NBER WORKING PAPER SERIES

IS LOTTERY GAMBLING ADDICTIVE?

Jonathan Guryan
Melissa Schettini Kearney

Working Paper 14742
http://www.nber.org/papers/w14742

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2009

The authors thank Botond Koszegi, Michael Grossman, Kevin Murphy, Richard Thaler, and Judy Hellerstein for helpful comments. We also thank Seth Freedman and Andriy Protsyk for valuable research assistance. Jonathan Guryan (jguryan@chicagobooth.edu), Melissa S. Kearney (kearney@econ.umd.edu).

Is Lottery Gambling Addictive?
Jonathan Guryan and Melissa Schettini Kearney
NBER Working Paper No. 14742
February 2009
JEL No. D14,D62,D81,H42,H71

ABSTRACT

We present an empirical test for the addictiveness of lottery gambling. To distinguish state dependence from serial correlation, we exploit an exogenous shock to local market consumption of lottery gambling. We use the sale of a winning ticket in the zip code, the location of which is random conditional on sales, as an instrument for present consumption and test for a causal relationship between present and future consumption. This test of addiction is based on the definition of addiction commonly used in the economics literature. It has two key advantages over previous tests for addiction. First, our test is unique in being based on an observed increase in consumption coming from a randomly assigned shock. Second, our approach estimates the time path of persistence non-parametrically. Our data from the Texas State Lottery suggests that after 6 months, roughly half of the initial increase in lottery consumption is maintained. After 18 months, roughly 40 percent of the initial shock persists, though estimates become less precise. These estimates provide an upper bound on the degree of addictiveness in lottery gambling. They also highlight the potential effectiveness of innovations and advertising campaigns designed to increase lottery gambling.

Jonathan Guryan
University of Chicago
Booth School of Business
5807 S. Woodlawn Ave.
Chicago, IL 60637
and NBER
jguryan@chicagobooth.edu

Melissa Schettini Kearney
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
and NBER
kearney@econ.umd.edu

I. Introduction

State lotteries are frequently promoted as an alternative to explicit taxation as a means of public finance. Lottery gambling in the United States is only legally available as a state government product. Private lotteries are illegal in all 50 states, but 42 states currently operate a state lottery. State lotteries constitute the most common form of gambling among American adults. In a 2007 Gallup poll, 65 percent of Americans reported participation in at least one form of gambling last year; 46 percent reported participation in state lottery gambling.[1]

Americans spend a great deal on lottery tickets. Lottery ticket sales totaled $41.4 billion in 2003, yielding gross revenues for states of $19.9 billion (Christiansen Capitol Advisors, 2004). This represents annual sales of $212 per adult living in a lottery state, or $372 per household nationwide. For lower-income households, the introduction of a state lottery appears to be associated with a 2.5 percent reduction in household non-gambling expenditures, including reductions in expenditures on food and on home mortgage, rent, and bills; there is a 3.1 percent reduction in non-gambling expenditures when instant games are offered (Kearney, 2005). Because lottery tickets are sold exclusively by state-run monopolies, it is important to ask whether the shift in expenditures due to the availability of lottery gambling is consumer-welfare enhancing. From the perspective of neoclassical economics, the answer to this question depends largely on whether

[1] http://www.gallup.com/poll/104086/One-Six-Americans-Gamble-Sports.aspx (last accessed July 14, 2008).

consumers appear to be informed, rational, and potentially addicted consumers of state lottery products.[2]

We test whether lottery gambling is addictive following the definition of addiction commonly used in the economics literature (see e.g. Becker and Murphy, 1988), and investigate the extent to which past lottery consumption causally increases current lottery consumption. The level of addiction of a good is higher the greater the reinforcement of past consumption on present consumption. The most serious empirical difficulty associated with testing for addiction of this type is that it is hard to tell the difference between serial correlation in consumption (which results from stable preferences) and a causal relationship between past and current demand. To distinguish addiction from serial correlation in lottery consumption, we exploit an exogenous shock to lottery gambling, as described below. We describe a test for addiction that can be implemented as a simple instrumental variables (IV) estimator, and which depends on the usual IV rank and exclusion assumptions. Intuitively, our empirical test for addiction amounts to determining how quickly an exogenous increase in lottery gambling dissipates.

Our test of addiction is based on two separate shocks to lottery demand: increases in sales both in and around winning stores after a winning jackpot ticket is sold. Conditional on the number of tickets sold in the zip code, the location of the winning ticket is random. Therefore, the increases in sales both at the winning store and at nearby stores are randomly-assigned shocks to demand. We then trace out the persistence of these two shocks to measure the extent of addiction. To our knowledge, this is the first

[2] As Becker and Murphy (1988) have argued, addiction itself need not imply irrationality, which has implications for optimal policy. We return to this point in Section II.

economic test of addiction based on an observed shock to consumption coming from a randomly assigned exogenous event.[3]

Our empirical analysis is based on detailed sales data from the Texas Lottery Commission. We demonstrate that in the week after a large-prize winning lottery ticket is sold in a zip code, ticket sales in that zip code are 13.2 log points (14.1 percent) higher than in non-winning zip codes. This increase at the zip code level, which we use as a proxy for local area market, reflects two different responses. First, the winning store itself experiences a 32 log point (38 percent) increase in sales of the winning game. In previous work (Guryan and Kearney, 2008), we argue that this demand response is a result of an erroneous belief that the winning store is lucky, something we deem "the lucky store effect". Second, non-winning stores in the zip code experience a 4.9 log point (5.02 percent) increase in ticket sales. This is clearly not driven by a lucky store effect, but perhaps is a response to a general advertising effect or an induced increase in the subjective probability of winning the lottery, coming from having observed someone in the area coming up a winner.

As described briefly above, to test for addiction, we use the sale of a winning ticket to generate two separate instruments for lottery consumption, and test whether the resulting short run demand shocks caused persistent increases in lottery demand. As we discuss in detail in Section III, the necessary exclusion restriction requires the dissipation of whatever caused these initial demand shocks in response to the winning ticket sale. In other words, in the case of the winning stores we must assume that the belief in the lucky

[3] Some authors have used arguably exogenous movements in prices to instrument for consumption with price in the empirical framework of Becker, Grossman, and Murphy (1994). Gruber and Koszegi (2001) use cigarette taxes and Farrell, Morgenroth, and Walker (1999) use jackpot rollovers in the UK National Lottery. We discuss these papers in Section II.

store ends before the time period for which we are drawing a conclusion about addiction, for example, 6 months, 12 months, or 18 months. If this assumption is not valid, our estimates of addiction are biased upward.

Importantly, the positive shock to lottery gambling at nearby non-winning stores provides a second instrumental variable whose source is different. In previous work, we show that the sales response at the winning store is caused by a mistaken belief that that particular store is temporarily lucky. Such a belief cannot drive increased sales at nearby stores since those stores did not sell a winning ticket. We compare our estimates of consumption persistence at winning stores to estimates of consumption persistence at non-winning stores in winning zip-codes and find very similar estimates. To interpret the persistence of heightened lottery consumption at non-winning stores as addiction requires that the particular mechanism behind that shock – perhaps general advertising – has worn off. The comparability of the two sets of estimates bolsters our confidence that our estimates of addiction are capturing more than simply the persistence of the so-called lucky store effect. There is no a priori reason to expect that the two responses would have the same degree of persistence, other than a general addictiveness (or habit formation) of lottery gambling.

The assumption that the initial shock wears off is not specific to this particular test for addiction. In fact, it is an assumption that is necessary for all economic tests for addiction. Consider a research design that instruments for current consumption using current prices (e.g. Becker, Grossman and Murphy, 1994; Gruber and Koszegi, 2001). In order to identify the effect of current consumption of future consumption—the test of addiction—it is necessary to assume that current prices do not affect future consumption directly. This could be violated, for example, through income effects or contextual relative price effects.

We view our empirical approach as having two main advantages over alternative empirical tests of addiction. First, we exploit the random assignment of a demand shock. Second, our research design allows us to estimate the time path of persistence non-parametrically. Whereas in other studies researchers are forced to make assumptions about the time-series properties of the consumption, we are able to allow the data to trace out the shape of addiction over time. We are able to do this in part because we can date the source of the demand shock.

Our empirical results show that in the month after a winning Lotto Texas ticket is sold, zip code level sales are elevated by 11 log points (standard error of 1.5). We aggregate to the month level to increase statistical precision and consider this the initial exogenous shock to demand. The decay of this shock is highly non-linear, but our IV estimates show that after 6 months and 12 months, roughly half of the initial shock to consumption is maintained. [4] Depending on one's priors, this estimate of a roughly 50 percent persistence rate after a year may be interpreted as a fairly sizable degree of addiction. As we discuss below, a cautious interpretation would be that these estimates are an upper bound of lottery gambling addiction. Furthermore, regardless of the reader's favored interpretation, these estimates certainly suggest that lottery winners, innovations,

[4] It is difficult to find estimates of gambling addiction in the clinical psychology or medical literatures that are comparable to what we estimate in this study. The research in those disciplines focuses on identifying behaviors and consequences as evidence of problem or pathological gambling, such as the tendency to engage in destructive behaviors, commit crimes, accumulate debt, or have strained relationship with family and friends. The American Psychiatric Association (APA) defines pathological gambling as an impulse control disorder that manifests itself with three dimensions: damage or disruption, loss of control, and dependence. The National Gambling Impact Study Commission (1999) estimated that six percent of American adult gamblers were problem or pathological gamblers. That report also notes that the APA uses

advertising campaigns, or other outside influences on lottery gambling can have sustained effects on the level of lottery consumption. They therefore have important implications for a normative assessment of state lotteries, in particular, state lottery practices that are designed to increase lottery gambling.

II. The Economics of Addiction

As early as Hicks (1965), economists have recognized that preferences for current consumption may be affected by past consumption levels. Pollack (1970) describes a formal model in which the marginal utility of current consumption depends on past consumption. Consider a simple model of lifetime consumption with two goods, where y_t is the numeraire and g_t is potentially addictive. Consumers maximize the discounted sum of period-specific utility $U = \sum_{t=0}^{\infty} (1+\rho)^{-t} u_t$, where $u_t = u(y_t, g_t, g_{t-1})$. Note that utility is a function of both current and past values of g. We say that g is addictive if $\frac{\partial^2 u_t}{\partial g_t \partial g_{t-1}} > 0$ and $\frac{dg_t^*(p_t, I_t, g_{t-1})}{dg_{t-1}} > 0$. In other words, good g is addictive if past consumption of g influences current consumption of g by increasing the marginal utility of current consumption.

This reduced-form definition of addiction is commonly used in the economics literature (e.g. Becker and Murphy (1988), Becker, Grossman and Murphy (1994), Gruber and Koszegi (2001)). It is distinguished from a physiological or psychological definition in that learning-by-consuming qualifies as addiction. We follow the economics

---

the terms abuse and dependence, not addiction, and that the scientific concept of dependence accords with what is referred to by the lay public as "addictive" or "compulsive" gambling (page 4-1).

literature and use this definition, though the distinction between learning about a product through consumption and physiological or mental addiction should be kept in mind when interpreting the empirical results. Note that the economics definition allows for both harmful $\left(\partial u_t / \partial g_{t-1} < 0\right)$ and beneficial $\left(\partial u_t / \partial g_{t-1} > 0\right)$ addictive goods. Addiction is determined by past consumption's effect on *marginal* utilities, not by its effect on the *level* of utility.

Recent empirical work on addiction in the economics literature has focused on the distinction between myopic and rational addiction.[5] Myopic addicts ignore the future implications of current consumption on marginal utilities while rational addicts, as they are called by Becker and Murphy (1988), take these effects into account in a fully time-consistent manner. One test often used to distinguish these two models is whether anticipated future increases (decreases) in the price of an addictive good decrease (increase) current consumption. Becker, Grossman, and Murphy (1994), hereafter BGM, test for addiction of cigarette smoking. The model specifies consumption in period *t* as a function of consumption in period *t-1*, consumption in period *t+1*, and current period price. The inclusion of the lagged and lead dependent variables potentially introduces serial correlation, and hence the approach calls for instrumenting for past and future consumption with past and future price. They report results that support the model; namely, cross-period price effects are negative and long-run responses are nearly twice as large as short-run responses.

[5] Pollack (1970) explicitly assumes that consumers do not foresee the effect current consumption will have on future preferences or choices. Ryder and Heal (1973), Stigler and Becker (1977), Boyer (1978), Iannacone (1986), and most rigorously Becker and Murphy (1988) incorporate forward-looking behavior into the model.

Gruber and Koszegi (2001) build on BGM in two ways. Their first contribution is empirical. They use as instruments exogenous changes in cigarette prices coming from increases in cigarette taxes and they use consumption data, as opposed to sales data. The empirical conclusions regarding the addictiveness of cigarette consumption are largely unchanged. Their second contribution is to incorporate time-inconsistent preferences to the model of rational addiction, which changes the optimality implications of the model. In the case of myopic addiction and a potentially harmful product, the optimal commodity tax rate would be higher than if the good were not addictive, to account for self-imposed harm as well as potential externalities. In the case of rational addiction along the lines of the Becker-Murphy model, higher taxes are optimal only to the extent that the addictive behavior imposes externalities. But, Gruber and Koszegi argue, if the rational addict has time-inconsistent preferences the optimal policy involves more regulation, higher taxes, and/or higher prices, to account for both externalities and internalities.

Mobilia (1993) uses the approach of BGM to test for rational addiction in the context of horse track betting. The Becker-Murphy predictions tested are as follows. First, a decline in the price of gambling in period *t* should increase consumption in period *t*, which should also increase gambling consumption in period *t+1*. And second, if the decline is anticipated in period *t-1*, then gambling consumption in period *t-1* should also rise. The long-run response in the BGM model is the effect on consumption of a change in the price in all periods; the short-run response holds past prices constant and considers only the effect on consumption coming from a change in price in the current and future periods. Using panel data on prices (defined by the state legislated takeout rate) and bet sales from 148 horse tracks from 1950 to 1987, Mobilia finds evidence consistent with rational addiction for amount bet per attendee, but not for horse track attendance. She finds long run price elasticities on the order of -0.68, which are nearly one-third larger than the estimated elasticities derived under the model specification without addiction.

Farrell, Morgenroth and Walker (1999) test for addiction in lottery participation in the context of the U.K. National Lottery, which was introduced in November 1994. Their goal is similar to ours: they argue that the widespread popularity of lotteries has generated concern over their potential addictiveness, but that “there has been no economic research into the extent to which lottery participation is subject to addiction”. They test the prediction of the Becker-Murphy addiction model that the long run elasticity is greater than the short run elasticity. They use exogenous variation in lottery bet price (expected value) that comes from the rollovers that increase the lotto jackpot. The authors find that the average rollover raises sales by 20 percent. They estimate a larger long-run price elasticity (-1.55) than short-run (-1.04), though the difference is not statistically significant. Their estimate of the coefficient on the lag of consumption is 0.33, compared to a corresponding estimate from BGM for cigarette addiction of 0.45. The authors interpret these findings as evidence that lottery gambling is less addictive than cigarette gambling.

III. An Empirical Test of Addiction

In contrast to the previous empirical literature described above, we do not employ the Becker-Grossman-Murphy empirical test of addiction. Our empirical approach is designed to trace out the persistence of an exogenous increase in consumption. To instrument for past consumption, we use an exogenous shock to past sales, which is driven by the sale of a winning lottery ticket. We then examine the path of sales following the random shock to lottery demand to document addiction. Our empirical approach allows us to trace out the dynamic path of persistence following the shock to demand non-parametrically. It also has a convenient interpretation that corresponds to what might be commonly considered evidence of addiction beyond the economics literature. Our approach does not, however, allow us to test for forward-looking behavior that distinguishes rational from myopic addiction because the random shocks to demand on which the tests are based are not predictable.

Motivated by the definition of addiction commonly used in the economics literature, the relationship of interest is

(1) $$g_{i,t+k} = \alpha(i,k) + \pi(i,k) g_{it} + \mu(i)_t + \varepsilon_{i,t+k}$$

Where where the i subscript indexes stores, t indexes weeks or months depending on the context, g is the log of the number of tickets sold, α is an intercept, π is the reduced form causal effect of demand in month or week t on demand in month or week $t+k$, μ is a set of month or week fixed effects, and ε is an error term. The parameters are indexed both by i and k because we are interested in the relationship at various durations (i.e. different values of k), and because we estimate relationships at both the store (i) and zip-code (z) levels. Henceforth, we drop the i or z notation on the parameters for convenience, and try to make it clear in context which is implied.

Any empirical test of addiction faces an obvious identification obstacle in that it is difficult in observational correlations between past and present consumption to distinguish serial correlation from true state dependence. Past and current consumption of a good are surely correlated simply because tastes remain fairly constant over time.

We surmount this identification issue by exploiting a temporary shock to lottery demand that results from the sale of a winning lottery ticket, which consumers may view as a change in the subjective probability of winning and thus as a change in the price. Our empirical strategy exploits the random variation in winner location. To test for addiction, we examine whether this temporary shock to lottery demand causes persistent increases in lottery consumption (i.e. beyond the length of the shock itself). It is the random shock to sales coming from the sale of a winning ticket that allows us to distinguish between serial correlation and persistence.

The research design can be viewed as a two-stage least squares design where the sale of a winning ticket is used to instrument for lottery sales immediately after the winning ticket is announced. This first-stage relationship is then used to identify whether short-run lottery demand responses causally lead to subsequent increases in lottery consumption. This causal relationship—as distinguished from a correlation—matches the commonly used economic definition of addiction that we described above. The key point is that the level of addiction is higher the greater the reinforcement of past consumption on present consumption.

The first stage of this two-stage test of addiction is of the following form

$$g_{i,t+1} = \alpha(1) + \gamma(1)\, w_{it} + \phi(1)\, g_{it} + \mu_{t+1} + e_{i,t+1} \tag{2}$$

where *w*, the instrument, is a dummy variable indicating that store *i* sold a winning ticket in week or month *t*, α, γ, and ϕ are parameters to be estimated and μ is a fixed week or month effect that among other things, captures temporal variation in the jackpot, which in practice is closely linked to sales. We present estimates of this specification at both the week and month level. We aggregate the two-stage least squares estimates to the month level to increase the power of the tests. At the week level, the estimated effect of selling a winner is thus the effect relative to other stores that week, controlling for the fact that all stores will sell more tickets when the jackpot is very high and fewer tickets when the jackpot is very low. This sales response immediately following the announcement of the winning ticket is the shock to lottery demand that we will use to identify the effect of current consumption on future consumption.

We additionally estimate this first stage equation, as well as the second stage described below, at the level of the zip code. Looking at the zip code level is crucial to confirming a first stage increase in market level zip code sales. The presumed mechanisms driving the initial increase in sales at the store level and the zip code level are different. Thus, estimating the system of equations at the two levels essentially provides us with two different instrumental variable estimates of the addictiveness of lottery gambling. For clarity of exposition, the discussion in this section focuses on the store level. All assertions about identification carry over to the zip code level.

The more tickets a store sells, the more likely it is to sell a winning ticket. Since sales are serially correlated, it follows that

$$E[\varepsilon_{it} \mid w_{i,t-1} = 1] \neq E[\varepsilon_{it} \mid w_{i,t-1} = 0]$$

and therefore a simple comparison of average sales at stores that sold and did not sell winners one week ago does not recover the causal effect of the winning ticket sale. Fortunately, since each lottery ticket has the same chance of winning, the probability of a store selling a winning lottery ticket is a linear function of the number of tickets it sells in a week.[6] Thus, conditional on the number of tickets sold in week *t*, each store has the same chance of selling a winning lottery ticket. Therefore,

$$E[\varepsilon_{it} \mid w_{i,t-1}=1, g_{i,t-1}] = E[\varepsilon_{it} \mid w_{i,t-1}=0, g_{i,t-1}].$$

Serial correlation is not a problem for the estimation of $\gamma(1)$ because any two stores with the same sales in week *t* have the same chance of selling a winning ticket in week *t*, regardless of whether sales have been high for a large number of weeks or if sales are only high for one week as a result of a temporary shock. Therefore, conditional on sales in week *t* (g_{it}), w_{it} is randomly assigned, and a simple Ordinary Least Squares (OLS) estimate of $\gamma(1)$ will be unbiased. Some readers may be troubled by the inclusion of a serially correlated lagged dependent variable as a regressor. While it is true that the resulting estimate of $\phi(1)$ is not the causal effect of lagged sales on current sales, the logic above still ensures that the estimate of $\gamma(1)$ is unbiased.

This specification allows for the estimation of longer lags, by game, simply by estimating

(3) $$g_{i,t+k} = \alpha(k) + \gamma(k) w_{it} + \phi(k) g_{it} + \mu_{t+k} + e_{i,t+k}$$

where *k* is the number of weeks after the winner is sold for which the effect is to be estimated. In the following section, we estimate the dynamics of the sales effects up 18 months after the sale of the winning ticket. To do this we estimate 18 different versions

[6] This is not entirely accurate. The probability of selling a winning ticket is a linear function of the number of unique number combinations sold. If store A sells X tickets, representing λX unique combinations, and store B sells X tickets, representing λ'X unique combinations, where λ' > λ, then store B has a higher probability of selling a winning ticket. We have no way of knowing in the data how many unique number combinations were sold, only how many tickets. Furthermore, there is no reason to suspect that the proportion of tickets that reflect unique combinations varies systematically across stores. We thus make the simplifying assumption that the number of unique number combinations a store sells is a fixed proportion of the number of tickets the store sells.

of the above specification (the first of which is the first stage described in equation (2) above), aggregating sales to the month level in order to increase statistical precision. These estimates at longer lags are the second stage regressions that we use to test for addiction.

Under assumptions that we discuss just below, the ratio $\gamma(k)/\gamma(1)$ is an instrumental variables estimate of $\pi(k-1)$, the effect of induced lottery consumption one period following the winning ticket sale on lottery consumption *k* periods following the winning ticket sale. This ratio captures how much of the increase in consumption is maintained *k* weeks later. The ratio defined for $k>1$ is a measure of the persistence of a shock. Since we do not impose a distributive lag structure on the path of persistence, the estimates provide a nonparametric description of the path of persistence in response to a shock to demand as opposed to the structural relationship of consumption across periods. We consider this measure to have intuitive appeal as it corresponds to addiction as it is popularly conceived: if something causes someone to gamble $100 today (perhaps an advertising campaign or a friend's influence), how much are they likely to be gambling *k* weeks later? The larger the amount, the more addictive we would consider gambling behavior to be.

For the winning ticket sale to be a valid instrument for short run sales—and equivalently, for the test of addiction to be valid—two conditions must be met. These conditions correspond to the usual conditions necessary for a valid instrument. First, the instrument must lead to a significant increase in lottery ticket consumption. In other words, there must be a shock to lottery demand and the first stage must be significant. In our case, this amounts to estimating equation (2), and then confirming that the short-run sales response at the winning store was not offset by substitution away from other forms of lottery gambling or by substitution of sales away from other nearby stores.

Second, the instrument must not affect future demand for lottery tickets except through its effect on current lottery demand. This is the typical exclusion restriction, which in the addiction case is satisfied subject to two assumptions. We can be sure that the sale of the winning ticket is not correlated with any outside determinants of lottery sales (i.e. unobservables) once we condition on sales contemporaneous with the sale of the winning ticket. This follows from the fact that each ticket has the same chance of being a winner.

One attractive feature of the empirical specification is that it allows for a clean and direct test of this assumption. If the location of the winning store is indeed conditionally random, estimates of lead effects (i.e. estimates of $\gamma(k)$ in equation (3) with negative values of *k*) should be indistinguishable from zero. Consider the one-week lead specification. Conditional on sales in week *t+1*, the fact that a store will sell a winning ticket in week *t+1* should be uncorrelated with sales in week *t*. In other words, the logic used above to argue that serial correlation was not a problem for estimating the $\gamma(k)$'s implies that last week's sales should not predict whether a store will sell a winning ticket this week, once we have conditioned on this week's sales. Much in the spirit of an event-study, the $\gamma(k)$'s should be zero in the weeks leading up to the sale of the winner (the lead effects). In the weeks following the sale of the winner the $\gamma(k)$'s (the lag effects) are free to follow whatever path lottery consumers choose.

The second condition required for the exclusion restriction to hold is that the winning ticket itself must not have a direct effect on future lottery sales. In other words, we must assume that the mechanism driving the short run increase in sales wears off. To be precise, say the sale of the winning ticket directly causes an increase in lottery consumption for k^* periods. This could be because having observed a winner, people believe the subjective probability of winning the lottery to be greater than previously thought. Our assumption is essentially that consumers respond to that new belief within k^* periods and in addition, that after k^* periods, they no longer believe it to be true. So, if sales remain high beyond period k^*, we can conclude that the increase is causally related to the increase in sales between period 0 and k^*. According to the definition described above, heightened consumption beyond period k^* would be evidence of addiction.[7]

To the extent that this latter assumption is violated, we would be erroneously attributing the persistence in increased consumption to addiction when it is actually reflecting a continued belief in a higher winning probability. Our estimate is thus an upper bound on addiction. As noted above, we estimate the model at both the store level and the zip code level. This has the advantage of reflecting two different shocks to consumption in period $t+1$, each of which is arguably a response to a different mechanism. Just to make things concrete, suppose it is a belief in a lucky store that is driving the initial sales increase at the winning store; then identification of addiction requires it to be the case that we are looking at sufficiently many periods for the lucky store belief to have worn off. Suppose it is a local market advertising effect driving the

[7] This conclusion is not dependent on the specific form of addiction described in the model above; it follows even if utility were a function of the full history of past consumption.

initial sales increase at the zip code level; then we need it to be the case that that this level of advertising has abated. We return to this point when we discuss our results.

A final note on our Instrumental Variables (IV) approach is that it is interesting to consider who the "treated" group is, in the sense of the Local Average Treatment Effect (Angrist and Imbens, 1996). In our previous work (Guryan and Kearney, 2008), we show that the increase in sales following the sale of a winning ticket is largest among zip codes with economically disadvantaged populations, including high school drop-outs, the elderly, and people living in poverty. It is likely the case that this contrasts to the Farrell et al. study of UK lottery gambling, described above, which uses rollover-induced variation in prices. Work by Oster (2004) has demonstrated that in the context of the Connecticut state lottery, rollovers induce higher sales among higher-income players. We thus suspect that the Farrell et al. IV approach is identifying effects off a relatively more economically advantaged population of UK lottery players as compared to our sample of Texas State Lottery players.

III. Data and Background on the Texas State Lottery

We have compiled a dataset that includes weekly store-level sales of lottery tickets by game, the location and jackpot size of winning tickets in three lotto games, and zip-code-level demographics for each lottery retailer. The data span the period from January 2000 to June 2002 and cover every lottery retailer active in the state of Texas during the period under study. Weekly counts of store-level sales by game were obtained through Open Records agreements with the Texas Lottery Commission. During the sample period, there are 24,400 active lottery retailers in Texas spread across 1,386 cities and 3,660 nine-digit zip codes. On average, there are 827 retailers per city and 30 per zip code.

We analyze the effects of the sale of winning tickets on Lotto Texas, which is the largest of Texas's lotto games. Lotto Texas offers multi-million dollar jackpots, which winners can choose to receive either as 25 annual payments or as one (present-discounted) cash payment. From its inception in 1992 until it was changed in mid-July 2000, Lotto Texas was played by choosing six numbers out of a field of 50, yielding odds of 15,890,700 to one of matching all six numbers. The field was later expanded to 54, yielding odds of 25,827,165 to one. The prize pool for each Lotto Texas drawing is comprised of 55 percent of sales for that drawing. Of this amount, 68 percent is allocated to the jackpot prize, plus any amount carried over from previous drawings. If no ticket bet matches the winning six numbers in a given week, the amount allocated for a top-prize winner is rolled over to the next draw. Portions of the prize pool are reserved each week to pay pari-mutuel prizes for five-of-six and four-of-six winners. A fixed prize of five dollars is paid to players who match three of the six numbers. Lotto Texas drawings occur twice a week. Players can purchase bets up to 10 drawings in advance, paying one dollar per drawings. Lotto Texas jackpots in our sample range from $1.03 million to $51 million. The Texas lottery brings in roughly $2.8 billion in sales annually. Over the period we observe, lottery retailers in Texas averaged weekly sales of $2,576. Retailers averaged weekly sales of $733 on Lotto Texas.

We link sales data to information about where and when winning tickets are sold. We observe the week in which a top Lotto Texas prize is won, the amount of the top prize, the retailer that sold the winning ticket, and the zip code and city in which the retailer is located. Our sample includes 68 winning Lotto Texas jackpots. The location of the vendors selling these winning tickets is publicly available and posted on the Texas Lottery's website.

The detailed sales data linked to information on where winning tickets were sold allow us to track sales at stores after they sell tickets that win large jackpots. Because we know the address of stores, we are able to track the sales at nearby stores to test whether there is spatial substitution by consumers.

IV. Short-Run Sales Effects

In this section, we present evidence on the first stage of the two-stage test of addiction – the immediate effects of the sale of a winning lottery ticket on lottery ticket sales. The short-run evidence reviewed in this section is also contained in Guryan and Kearney (2008). We review it again here as it is crucial to the validity of our IV strategy to demonstrate that sale of a winning ticket in fact causes a shock to local gambling consumption.

A. Short-run effects on store level sales

Table 1 presents the one-week results from OLS estimation of equation (2). Each entry in the table is the estimated effect of selling a winning ticket from a separate regression. All regressions control for contemporaneous (i.e. measured the same week the winning ticket was sold) log sales, and week effects. Column (1) reports the results for estimating equation (2) with the dependent variable defined as the log of store-level sales of Lotto Texas tickets. These estimates suggest that same-store sales increase by 32 log points the week after a store sells a winning ticket. The estimated sales increase is statistically significant.

Our model facilitates a direct test of the identifying assumption that stores are randomly selected conditional on contemporaneous sales. If this identifying assumption is satisfied, sales in week *t-k* should not predict whether a store sells a winner at a future week *t*, controlling for sales in week *t*. In other words, contemporaneous sales should be a sufficient statistic for the probability that a store sells a winner. A direct test of this prediction is to estimate equation (3) for negative values of *k*. These are the *lead* versions of the *lags* specifications, and the results are presented in tables 1 and 2 along with the one period lag estimates. Consistent with the prediction of the identifying assumption, none of the lead estimates are significantly different from zero.

The increase implied by the 0.320 log point effect on same-store sales of Lotto Texas tickets the week following a winning ticket sale is large in both economic and statistical terms. The average store sells an additional 276 tickets the week after selling a winning Lotto Texas ticket. This increase is about 11 percent of the week-to-week standard deviation in total retail-level lottery ticket sales; about 38 percent of the week-to-week standard deviation in retail-level Lotto Texas ticket sales; and 50.2 percent of the week-to-week standard deviation in the change in retail-level Lotto Texas ticket sales.[8]

As we described above, in order to test for addiction, the sale of a winning ticket must cause an increase in consumption of lottery tickets. The increase in game-specific sales after the sale of a winner could reflect an aggregate increase in sales at the winning store, or it could be completely offset by declines in sales of other games at the store. To

[8] Our previous paper (Guryan and Kearney, 2008) shows that sales response increases with the proportion of the zip code population with low levels of education, living in poverty, and above the age of 64.

look directly at aggregate retailer sales, we modify equation (2) to define the dependent variable as the log of total retailer lottery sales, including the sales on smaller-jackpot lotto games, daily numbers games and scratch tickets. The regressor of interest is still a binary indicator for whether the store sold a winning Lotto Texas ticket, controlling for lagged sales. As reported in column 2, total retailer lottery sales are 18.4 log points higher the week after a store sells a winning Lotto Texas ticket. The increase is smaller in magnitude than the game-specific effect, suggesting that the initial response is concentrated in sales of the specific game for which a winner was sold. Importantly, there is a significant increase in lottery consumption at the winning store net of any substitution across games.

Another dimension of potential substitution is shifting across lottery vendors. If the increase in lottery sales at the winning store is driven entirely by a decrease in sales at nearby stores, then there is no shock to lottery consumption, but merely a shifting of the purchase location. To investigate this possibility, we compare the estimates for the winning store to estimates of the effect of the winning ticket sale for non-winning stores in the same zip-code as winners. To do this, we estimate a version of equation (2) with three changes: (a) we restrict the sample to stores that did not sell a winning ticket in week t, (b) the regressor of interest is an indicator for whether any store in the zip-code sold a winning Texas Lotto ticket in week t, and (c) we include controls sales at both the zip-code and store level in week t. The latter change is necessary because the sample is restricted to non-winning stores. However, conditioning on store-level sales in week t perfectly controls for selection because selling a winning ticket is random conditional on this variable.

The results, presented in column 3 of table 1 show that non-winning stores in winning zip-codes experience significant positive increases in sales the week after the winning ticket was sold nearby. There is no net substitution away from nearby stores to explain the increase in sales at the winning store, suggesting that there is indeed an aggregate increase in total consumption of lottery tickets. Furthermore, the increase in sales at stores near the winning store is smaller than at the winning store but significantly positive. Below, we take advantage of the fact that this is evident of a second randomly assigned shock to lottery demand. The shock to nearby stores is different in a number of ways from the shock to the winning store. It is smaller in magnitude, it affects a different set of people, and it is driven by different motivations. In particular, we can be sure that the sales increase at nearby stores is not driven by the lucky store effect; the nearby stores did not win so they cannot be lucky.[9] Instead, this response is likely driven by increased attention to the lottery, akin to advertising, or a revised subjective estimate of the likelihood of an individual winning the lottery. Below, we use both of these demand shocks to generate two independent series of IV estimates of addiction. The comparability of the two series of estimates addresses some of the worries about potential biases described in the previous section.

B. Short-run effects on sales in the winning zip code

We next estimate the first stage regression at the level of the 9-digit zip code, which we use as a measure (albeit an imperfect one) of the local market. The regressor of interest is an indicator for whether any retail outlet in zip code *z* sold a winning ticket in

---

[9] See Guryan and Kearney (2008) for more evidence supporting the conclusion that the additional sales at the winning store in the week following the winner are driven by a mistaken belief that the winning store is

week *t*. We define the dependent variable as the log of total Lotto Texas ticket sales in the zip code in week *t+1*. The estimating equation is the same as above, except that all variables are aggregated to the zip-code level rather than the retailer level.

Table 2 presents the results for zip code level sales. We begin by looking at all zip codes. The estimates indicate that zip codes with Lotto Texas winners experience a 13.2 log point increase (standard error of 1.9 log points), or 14.1 percent, in Lotto Texas sales. All of the lead estimates are insignificant, consistent with the fact that the location of the winner is conditionally random.

Zip-codes are not perfect measures of lottery ticket markets. Zip-codes differ both in geographic size and in number of lottery retailers. Even if there were no geographic substitution of sales, if the sales response is concentrated at the winning store we might expect the effect of the winning ticket to be diluted as the number of stores in the zip code gets large. We present estimates where we break zip codes into groups based on the number of lottery retailers. The aggregate sales response is predictably largest in zip codes with fewer stores.

We conclude from the results presented in this section that the sale of a winning ticket leads to an aggregate increase in lottery sales. It is thus a valid candidate for an instrument for lottery gambling consumption.

V. Evidence on persistence

Having established that the sale of a winning ticket causes a shock to lottery ticket demand a week after the sale of the winner, we now turn to a test for whether that shock has persistent effects on subsequent lottery ticket demand. To increase the precision of

---

temporarily lucky, rather than say an advertising effect.

the test, from here on we aggregate all estimates to the month level, rather than the week level. Our test follows directly from the economic definition of addiction, that a good is addictive if current consumption causally increases future consumption by increasing the marginal utility of future consumption. That definition implies that a test for addiction is whether there is a causal relationship between current and future consumption.

Recall that the relationship of interest is equation (1) above. We are interested in the relationship between lottery sales in month *t+1* and lottery sales in month *t+k* . These causal effects of current demand on future demand are summarized by a series of parameters, $\pi(k-1)$. Here we estimate this series as *k* runs from 2 to 18 months. Lottery tickets are defined as addictive if $\pi(k-1)>0$ for some $k>2$. Figure 1 plots the reduced form estimates. The figure shows a fairly steep initial decline, followed by a relative steady level of heightened consumption. As can be seen in the figure, between 6 and 12 months after the sale of a winning ticket, zip code sales are elevated by approximately five log points, compared to 10 log points in the first month following the winner. The point estimate at 18 months is similar, though the confidence interval around the estimate is widened.

We implement an IV approach that instruments for zip-code-level sales in month *t+1* with a dummy for whether any store in the zip-code sold a winning ticket in month *t*. As described above, the first stage of this two-stage least squares test is therefore the month-level version of equation (2) above. We estimate the effect of a shock to lottery consumption at various horizons by estimating $\pi(k-1)$ for different values of *k* starting with 2. In each case, the sale of a winning ticket in the zip-code in month *t* is used as an instrument for lottery sales in the zip-code in month *t+1*, after conditioning on lottery sales in the zip-code in month *t*. For any length *k*, $\pi(k-1)$ is then estimated separately using IV in which the dependent variable is lottery sales in the zip-code in month *t*+*k*. These IV estimates are almost exactly the ratios of the reduced form estimate of $\gamma(k)$ at the month level to the first-stage estimate of $\gamma(1)$. The only difference is that the IV is estimated on the smaller sample that has observations of sales both one and *k* months after month *t* (the point estimates of the IV are almost identical to these ratios).

These zip-code IV persistence estimates along with the corresponding reduced form estimates are presented in Table 3. The full series of IV persistence estimates one to 17 months after the demand shock (i.e. for *k*=2 to 18) are shown in Figure 2. The top entry in the left column of table 3 shows the first-stage estimate for the month-level regression. This estimate implies that sales of Texas Lotto tickets are 0.11 log points higher in the zip-code where a winning Texas Lotto ticket was sold in the month following the sale of the winner. The reduced form estimate in the second row in the left column implies that two months after the sale of the winner, sales in the winning zip-code are elevated 0.070 log points relative to non-winning zip-codes. Comparing the estimated 0.070 to 0.110, we tentatively conclude that the initial shock exhibits a persistence rate of 0.636 (0.070/0.110) after one month. Five months later, roughly half of the initial shock remains (0.054 of 0.110). And 17 months later, 40 percent remains.

The IV estimates are shown in the right column. The first entry is blank because the first month following the sale of the winner is when the shock to demand occurs. The *k=2* estimate implies that one month after the exogenous shock to lottery demand, 0.639 of the initial shock remains, which is what the ratio of the reduced form estimates indicated. The estimate is statistically distinguishable from zero. Under the identifying assumptions, we can attribute all of this increase to a causal result of increased lottery ticket consumption one month earlier. After 5 months (*k=6*), nearly half of the initial shock to demand persists. The point estimate is 0.492 with a standard error of 0.157. After 11 months the point estimate actually rises to 0.635, and by 17 months after the shock, close to 40 percent of the initial sales response remains.

The pattern of persistence can be seen clearly in Figure 2. The persistence fades gradually over a year and a half. Many of the point estimates at longer lags are not individually significant and exhibit somewhat more variation than the shorter lag estimates. This is directly related to the decline in sample size as the lags get longer. There are simply not very many winning stores for which we have data 18 months subsequent.

As discussed above, the primary threat to the validity of these IV estimates is the possibility that the sale of the winning ticket has a direct effect on sales long after the winner was announced (specifically more than one month). The estimates above treat the unit of analysis as the zip code to ensure that the initial shock to demand is representative of a real increase in lottery ticket consumption rather than a shift in the location of lottery purchases. However, looking at the data this way combines the experiences of two very different types of stores: the winning stores themselves and other stores in the same zip code that did not sell a winning ticket. It is possible to examine the experiences of these stores separately.

Looking at these two sets of stores separately is useful because the reason for the initial sales response – the exogenous shock to demand – is different. In the case of the winning store, we have argued elsewhere that the initial response is due to an erroneous belief that the store is lucky. In the case of nearby stores, there can be no such belief, and the sales response is likely a combination of general advertising effects and a local realization that real people win the lottery. Importantly, for the purposes of this study, the shock is due to a different phenomenon. Furthermore, the shock to the non-winning stores in the winning zip-code is smaller in magnitude. A test of the identifying assumption is therefore whether this separate shock produces estimates of persistence that are similar to those produced by the shock induced at the winning store itself. Since we estimate the same parameter with two different instruments, this test can be thought of as an overidentification test on the vector of IV estimates.

To estimate the effects of these separate shocks we estimate two versions of the specifications shown in Table 3. In both cases, the dependent variable is the natural log of *store-level* Lotto Texas sales in month $t+k$. In the first version, we estimate the effect of a store selling a winning Lotto Texas ticket in month t on the log of the winning store's Lotto Texas ticket sales in week $t+k$. The first stage is almost the same as the regression presented in the first column of table 1. The only difference is that the estimates are aggregated to the zip-code level. As before, we must control for the natural log of store-level sales in month t to ensure that the winning ticket dummy is conditionally random. The IV estimates use the dummy for whether the store sold a winning Texas Lotto ticket in week t to instrument for the log of store-level sales in month $t+1$.

To estimate the effect of the shock experienced by non-winning stores in winning zip codes, we restrict the sample to stores that did not sell a winning ticket in month t. We estimate the effect of a winning ticket being sold in the store's zip-code in month t on store-level sales in month $t+k$. We control for the natural log of both zip-code level *and* store-level sales in month *t*. It is necessary to control for zip-code level sales in month *t* because the more tickets are sold in the zip-code, the more likely it is that a winner is sold there in that month. It is necessary to control for store-level sales in month *t* because we restrict the sample to non-winning stores. Because selling a winning ticket in month *t* is random conditional on the store's sales in month *t*, conditioning on this variable corrects perfectly for selection.

The left side of Table 4 shows the reduced form estimates and the right side shows the IV estimates. One month after selling a winning ticket, the winning store experiences increased sales of 0.225 log points. The non-winning stores in the same zip-code as winners experience a 0.066 log point increase. The IV estimates show what fraction of each of these exogenous demand shocks remain months later. Strikingly, the IV estimates using these two different shocks to demand yield very similar estimates. One month after the shock ($k=2$), the winning store retains 0.716 of its initial shock to lottery demand. After the same month has passed, the non-winning stores in the winner's zip-code retain 0.665 of their initial shock. After two months ($k=3$), the estimates are 0.585 and 0.502. The similarities in the patterns of these two series of estimates can be seen clearly in Figure 3.

The comparability of the two IV estimates gives us confidence that we are identifying a true causal effect of past consumption on future consumption. These estimates suggest that a shock to lottery consumption reaches its half-life at roughly between three and five months. In other words, about 50 percent of the initial shock to lottery consumption remains after six months. Over the course of the following 12 months, that level of persistence does not measurably change.

Though the estimates are not statistically distinguishable, it does appear that six months after the winning ticket sale and beyond, the shock to the non-winning stores in winning zip codes is slightly less persistent than in the winning stores. If this difference is real rather than a result of sampling variance, it suggests that the lucky store effect might itself persist at low levels in the winning stores. If this is the case, we would take the non-winning store persistence estimates as estimates of addiction. Similarly, an alternative reading of the pattern of estimates is that the somewhat rapid initial decay in the persistence estimates could be considered the diminishment of the psychological phenomenon driving the initial shock – what we have argued in Guryan and Kearney (2008) to be a belief in a "lucky store". Among non-winning retailers in the zip code level, it could be the tapering off of the effect coming from the general attention and advertising of the local winning ticket. Persistence at the longer intervals would still be interpretable as a measure of addiction.

VI. Conclusion

In this paper we present a test for the addictiveness of lottery gambling. We exploit a random shock to local market consumption of lottery gambling coming from the sale of a winning ticket. We use the sale of a winning ticket, the location of which is random conditional on sales, as an instrument for present consumption and test for a causal relationship between present and future consumption. This definition of addiction accords with the definition commonly used in the economics literature. We argue that under reasonable assumptions, persistence of the initial sales shock observed at longer intervals of time is evidence of addiction.

Data from the full set of Texas Lottery retailers over the years 2000 to 2002 provide clear evidence that the sale of a winning ticket leads to an initial increase in ticket sales at both the store and local market level. Ticket sales in the winning zip code increase by 11 log points in the month following the sale of a winner. This increase is driven by a large localized increase in sales at the winning store (0.225 log points) and a corresponding positive spillover to non-winning stores in the zip code (0.066 log points). Instrumental variables estimates of the persistence of these initial shocks suggest that around half of the initial increase remains after 6 months. The estimated degree of persistence (or addiction) after 12 and 18 months are 0.635 (standard error of 0.207) and 0.394 (standard error of .235) respectively.

This finding has clear implications for a normative evaluation of lottery gambling. Our findings suggest that outside influences on lottery gambling could potentially have long-lasting effects on the level of gambling; that influence could take the form of a winning ticket, but more generally, of an advertising campaign or a new game or lottery initiative. We are careful to define addiction as it has been treated traditionally in the economics literature. However, it should be noted that given the data available we have no way of distinguishing the mechanism behind the observed persistence. It could be driven for example by clinical addiction or simply learning-by-doing and the welfare implications clearly vary. Furthermore, as with previous economic tests of addiction, we are unable to document directly a change in marginal utilities. For example, a shock to past consumption might expand a boundedly rational consumer's consideration set to include lottery tickets, thereby increasing the likelihood of future consumption.

An additional important caveat to the interpretation of our results is that we are unable to explore heterogeneity across lottery consumers. It is certainly true that there is great variation in the degree of addictiveness among lottery gamblers. Unfortunately our empirical study relies on aggregate ticket sales, so we must make a representative agent assumption and cannot characterize our results in terms of varied responses across consumers.

A consideration of the welfare implications of our findings would require an estimate of any harm associated with addictive lottery gambling. If consumers who are enticed to buy lottery tickets by the sale of a winner then maintain that level of consumption because of a myopic addiction, then any resulting internal harm would imply a lower level of the optimal provision of this state-provided good, all else equal. If they are rational addicts in a Becker-Murphy sense, optimal provision and pricing would depend only on the external harm imposed, for example, on family members from the displaced consumption of other household goods (see Kearney, 2005). Furthermore, in light of previous research demonstrating that consumers do not fully understand the random processes underlying the good (Guryan and Kearney, 2008; Clotfelter and Cook, 1993), there could be an interaction between misperceptions and addiction that would have further implications for the optimal provision of state lottery products.

References

Becker, Gary S., Michael Grossman, and Kevin M. Murphy (1994). "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84(3): 396-418.

Becker, Gary S. and Kevin M. Murphy (1988). "A Theory of Rational Addiction," *Journal of Political Economy*, 96(4): 675-700.

Becker, Gary S. and Kevin M. Murphy (1993). "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics*, 108(4): 941-964.

Boyer, Marcel. (1978). "A Habit Forming Optimal Growth Model," *International Economic Review*, 19(3): 585-609.

Clotfelter, Charles T. and Philip J. Cook (1993). "The 'Gambler's Fallacy' in Lottery Play," *Management Science* 39 No. 12 (December): 1521-1525.

Farrell, Lisa, Edgar Morganroth and Ian Walker (1999). "A Time Series Analysis of U.K. Lottery Sales: Long and Short Run Price Elasticities," *Oxford Bulletin of Economics and Statistics*, 61(4): 513-526.

Gruber, Jon and Botond Koszegi (2001). "Is Addiction 'Rational'? Theory and Evidence," *Quarterly Journal of Economics*, 116(4): 1261-1303.

Guryan, Jonathan and Melissa S. Kearney (2008). "Gambling at Lucky Stores: Evidence from State Lottery Sales," *American Economic Review* 98 (1), March p 458-73.

Hicks, John. (1965). *Capital and Growth* (New York: Oxford University Press).

Iannacone, Laurence R. (1986). "Addiction and Satiation," *Economics Letters*, 21: 95-99.

Kearney, Melissa S. (2005). "State Lotteries and Consumer Behavior" *Journal of Public Economics* 89: 2269-2299.

Mobilia, Pamela (1993). "Gambling as a Rational Addiction," *Journal of Gambling Studies* 9(2): 121-151.

National Gambling Impact Study Commission (1999). *Final Report*. Washington, D.C.

Oster, Emily (2004). "Are All Lotteries Regressive? Evidence from the Powerball." *National Tax Journal* 57: 179-87.

Pollack, Robert A. (1970). "Habit Formation and Dynamic Demand Functions," *Journal of Political Economy*, 78(4): 745-763.

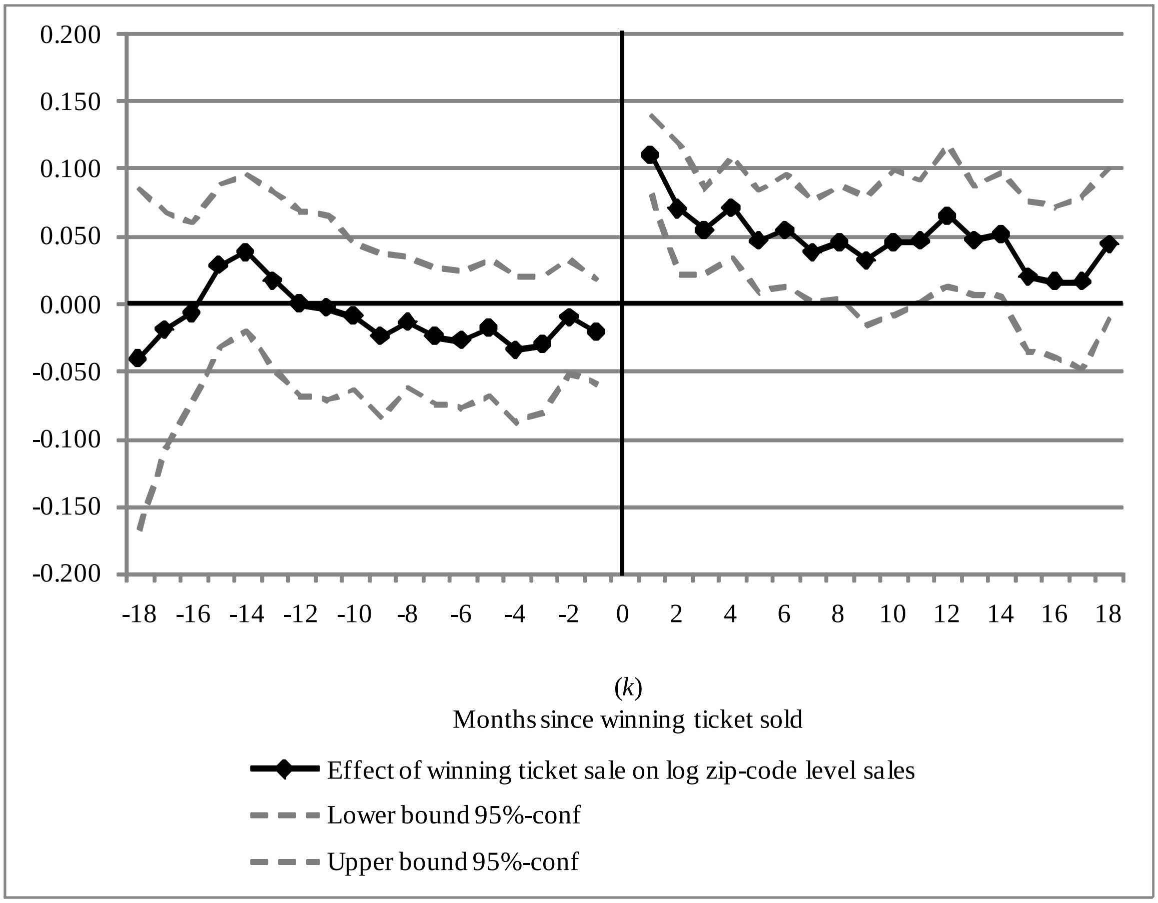
Ryder, Harl E. and Geoffrey M. Heal (1973). “Optimum Growth with Intertemporally Dependent Preferences,” *Review of Economic Studies*, 40(1): 1-33.

Stigler, George J. and Gary S. Becker (1977). “De Gustibus Non Est Disputandum,” *American Economic Review*, 67(1): 76-90.

Terrell, Dek (1994). “A Test of the Gambler’s Fallacy – Evidence from Pari-Mutuel Games,” *Journal of Risk and Uncertainty* 8 No. 3: 309-317.

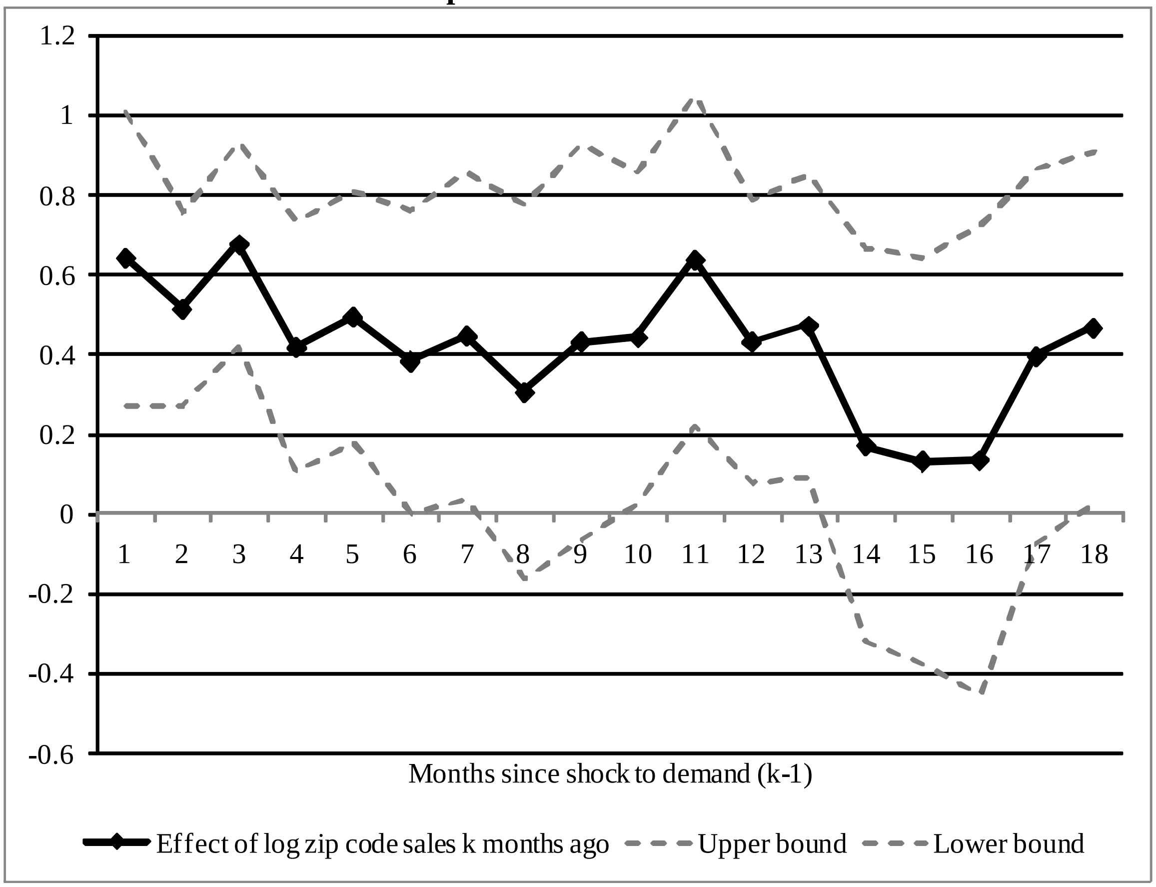
Walker, Ian and Juliet Young (2001). "An Economist's Guide to Lottery Design." *The Economic Journal*, 111 (November): F700-F722.

Figure 1:
Effect of a winning ticket sale on zip-code-level Texas Lotto tickets sales



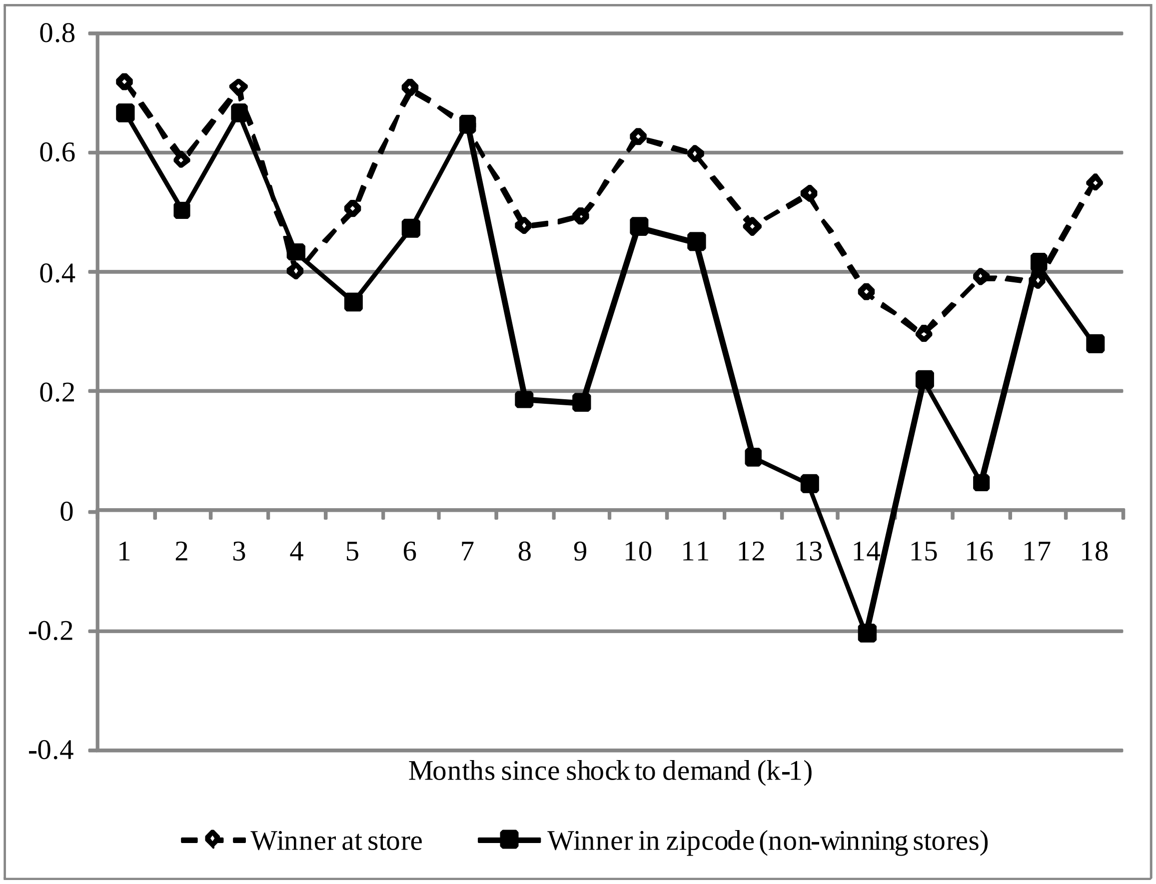
Note: The series of black solid diamonds plots the point estimates from 48 regressions, each estimating the effect of the sale of a winning Texas Lotto ticket in the zip code in month *t* on the natural log of sales of Texas Lotto tickets in the zip code in month *t+k*. All regressions control for the natural log of Texas Lotto sales in month *t*, and on month fixed effects. The dashed lines represent the upper and lower bounds of the 95-percent confidence intervals of the point estimates.

Figure 2:
Instrumental variables estimates of the effect of Texas Lotto sales in the zip code on future Texas Lotto sales in the zip code



Note: The series of black solid diamonds plots the IV estimates of the effect of the log of zip-code-level Texas Lotto sales in month *t+1* on the log of zip-code-level Texas Lotto sales in month *t+k*. The instrument for sales in month *t+1* is a whether a winning Texas Lotto ticket was sold in the zip-code in month *t*. All regressions control for the log of Texas Lotto sales in month *t* and month dummies. The dashed lines represent the upper and lower bounds of the 95-percent confidence intervals of the point estimates.

Figure 3:
Instrumental variables estimates at the retailer level from two different shocks to Texas Lotto demand



Note: Both series plot IV estimates similar to those shown in Figure 2, but where the dependent variable is the log of Texas Lotto sales at the retailer level, rather than the zip-code level. The instrument for the series of open diamonds is whether the retailer sold a winning Texas Lotto ticket in month *t*. The instrument for the series with the solid squares is whether there was a winning ticket sold in the zip-code in month *t*. The regressions that produce the solid squares estimates are restricted to retailers that did not sell a winning ticket in month *t*, and control for the log of retailer sales of Lotto Texas tickets in month *t*. Regressions in both series also control for the log of zip-code level sales of Lotto Texas in month *t* and for month effects.

Table 1:
Impact of a winning Lotto Texas ticket on store level sales: one-week effect

Dependent Variable:	Ln(retailer weekly Lotto Texas ticket sales) (1)	Ln(retailer weekly total lottery ticket sales) (2)	Ln(retailer weekly Lotto Texas ticket sales) (3)
	Winner at store		*Winner in zip code (non-winning stores)*
1 week after	.320	.184	.051
	(.033)	(.054)	(.010)
1 week lead	-.052	-.027	-.021
	(.027)	(.062)	(.010)
2 week lead	-.041	-.137	-.008
	(.028)	(.080)	(.011)
3 week lead	-.047	-.041	.002
	(.031)	(.059)	(.011)

Note: Each cell reports the OLS regression coefficient on a binary indicator for a Lotto Texas jackpot winner. In columns (1) and (2) the indicator is for a winner at the store. In column (3), the specification includes separate indicators for a winner at the store and a winner at another retailer in the zip code. The regression model estimates $\ln(\text{sales})_{jt}$ as a function of $\ln(\text{sales})_{i(t-k)}$ and win $_{i(t-k)}$, where *i* indexes retailers, *t* indexes weeks, and *k* indexes weeks lag or lead. Each row reports the results for a different *k*. All regressions control for week fixed effects. Standard errors are robust standard errors. The initial sample of retailers consists of 2,031,395 observations. The sample for non-winning stores is limited to retailers in zip codes with more than one retailer; there are 1,387,413 such observations.

Table 2:
Impact of a winning Lotto Texas ticket on zip-code sales: one-week effect

Dependent Variable:	ln(zip-code weekly Lotto Texas sales)			
	All zip-codes	Zip-codes with 1-5 retailers	Zip-codes with 6-10 retailers	Zip-codes with >10 retailers
1 week after	.132	.375	.196	.056
	(.019)	(.050)	(.066)	(.014)
	n=396,929	*n=287,508*	*n=29,718*	*n=79,883*
1 week lead	-.017	-.157	.046	-.012
	(.019)	(.081)	(.012)	(.010)
2 week lead	.004	-.090	.012	-.002
	(.016)	(.065)	(.032)	(.010)
3 week lead	-.004	-.125	.034	-.006
	(.022)	(.101)	(.018)	(.009)

Note: Each cell corresponds to a unique regression estimating $\ln(\text{sales})_{zt}$ as a function of $\ln(\text{sales})_{z(t-k)}$ and $\text{win}_{z(t-k)}$, where *z* indexes zip codes, *t* indexes weeks, and *k* indexes weeks lag or lead. All regressions control for week fixed effects. Standard errors are robust standard errors.

Table 3:
The Persistence of the Effect of a Winning Ticket on Zip-code Sales of Lotto Texas

Dependent variable ln(zip-code monthly sales of Lotto Texas)

	(1)	(2)
k	*Reduced Form:* γ_k^z	*IV:* π_k^z
1 month lag	0.110	---
	(0.015)	
2 month lag	0.070	0.639
	(0.024)	(0.184)
3 month lag	0.054	0.512
	(0.016)	(0.120)
6 month lag	0.054	0.492
	(0.021)	(0.157)
12 month lag	0.065	0.635
	(0.026)	(0.207)
18 month lag	0.044	0.394
	(0.028)	(0.235)

Note: The first column shows reduced-form estimates of the effect of a zip-code selling a winning Texas Lotto ticket in month t on the log of zip-code-level sales of Texas Lotto tickets in month $t+k$. Regressions control for the log of zip-code-level Texas Lotto sales in month t and month fixed effects. The second column shows IV estimates of the effect of the log of zip-code-level Texas Lotto sales in month $t+1$ on the log of zip-code-level Texas Lotto sales in month $t+k$ (i.e. $k-1$ months later). The instrument is selling a winning Lotto Texas ticket in the zip-code in month t, and the controls are the same as the estimates in the first column.

Table 4:
The Persistence of the Effect of a Winning Ticket on Retailer Sales, Separately by Winning Store and Non-Winning Stores

	Dependent Variable: ln(retailer monthly Lotto Texas sales)			
	(1)	(2)	(3)	(4)
	Reduced Form:		*IV:*	
	γ_k^i	$\gamma_k^{z'}$	π_k^i	$\pi_k^{z'}$
k	winner at store	winner in zipcode, non-winning stores	winner at store	winner in zipcode, non-winning stores
1 month lag	0.225	0.066	---	---
	(0.036)	(0.010)		
2 month lag	0.169	0.046	0.716	0.665
	(0.045)	(0.013)	(0.145)	(0.164)
3 month lag	0.133	0.031	0.585	0.502
	(0.030)	(0.010)	(0.107)	(0.132)
6 month lag	0.126	0.022	0.504	0.349
	(0.041)	(0.010)	(0.137)	(0.143)
12 month lag	0.159	0.032	0.596	0.449
	(0.047)	(0.013)	(0.153)	(0.168)
18 month lag	0.088	0.032	0.384	0.414
	(0.053)	(0.017)	(0.199)	(0.202)

Note: The first column shows reduced-form estimates of the effect of a store selling a winning Texas Lotto ticket in month *t* on the log of retailer sales of Texas Lotto tickets in month *t+k*. Regressions control for the log of store-level Texas Lotto sales in month *t* and month fixed effects. The second column shows reduced-form estimates of the effect of a winning Texas Lotto ticket being sold in the same zip-code in month *t* on the log of retailer sales of Texas Lotto tickets in month *t+k*, for the sample of stores that did not themselves sell a winning Texas Lotto ticket in month *t*. Regressions control for the log of store-level Texas Lotto sales in month *t*, the log of zip-code level sales in month *t*, and month fixed effects. The third and fourth column show IV estimates of the effect of the log of store-level Texas Lotto sales in month *t+1* on the log of store-level Texas Lotto sales in month *t+k* (i.e. *k-1* months later). In column 3, the instrument is selling a winning Lotto Texas ticket in month *t*, and the controls are the same as the estimates in column 1. In column 4, the instrument is a winning Texas Lotto ticket being sold in the zip-code in month *t*. The estimates in column 4 restrict the sample to stores that did not sell a winning ticket in month *t*, and include the same controls as the estimates in column 2.