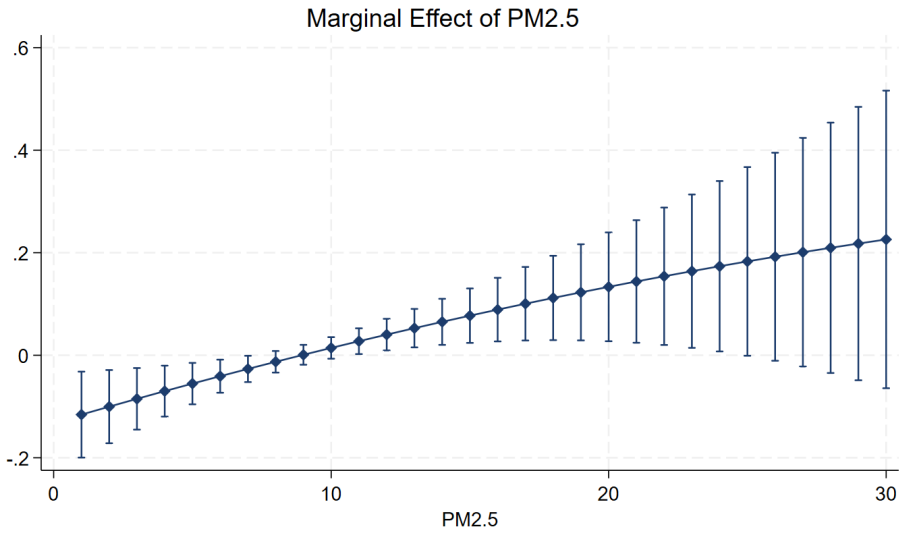

Figure 9: IV Results on Selected Sample

Regression coefficients for the marginal effect of PM2.5 on lottery sales at the county-day level using an instrumental variables method. The sample only includes counties with a PM2.5 monitor. A 3rd degree polynomial of PM2.5 is instrumented for by a 3rd degree polynomial of both one day lagged precipitation and wind speed. Standard errors are robust and vertical lines represent 95% confidence intervals. Fixed effects are included for year, month, day of week, and county. Jackpot size, unemployment rate, current precipitation, and temperature are included as controls.

7 Negative Associations with Lottery Sales

The North Carolina State Lottery is regressive. Using county GDP, the unemployment rate, population growth, income, and housing prices, the Department of Commerce assigns counties to one of three "distress" tiers. With exactly 100 counties, Tier 1 and 2 contain 40 counties while the least distressed tier 3 contains only 20. As depicted in Figure 10, Tier 1 (most distressed) has higher median per capita spending than the other tiers. The state legislature also favors Tier 1 counties when redistributing lottery revenue from education funding. Nonetheless, Tier 1 counties receive back a smaller ratio of their lottery expenditure.

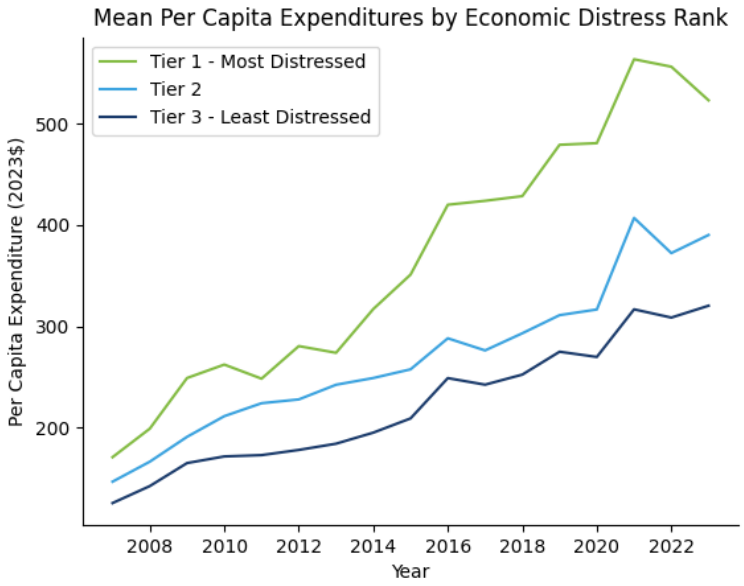

Figure 10: Lottery Spending by Distress Tier

Figure depicts mean per capita expenditure on lottery tickets by economic distress tier. Dollar values are converted to January 2023 using the CPI. Means are a population weighted average.

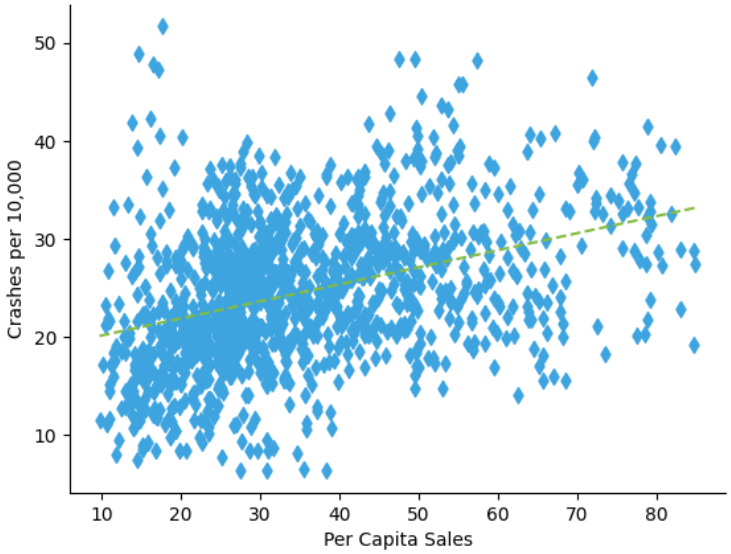

Figure 11: Correlation of Per Capita Sales and Vehicle Crashes per 10,000 Residents

Scatterplot depicts per capita car crashes and per capita sales of lottery tickets in North Carolina in 2022. Each data point represents a county-month.

7.1 Unsafe Driving

Using monthly data on vehicle crashes in 2022 from the North Carolina Department of Transportation, I examine at the correlation between per capita lottery sales and vehicle crashes per 10,000 residents. Figure 11 displays the scatter plot of county-month observations with a line of best fit overlaid.

Table 1: Correlation of Per Capita Sales and Crashes

	(1) Crashes Per 10k	(2) Crashes Per 10k	(3) Crashes Per 10k	(4) Crashes Per 10k
Intercept	18.413***	18.498***	15.693***	19.930***
	(0.507)	(0.475)	(2.014)	(2.708)
Per Capita Sales	0.174***	0.172***	0.252***	0.131*
	(0.013)	(0.012)	(0.058)	(0.077)
County Fixed Effects			X	X
Month Fixed Effects		X		X
R-squared	0.139	0.150	0.022	0.005

Table shows how the correlation between car crashes and per capita lottery spending is impacted by the inclusion of county and month fixed effects. $* = p < .1, ** = p < .05, *** = p < .01$.

The correlation coefficient is 0.37 for the raw data. A correlational regression table including fixed effects for month and county to capture seasonality in travel and sales is presented below. While the fixed effects mitigate this concern to some extent, I caution that county-level vehicle miles traveled (VMT) data are unavailable. Preferably my crash data should be per VMT rather than per population. Since VMT and per capita lottery sales tend to be higher in rural counties, I believe this specification overstates the relationship to some extent.

7.2 Habit Formation

Surveys estimate that 4.6% of the US population qualify as problem gamblers, while 1% of adults are considered pathological gamblers (Clotfelter 2024). The State of North Carolina recently increased spending on the gambling help hotline to $3 million each year. As such, I provide some evidence that gambling is habit forming at an aggregated spatial scope. I estimate the following equation for $j = 100$ in steps of 5 and t denoting weeks. County level

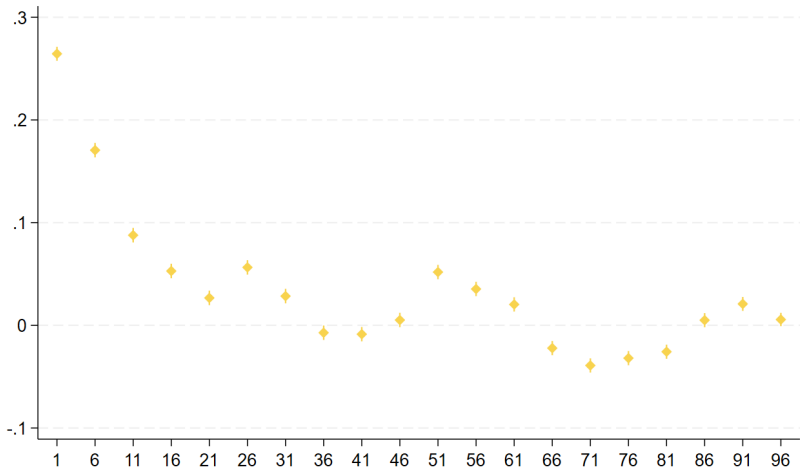

Figure 12: Habit Formation in Lottery Sales

Figure depicts coefficients of a regression of lottery sales on past lottery sales. Vertical lines denote 95% confidence intervals. Each week from 1 to 96 by 5 is included. Jackpot size and unemployment rate are included as controls, as are fixed effects for county and week.

sales are summed to the week level.

$$ln(sales_{i,t}) = \beta_0 + \sum_{j}^{J} \beta_j ln(sales_{i,t-j}) + ln(\psi_{i,t}) + \nu_i + \nu_t + \nu_k + \epsilon_{i,t} \tag{5}$$

The results in Figure 12 show that recent weeks are highly correlated with sales, but that all recent weeks are included, the effect drops to 0 after 30 weeks. Figure 13 displays coefficients when each regression is run separately, so I estimate Equation 6 for $j \in \{1,6,11,\ldots,96\}$.

$$ln(sales_{i,t}) = \beta_0 + \beta_j ln(sales_{i,t-j}) + ln(\psi_{i,t}) + \nu_i + \nu_t + \nu_k + \epsilon_{i,t} \tag{6}$$

Using this specification, lagged lottery sales in a county have a lasting effect, increasing current lottery sales for up to 70 weeks, or about a year and a half. Lottery sales grow consistently across time. In order to ensure that I am not simply picking up this growth effect, I also run this specification using a detrended outcome. I calculate the deviation from the trend as given in Equation 7 where predicted sales are the fitted values from Equation

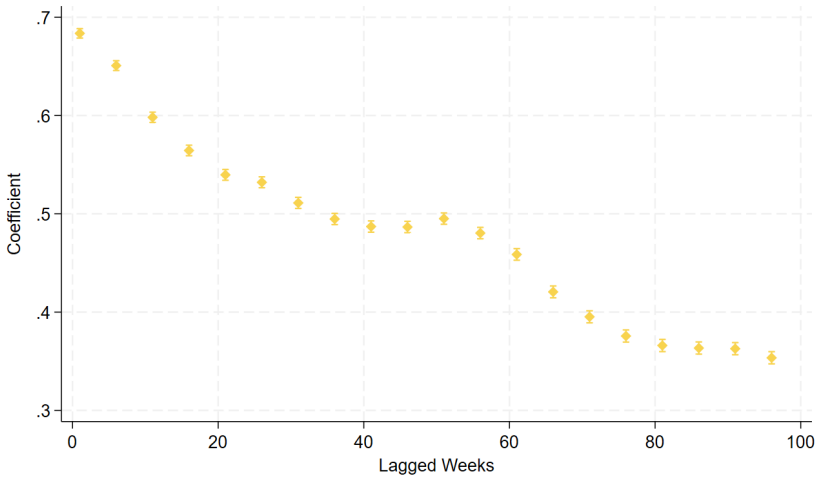

Figure 13: Habit Formation in Lottery Sales

Figure depicts coefficients of a regression of lottery sales on past lottery sales. Vertical lines denote 95% confidence intervals. Each coefficient comes from a separate regression. Jackpot size and unemployment rate are included as controls, as are fixed effects for county and week.

8 with $f(week_t)$ being a third degree polynomial function of weeks since the beginning of the North Carolina State Lottery.

$$deviation_{it} = \frac{log(sales_{it})}{log(predictedsales_{it})} \tag{7}$$

$$ln(sales_{it}) = \beta_0 + \Sigma f(week) \times \mathbf{1}(county_i) + \Gamma X_{it} + \epsilon_{it} \tag{8}$$

If lottery sales are habit forming, this effect would contribute to the trend. I do not attempt to correct for this feature, so my result underestimates the habit forming effect of lottery sales. The results from this regression are presented in Figure 14. The results of these regressions also denote a lasting effect of 70 weeks.

8 Conclusion

This paper examines the relationship between environmental factors and risk-taking behavior using lottery sales in North Carolina as a proxy. Using county-level daily sales from 2006

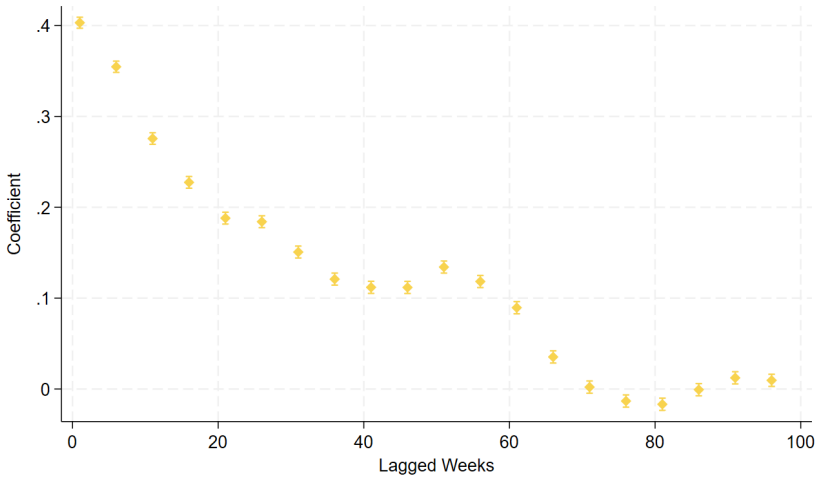

Figure 14: Habit Formation Regression on Detrended Sales

Figure depicts coefficients of a regression of detrended lottery sales on lagged detrended lottery sales. Vertical lines denote 95% confidence intervals. Each coefficient comes from a separate regression. Jackpot size and unemployment rate are included as controls, as are fixed effects for county and week.

to 2016, I document a nonlinear relationship between particulate matter concentrations and per capita lottery sales. Each additional day in the 20-25 $\mu g/m^3$ range results in an increase of $100,000 million of lottery revenue. Temperature has a hump-shaped effect, where highest sales occur on days with moderate temperature. This result suggests that mobility dominates any psychological effect.

I employ an instrumental variables strategy using wind speed and a one day lag of precipitation. My results confirm the main specification results. I also find heterogeneous effects, with economically distressed counties being less sensitive to pollution. Whether lottery tickets function as a normal good or a good with negative externalities is unclear. I provide evidence of habit formation in lottery sales and share evidence of positive correlation of lottery sales with car crashes. Insofar as lottery tickets serve as an adequate proxy for risk taking, my results demonstrate additional costs of air pollution concentrations. These negative impacts are mitigated somewhat by the side effect of increased funding for North Carolina schools through lottery sales.

References

Burke, M., S. M. Hsiang, and E. Miguel (2015). "Climate and Conflict". *Annual Review of Economics* 7.Volume 7, 2015. Publisher: Annual Reviews, pp. 577–617. DOI: 10.1146/annurev-economics-080614-115430.

Clotfelter, C. T. (2024). "Better State Lotteries". *Public Finance Review* 52.3. Publisher: SAGE Publications Inc, pp. 287–313. DOI: 10.1177/10911421231206030.

Di, Q. et al. (2019). "An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution". *Environment International* 130, p. 104909. DOI: 10.1016/j.envint.2019.104909.

Dynan, K. E. (2000). "Habit Formation in Consumer Preferences: Evidence from Panel Data". *American Economic Review* 90.3, pp. 391–406. DOI: 10.1257/aer.90.3.391.

Heilmann, K., M. E. Kahn, and C. K. Tang (2021). "The urban crime and heat gradient in high and low poverty areas". *Journal of Public Economics* 197, p. 104408. DOI: 10.1016/j.jpubeco.2021.104408.

Herrnstadt, E., A. Heyes, E. Muehlegger, and S. Saberian (2021). "Air Pollution and Criminal Activity: Microgeographic Evidence from Chicago". *American Economic Journal: Applied Economics* 13.4, pp. 70–100. DOI: 10.1257/app.20190091.

Kim, H. ((2013). "Situational Materialism: How Entering Lotteries May Undermine Self-Control". *Journal of Consumer Research* 40.4, pp. 759–772. DOI: 10.1086/673191.

Molitor, D., J. T. Mullins, and C. White (2023). "Air pollution and suicide in rural and urban America: Evidence from wildfire smoke". *Proceedings of the National Academy of Sciences* 120.38. Publisher: Proceedings of the National Academy of Sciences, e2221621120. DOI: 10.1073/pnas.2221621120.

Mullins, J. T. and C. White (2019). "Temperature and mental health: Evidence from the spectrum of mental health outcomes". *Journal of Health Economics* 68, p. 102240. DOI: 10.1016/j.jhealeco.2019.102240.

Zhang, X., X. Chen, and X. Zhang (2018). "The impact of exposure to air pollution on cognitive performance". *Proceedings of the National Academy of Sciences* 115.37. Publisher: Proceedings of the National Academy of Sciences, pp. 9193–9197. DOI: 10.1073/pnas.1809474115.