An Equilibrium Model of Rollover Lotteries

Giovanni Compiani Lorenzo Magnolfi Lones Smith*

March 2024

Abstract

In a rollover lottery, buyers pick their own numbers, and a jackpot not won rolls over to the next draw. Since they are a major source of government revenue globally, we develop an equilibrium model of this lottery to shed light on its design. Buyers care about the lottery enjoyment level and expected winnings, and the market-clearing price is not the ticket price but the expected monetary loss on a lottery ticket. The supply curve captures the relation between tickets sold and expected loss implied by the rules of the game. We use this equilibrium model in two empirical applications. First, we test the model's predictions on the optimal relationship between odds and population size using data from many countries, and across U.S. states. Second, we propose a structural empirical implementation of the model and nonparametrically estimate demand for U.S. national rollover lotteries by exploiting the randomness inherent in the rollover mechanism. We find that the model predicts well out of sample and show how to use it to inform lottery design.

*Compiani: University of Chicago - Booth School of Business, giovanni.compiani@chicagobooth.edu; Magnolfi: Department of Economics, University of Wisconsin, Madison, WI 53706, magnolfi@wisc.edu; Smith: Department of Economics, University of Wisconsin, Madison, WI 53706, lones.smith@wisc.edu. We thank seminar and conference participants at U Chicago, the 2022 Tinos IO Conference, UVA, Conference on Models and Econometrics of Strategic Interactions (Vanderbilt), and at the Frontiers in Empirical Industrial Organization Workshop (Mannheim).

1 Introduction

Lotteries have a long history across the world and are very popular by many measures. For example, American consumers spent almost $100 billion on state-run lotteries in 2022. These revenues exceeded sales of other forms of entertainment (such as video streaming services, concert tickets, books, and movie tickets) combined. The national reach of lotteries is tremendous — half of U.S. adults play the lottery at least once a year, and one-eighth play at least once a week (Cohen, 2022). The popularity of lotteries makes them a significant source of revenue for governments, generating $31 billion in revenue for state and local governments in the U.S.in 2021.[1]

Given the sheer size of this market, and its contribution to government revenue, lottery design questions are important. Most significant among formats is the Genoese *rollover lottery*, in which (a) ticket buyers pick their own number combinations, (b) all tickets matching a randomly drawn winning combination win an equal share of a jackpot prize, and (c) if no one wins, the jackpot rolls over to the next draw. The lottery chooses the odds (the number of possible combinations) as well as the share of ticket sales that rollover to the next jackpot. Such lottery changes may dramatically impact ticket sales and government revenues. For instance, after Powerball lengthened the odds in October 2017, average per-draw revenues increased from $16.8 million to over $26 million.[2] How can researchers make predictions on alternative lottery designs?

In this paper, we propose a novel equilibrium model of the lottery market based on supply and demand. We assume that the good purchased is lottery thrill. One buys a lottery ticket if the expected monetary loss — the ticket price minus the expected winnings — is less than the gambling thrill. Potential ticket buyers vary by thrill, and this heterogeneity generates an aggregate demand curve. In our novelty, the market-clearing price in a lottery equilibrium is *not the stated ticket price but instead the expected monetary loss on a lottery ticket* — since that is formally the cost of experiencing the lottery thrill. Next, lottery rules mechanically imply an (inverse) supply function that maps ticket sales to the expected monetary loss, conditional on the jackpot inherited from the previous draw. Of course, the new jackpot incorporates a fraction of the current draw's ticket sales. The intersection of lottery supply and rationally forecast demand determines the new ticket sales and expected loss. For buyers to compute the expected loss from buying a ticket, it must be that they form an expectation of the overall number of tickets sold. We adopt the standard

[1]Source: https://www.statista.com/statistics/249128/us-state-and-local-lottery-revenue/.

[2]This change will be further discussed in Section 5.5.

assumption that their expectations are rational.

A range of comparative statics analyses immediately follow, illuminating the economics of rollover lotteries. For example, each rollover improves the expected payoffs of lottery ticket buyers — i.e., it lowers the inverse supply curve of losses, acting as a subsidy — and therefore induces more people to play the lottery. From a revenue perspective, higher inherited jackpots have a cost — as the jackpot may be won and paid out —but also a benefit due to increased sales. For elastic demand (which corresponds to what we find in our application), we show that higher rollovers have a positive effect on revenues up to very large levels of jackpot, thus helping explain the success of rollover lotteries. Besides providing insights into the existing lotteries, our modeling framework allows us to extrapolate out of sample, and thus evaluate counterfactual lottery designs.

Our model makes predictions of ticket sales, individual losses, and government revenues, and skips a comprehensive model of individual decision-making underlying the demand curve. We take no stand on how many tickets anyone purchases, for instance. Instead, the demand curve is our basic element. For we estimate our model with the type of data commonly available to lottery authorities, i.e., market-level data on ticket sales and jackpots. This simple framework explains the data and performs well in predicting revenues both in and out of sample. Second, our model does not speak to a range of normative questions, such as the regressivity of lotteries — again, because we are silent on the underlying consumer optimization. Nevertheless, our positive framework complements normative approaches. For instance, we can quantify how much revenue the government must give up if it chooses lottery parameters according to any candidate welfare-maximizing policy. Alternatively, given socially optimal levels of expected loss or jackpots (e.g., Lockwood et al., 2021), our model can be used to devise lottery designs that achieve these outcomes. This requires a notion of market equilibrium, which is our main theoretical contribution.

We explore the model's empirical implications in two applications. First, we derive a relationship between lottery odds and market size. If lottery ticket demand scales in proportion to market size, we find that the revenue-optimal lottery has odds that scale linearly with the market size. To test this prediction, we construct two separate datasets: one on rollover lotteries across developed nations all over the world, and another on state-level lotteries across U.S. states. When using population as a proxy for market size, we find in both datasets that the model's predictions hold. In particular, across U.S. states — where we expect demand conditions to be quite similar — we find an R^2 coefficient above 0.75, suggesting that the model explains the relationship between odds and market size in the data very well.

We then consider a structural empirical implementation of the model, and nonparametrically estimate demand for the two U.S. national rollover lotteries: Mega Millions and Powerball. To identify their demands, we exploit the peculiar nature of the supply functions in this market. In a typical oligopoly with differentiated products, supply involves strategic interaction among firms; in contrast, in our case supply is entirely pinned down by lottery rules, and is thus deterministic. Furthermore, the rollover mechanism shifts the supply function: as the inherited jackpot increases, supply shifts down, implying lower equilibrium losses at any level of ticket sales. We exploit the randomness inherent in these shifts in supply to trace out the demand functions. This is analogous to the classic idea of using cost shocks to identify demand. While our theory provides a supply and demand framework to study the lottery market, it does not suggest or require a specific form of consumer utility. We thus use nonparametric methods (Compiani, 2022) to estimate the demand system as a flexible function of expected losses and thrill components. We find average own-loss elasticities in the range of 1.5–1.7, and little substitution across the two national lotteries, consistent with the existing literature. When we apply the estimates to predict equilibrium outcomes out of sample, we find that our approach fits the data well.

These two applications highlight how our modeling framework can either suggest stylized theoretical implications, or provide quantitative counterfactual predictions. In particular, one example of the latter use is to predict changes in sales and revenues corresponding to any changes in lottery design. In our framework, changes in lottery design affect the supply function and thus equilibrium outcomes. We illustrate this in our last section, by considering changes in the fraction of ticket sales accruing to the jackpot prize pool. We find that average per-draw revenues would increase for U.S. lotteries if the government increased this fraction above the status quo level.

Our paper relates to several research strands. A large literature investigates the public economics aspects of lotteries, including their potentially regressive nature (e.g., Clotfelter and Cook, 1987; Oster, 2004; Kearney, 2005), the addictive behavior they may generate (Guryan and Kearney, 2010), the competition that ensues when neighboring jurisdictions offer different lotteries (Knight and Schiff, 2012), and optimal lottery design (Lockwood et al., 2021). With respect to this literature, our work is complementary as it focuses on positive questions, characterizing the lottery market equilibrium and performing counterfactual analyses. Related to our findings on lottery odds across nations and U.S. states, Cook and Clotfelter (1993) are the first to note the empirical regularity that lottery odds tend to scale with population.

While they propose an explanation based on prospect theory, we show that this linear relationship is also predicted by a theory with rational, risk-neutral players.

Another important literature empirically investigates the individual determinants of risk preferences using either individual-level data on insurance choice (e.g., Cohen and Einav, 2007; Barseghyan et al., 2013) or aggregate data on betting (Aruoba and Kearney, 2011; Gandhi and Serrano-Padial, 2015; Chiappori et al., 2019). In contrast, our focus is not on risk preferences, but rather on building an equilibrium model to explain outcomes in lottery markets.

This paper is related to previous studies that estimate demand for lotteries. Gulley and Scott (1993) and Forrest et al. (2000b) estimate parametric models of demand for U.S. state lotteries and the UK national lottery, respectively. With respect to these existing studies, we adopt a nonparametric approach and estimate demand in the broader context of market equilibrium, thus generating more accurate predictions out of sample, and enabling counterfactual analyses.[3] In this latter aspect, this paper adopts the spirit of recent work that leverages state-of-the-art tools in empirical industrial organization to address policy questions where equilibrium effects and supply responses are first-order. For instance, in the case of Miravete et al. (2018) the response of oligopolistic firms significantly affects the analysis of optimal taxation. Similarly, in our case, considering the supply response (arising from the rules of the game as opposed to strategic behavior) is fundamental to studying lottery design.

Lotteries were a famous first application of the 1944 expected utility model of von Neumann and Morgenstern. Indeed, Friedman and Savage (1948) posed the basic puzzle of why the same people both gamble and buy insurance. A vast literature has since followed up exploring this natural question, which we speak to. We assert that people gamble or buy lottery tickets for both the expected winnings and the thrill. As a result, people buy tickets despite having an expected loss. In Section A.1, we use the tremendously large jackpots to rule out substantial risk preference or risk aversion since it is incompatible with the responsiveness we observe to increasing jackpots.

The paper is organized as follows. Section 2 gives the background on rollover lotteries. Section 3 develops our model and equilibrium analysis. Section 4 presents a first empirical application to the variation of lottery odds across nations. Section 5 introduces a second application to the U.S. national lottery market. Section 6 discusses how the model can be used for counterfactual analysis. Section 7 concludes.

[3] Walker and Young (2001) consider questions of lottery design akin to those we examine in this paper. They use a regression of sales on moments of the prize distribution to comment on optimal lottery design; in contrast, we assess counterfactuals with a structural model.

2 Background: Rollover Lotteries

Since ancient times, lotteries have been used to randomly allocate prizes. Examples can be found in the Bible,[4] in classical Rome, and in the Han Dynasty in China (Haigh, 2008). The concept of a lottery as a form of gambling, where a state entity sells chances and awards monetary prizes, originated in Europe during the Renaissance period: there are records of this practice in Venice and Florence during the 1520s (see Bradley, 2001, and its sources). Early lotteries adopted a similar format to modern day fundraising raffles. Tickets sold come with a receipt that identifies the buyer, and for each prize in a pre-determined set, the winning ticket is randomly drawn from a container. This format is impractical for anything but small-scale lotteries.[5]

The key innovation that would enable large-scale lotteries originated in the 1530s in Genoa, Italy (Bellhouse, 1991). In a *Genoese lottery*, buyers pick their a set of different numbers, and winning numbers are drawn without replacement. So selecting winners straightforward, as it involves drawing a few numbered balls from an urn, as opposed to drawing winning tickets that correspond to the set of players in each raffle. Prizes (including the jackpot, corresponding to guessing all numbers correctly) may be won by no one, or shared by many. Early Genoese lotteries offered fixed prizes set in advance, which entailed a risk of bankruptcy if sales were insufficient to cover prizes. This problem spurred further innovation.

Today's lotteries adopt a pari-mutuel payout, at least for jackpots and the largest prizes, in which all winning tickets share the prize money equal to a fraction of ticket sales. Moreover, a jackpot not won rolls over to the next draw. This format has been helped by the advent of online vending, which facilitates keeping an electronic ledger of bets and computing the prize pool in real-time. Not surprisingly, large-scale rollover lotteries became widespread in the late 20th century, as retail locations started being connected to computerized networks. Several U.S.states introduced new lotteries in the 1980s.[6] In 1982, Canada introduced the first modern national rollover lottery, Lotto 6/49, replacing a standard fixed prize national lottery (Bellhouse, 1991).

Modern rollover lotteries currently exist in many countries across the world. Their rules are largely similar: players choose numbers before each draw. Correctly guessing all numbers wins the jackpot, while identifying some of the numbers wins lesser fixed prizes. Lotteries allocate a fixed share of ticket sales to their prize pools (the *take*

[4] In Numbers, 34:13: "Moses commanded the Israelites: 'Assign this land by lot as an inheritance'."

[5] When the first English state lottery was held in 1567, adopting the raffle format and resulting in more than 400,000 tickets sold, the draw took four months to be completed (Haigh, 2008).

[6] For instance, New Jersey introduced Pick-6 Lotto in 1980 after establishing a retail network of up to 2,000 locations. See https://www.njlottery.com/en-us/aboutus/history.html.

rate), and unwon jackpots roll over. The remaining portion of sales is withheld by the lottery authority and is what we refer to as government *revenue*.

The two national U.S. lotteries, Powerball (1992–) and Mega Millions (2002–), have similar rules. Powerball (Mega Millions) players pick five numbers from 1 to 70 (69) and another from 1 to 25 (26). Both lotteries had bi-weekly draws, one midweek and one on the weekend, until Powerball introduced a third draw in 2021. These lotteries add more to state revenues than the corporate income taxes in many states.[7]

States face complex dynamic tradeoffs when setting lottery policy. Consider for instance that a higher take rate yields more revenue for any given ticket sales, but deters ticket sales and so hampers jackpot growth. Similarly, longer odds reduce the winning probability for any given jackpot, discouraging playing the lottery; on the other hand, they raise the chance of long rollover streaks with high eventual jackpots, boosting later ticket sales. Our equilibrium model allows us to use lottery data to estimate demand and smartly make these tradeoffs. It can thus guide lottery design.

3 Model

3.1 Setup

In our theoretical framework, ticket sales and jackpots arise from an *equilibrium* crossing of demand and supply for lottery tickets.

On the demand side, we assume that consumers cannot individually impact the current outcomes or future trajectory of jackpots. They thus act myopically for the current draw. Consumers trade off two lottery attributes: the gamble over the monetary loss — calculated using a risk-averse, neutral, or loving utility function — and the non-monetary *thrill* of the lottery experience. In what follows, we assume risk neutrality[8] in which case, the expected monetary *loss* from the lottery is decisive. Someone with a negative thrill dislikes the lottery experience but might play if they expect to win a positive amount. But anyone who enjoys a positive thrill is willing to play if they don't expect to lose too much money. As the expected loss increases, the number of lottery tickets sold Q falls. Integrating over the distribution of thrill values gives rise to a standard downward-sloping inverse demand function loss $\Lambda(Q)$. Notably, the loss plays the role of the price in our market, and not the lottery ticket face value. Here, we can act without loss of generality as if everyone buys at most one

[7] Source: U.S. Census Bureau, State Government Finances, https://www.census.gov/topics/public-sector/government-finances/data.html.

[8] We discuss preferences with risk loving or risk aversion in Appendix A.1.

lottery ticket. For given our exclusive focus on aggregate demand, a consumer buying n tickets yields the identical demand of n individuals all with the same thrill level buying one ticket each. Assume that $\Lambda(Q)$ is differentiable. Lottery demand functions are the main primitives estimated empirically in the application of Section 5.

We now turn to the supply of tickets. As usual, a supply function summarizes purchase opportunities available to consumers. But unlike standard markets, supply does not reflect active choices made by firms but rather embeds the lottery rules. The *inverse supply curve* maps the quantity Q of tickets sold in a given draw to the corresponding expected monetary loss implied by the laws of probability.[9] Under risk neutrality, the expected monetary loss is the ticket price p minus the expected winnings — namely, the sum of the expected jackpot winnings and the expected payoff $w > 0$ from lesser, non-jackpot prizes.[10]

The expected jackpot winnings depend on the *final jackpot* at the end of the draw period and the probability of winning inclusive of possible ties. A fraction τ (the *take rate*) of ticket revenues is withheld to pay for lesser prizes, lottery expenses, and government revenue.[11] Thus, the final jackpot is the *inherited jackpot* J from any rollover plus a share $1 - \tau$ of the current draw revenues not withheld by the lottery authority. So the jackpot for the current draw is $J + (1 - \tau)pQ$, where J is zero if someone won the jackpot in the previous draw.[12] Each ticket has an equal probability π of winning the jackpot, equally split among all winners. To facilitate our derivations, we let $\alpha = -\log(1 - \pi) \approx \pi$; the approximation error $O(\pi^2)$ is very small — since π is of order 10^{-8} for national lotteries in the U.S. For simplicity, we hereafter call α the *win chance*, and $1/\alpha$ the lottery *odds*.

Combining these pieces yields an *inverse supply curve* $L(Q|J)$, mapping the ticket quantity sold Q into the uniquely associated expected monetary loss based on the lottery rules, given inherited jackpot J. The expected loss L also depends on the take rate τ, winning chance π, and ticket price p, but we often suppress these arguments. The rollover feature plays a key role in identifying our demand, since the lottery changes draw by draw. Theorem 1 formulates the precise supply curve.[13]

[9] Throughout the analysis, we treat Q as a real number as opposed to an integer for simplicity.

[10] These are fixed probability prizes, excluded from the rollover mechanism, awarded to whomever matches three to five of the six winning numbers.

[11] Consistent with the data for Powerball and Mega Millions, we maintain that $w < p\tau$, which ensures that the non-jackpot prizes are fully covered, meaning that the lottery does not lose money.

[12] Some rollover lotteries, including Powerball and Mega Millions, reset to a positive prize floor after someone wins the jackpot. We abstract from this aspect to streamline the presentation of the model.

[13] A formula for the expected value of the jackpot in a rollover lottery appears in Cook and Clotfelter (1993). The novelty in Theorem 1 is formulating loss as a supply function, enabling equilibrium analysis.

Theorem 1 (Risk Neutral Supply) *The inverse supply is:*

$$L(Q|J) = p - w - [J/Q + p(1-\tau)][1 - e^{-\alpha Q}]. \tag{1}$$

For a quick proof, observe that the expected winnings per ticket equal w plus the expected per-ticket jackpot winnings, namely $J + p(1-\tau)Q$ times the chance $1 - e^{-\alpha Q}$ that the jackpot is won this draw,[14] divided by the quantity Q of tickets sold. Appendix B.1 offers an instructive alternative derivation of (1) that accounts for the many ways the jackpot can be multiply shared among 2,3,4,... winners.[15]

The inverse supply function depends on the quantity of tickets Q via two channels. As Q grows, more revenues inflate the final jackpot, thus lowering the expected loss. On the other hand, ties among multiple winners are more likely if more tickets are sold, which raises the expected loss. Thus, unlike textbook cases, here the inverse supply curve might *fall* in Q. This occurs when the first force dominates, which intuitively happens at small inherited jackpots J when new ticket sales matter more.

The supply curve depends on the inherited jackpot J. First, inverse supply shifts down as J rises. For the expected loss intuitively falls as the inherited jackpot rises, all else equal. Also, as J increases, inverse supply transitions from decreasing and convex in Q (for low J), to decreasing and then increasing in Q (for medium J), to increasing and concave (for large J). For any J, the inverse supply tends to $p\tau - w$ as $Q \to \infty$, since a large enough number of tickets sold for a draw swamps any inherited jackpot J. Theorem 2 formalizes these claims if inverse supply, depicted in Figure 1.

Theorem 2 (Shape of Inverse Supply Curve)
(*a*) *Inverse supply starts positive at $Q = 0$ for low inherited jackpots: $J < (p - w)/\alpha$.*
(*b*) *At inherited jackpot $J = 0$, supply is both globally falling and strictly convex.*
(*c*) *If $0 < J < 2p(1-\tau)/\alpha$, then supply is initially positive, falling and strictly convex, then rising and strictly convex, and finally rising and strictly concave. The supply minimum and inflection points fall in the inherited jackpot J.*
(*d*) *If $J > 2p(1-\tau)/\alpha$, then supply is rising and strictly concave.*
(*e*) *For any inherited jackpot J, supply tends to $p\tau - w$ as $Q \to \infty$.*

[14] This is the win chance for integers $Q > 0$, for no one wins with chance $(1-\pi)^Q = \left(e^{\log(1-\pi))}\right)^Q = e^{-\alpha Q}$. For instance, if ticket sales equal the lottery odds $1/\alpha$, then no one wins the lottery with chance $1/e \approx 37\%$.

[15] Our characterization of supply implicitly assumes that players choose numbers at random. Although there is evidence that this is not the case (see, e.g., Thaler and Ziemba, 1988), previous research finds that this is unlikely to matter quantitatively. For instance, Cook and Clotfelter (1993) find the correlation between actual coverage (i.e., the number of combinations played at least once) and random coverage to be almost 1 in Illinois lottery data.

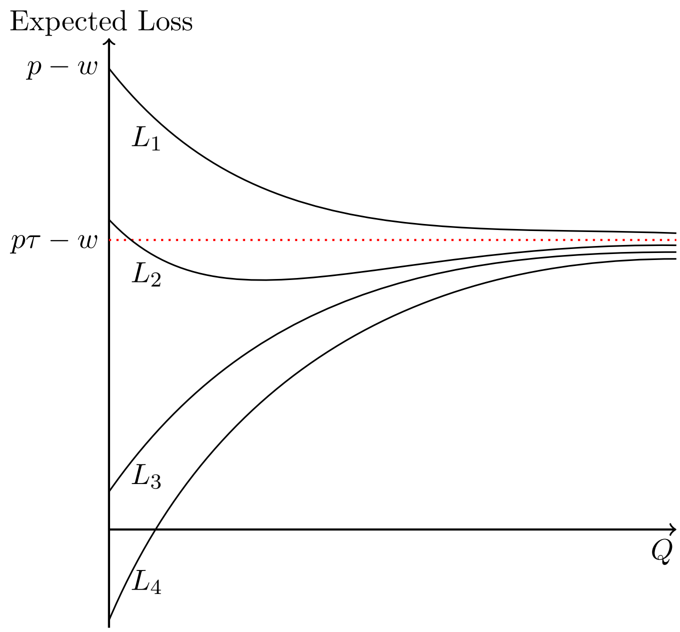


Figure 1: **Inverse Supply Curves as the Jackpot Rises.** We depict Theorem 2, with schematic inverse supply curves $L_i = L(\cdot|J_i)$ for respective jackpots $J_1 < J_2 < J_3 < J_4$. For low jackpots $J < 2p(1-\tau)/\alpha$, supply is first convex then concave, and first decreasing then increasing, while supply is concavely increasing for high jackpots $J > 2p(1-\tau)/\alpha$.

A classic fixed jackpot lottery has monotonically rising supply curve $L(Q|J) = p - J/Q$. A pure parimutuel has an infinitely elastic inverse supply $L(Q|J) = \tau p$ (from the take rate). The original Genoese lottery with no rollover has an increasing supply curve $L(Q|J) = p - J[1 - e^{-\alpha Q}]/Q$, while a parimutuel Genoese lottery has a falling supply curve $L(Q|J) = p - w[1 - e^{-\alpha Q}]$. The rollover lottery yields a supply curve like a parimutuel Genoese at low jackpots and original Genoese at high jackpots.

Proof of Theorem 2: For the vertical intercept of supply in part (a), consider:

$$L(0|J) = p - w - [J/Q + p(1-\tau)][1 - e^{-\alpha Q}]\Big|_{Q=0} = p - w - J\alpha. \tag{2}$$

The proofs of (b)–(d) are in Section B.2, but the rely on the derivative of (1):

$$L'(Q|J) = J/Q^2 - e^{-\alpha Q}\left([J/Q + p(1-\tau)]\alpha + J/Q^2\right). \tag{3}$$

By rewriting this and $L''(Q|J)$ as Taylor series, we compute thresholds $\overline{Q} \geq \underline{Q} \geq 0$, such that supply shifts from falling to rising at $\underline{Q}$, and has an inflection point $\overline{Q}$. Also, $0 < \underline{Q} < \overline{Q}$ for small inherited jackpots, $0 = \underline{Q} < \overline{Q}$ for larger inherited jackpots, and $\underline{Q} = \overline{Q} = 0$ for the largest inherited jackpots. Part (e) follows by inspecting (1). $\square$

3.2 Equilibrium Analysis

A *lottery equilibrium* for a draw is a crossing of supply and demand — namely, the *equilibrium quantity* $\mathcal{Q}(J)$ at which inverse supply (1) equals inverse demand $\Lambda(Q)$:

$$L(\mathcal{Q}(J)|J) \equiv \Lambda(\mathcal{Q}(J)). \tag{4}$$

In words, the equilibrium quantity $\mathcal{Q}(J)$ is a market-clearing expected loss — given by the inverse supply function $L(\mathcal{Q}|J)$. Since consumers must correctly forecast how many tickets will be sold to compute the expected ticket loss, a lottery equilibrium is a rational expectations equilibrium. While a strong assumption, this is supported by the good fit of the model in our empirical applications (Sections 4 and 5). In addition, Forrest et al. (2000a) find empirical evidence for rational expectations in the context of the UK national lottery.

We now provide conditions under which the equilibrium is unique and *stable* — namely, such that, given a small over- or under-purchase of tickets, market forces push demand back toward equilibrium. As usual, a lottery equilibrium is stable if the (possibly downward-sloping) inverse supply cuts the inverse demand curve from below at any crossing, i.e. $\Lambda(Q) = L(Q|J)$ implies $\Lambda'(Q) < L'(Q|J)$. For then, if slightly more tickets sell (so $Q > \mathcal{Q}(J)$), the available inverse supply loss exceeds the inverse demand (see Figure 2), discouraging marginal ticket purchases.

We can now establish the existence and uniqueness of a stable equilibrium.

Theorem 3 *Assume inverse demand $\Lambda(Q)$ exceeds inverse supply (2) at $Q = 0$, is below $p - w$ for large Q, and is steeper than supply at any crossing where $L' < 0$. Then there is a unique equilibrium, and it is stable.*

The proof in Section B.3 uses the intermediate value theorem, a falling demand curve, and a unique supply inflection point. We will see empirically that the assumptions on demand in Theorem 3 are met in our application to U.S. rollover lotteries. Uniqueness of the equilibrium is helpful in empirical analysis as it implies that one need not worry about multiple equilibria both in estimation and in counterfactual analysis. In addition, as Samuelson's correspondence principle makes clear, stability in partial equilibrium models is critical for intuitive comparative statics predictions. In our case, it ensures that the equilibrium lottery loss falls in the inherited jackpot, since the inverse supply L falls in J, as depicted in Figure 2. We record this and other comparative statics predictions in the following theorem.

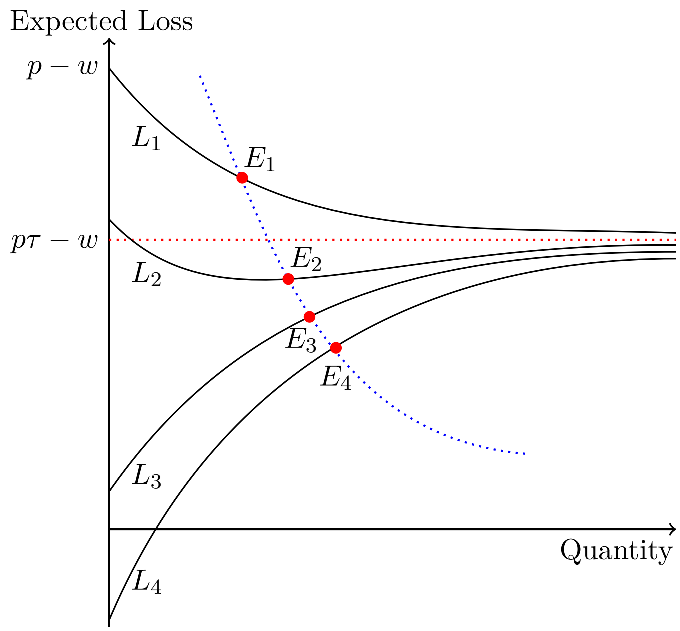


Figure 2: **Lottery Equilibria as Inherited Jackpot Rises.** This figure shows lottery equilibria for supply curves with inherited jackpots $J_1 < J_2 < J_3 < J_4$ from Figure 1.

Theorem 4 *The stable equilibrium quantity $\mathcal{Q}(J)$ falls in the ticket price p and take rate τ, and rises in the inherited jackpot J and winning chance α.*

Proof: Easily, from (1), the supply loss falls in the inherited jackpot J:[16]

$$L_J(Q|J) \;\;=\;\; -[1-e^{-\alpha Q}]/Q < 0.$$

Differentiate (4) in J. Since $L_J(\mathcal{Q}(J)|J) < 0$ and $\Lambda'(\mathcal{Q}(J)) - L'(\mathcal{Q}(J)|J) < 0$ by stability:

$$\mathcal{Q}'(J) = \frac{L_J(\mathcal{Q}(J)|J)}{\Lambda'(\mathcal{Q}(J)) - L'(\mathcal{Q}(J)|J)} > 0. \tag{6}$$

We can proceed likewise with our other parameters. Differentiating (1):

$$\begin{aligned}
L_\alpha(Q|J,\tau,\alpha,p) &= -Q[J/Q + p(1-\tau)]e^{-\alpha Q} < 0\\
L_p(Q|J,\tau,\alpha,p) &= 1-(1-\tau)[1-e^{-\alpha Q}] > 0\\
L_\tau(Q|J,\tau,\alpha,p) &= p[1-e^{-\alpha Q}] > 0.
\end{aligned} \tag{7}$$

These formulas allow us to analyze shifts in p, α, and τ. For example, we can

[16]Here, when necessary, we make explicit the dependence of the inverse supply and of the equilibrium quantity on the take rate τ, the win chance α and the ticket price p. We denote partial derivatives of the loss function in J, τ, α or p by subscripts throughout the paper.

differentiate (4) in τ at $Q = \mathcal{Q}(J, \tau, \alpha, p)$, and sign it with (7) and stability:

$$\mathcal{Q}_\tau(J, \tau, \alpha, p) = \frac{L_\tau(Q|J, \tau, \alpha, p)}{\Lambda'(Q) - L'(Q|J, \tau, \alpha, p)} = \frac{p[1 - e^{-\alpha Q}]}{\Lambda'(Q) - L'(Q|J, \tau, \alpha, p)} < 0.$$

We likewise can conclude $\mathcal{Q}_\alpha > 0 > \mathcal{Q}_p$ at any stable equilibrium. □

This result sheds light on the economics of the jackpot rollover. Why would a rollover lottery be preferable to a fixed prize lottery? Loosely, the rollover acts like a per unit subsidy on the next lottery draw, since it reduces the ticket loss. When taxing an activity, a greater subsidy is profitable when the response is elastic — namely, above the Laffer curve peak. So for a supply elasticity $\eta = \infty$ — such as for a standard parimutuel lottery with a constant take rate — a demand subsidy is profitable with elastic demand — namely, demand elasticity $\varepsilon < -1$.

In contrast, the supply curve for a rollover lottery is not infinitely elastic. Adapting (6), in equilibrium, the slope of lottery revenue in the inherited jackpot J is:

$$MR_J = (p\tau - w)\mathcal{Q}'(J) = (1 - e^{-\alpha Q})\frac{(p\tau - w)/L}{1/\eta - 1/\varepsilon}. \tag{8}$$

Meanwhile, the marginal cost of a rising jackpot J is $MC_J = 1 - e^{-\alpha Q}$, namely, the probability that the jackpot is won next draw.[17]

By Theorem 2, at low jackpots J, supply is falling and so its elasticity η is negative, but absolutely smaller than ε. In this case, the J-marginal revenue (8) exceeds $|\varepsilon| MC_J$ — since the numerator in (8) exceeds one at large jackpots, by Figure 2. We will see later in Figure 6 that demand is everywhere elastic for the two national rollover lotteries in the U.S., and thus if the lottery owner could deviate from the rules at low jackpots, a small jackpot increment would be strictly profitable.[18]

But for large enough jackpots J, supply is forever rising by Theorem 2, and so has a positive elasticity $\eta > 0$. So the J-marginal revenue (8) is strictly below $|\varepsilon| MC_J$, and thus potentially below MC_J. The lottery authority at this point would strictly desire a lower jackpot subsidy starting at that point, if it could do so. This suggests why the jackpot rollover is profitable, and clarifies the limits on its profitability.

Theorem 4 is also important because it motivates our identification strategy in Section 5.3: exogenous increments in inherited jackpots shift the inverse supply down and thus help trace out the demand curve, as illustrated in Figure 2.

[17]To see that equality of MC_J and MR_J determines static optimality, note that a \$1 increase in J has a net payoff of $(p\tau - w)\mathcal{Q}'(J) - 1$ if someone wins (chance $1 - e^{-\alpha Q}$) and of $(p\tau - w)\mathcal{Q}'(J)$ if no one wins.

[18]Not surprisingly, in many cases including the large U.S. rollover lotteries we examine in our application, the lottery subsidizes early draws by guaranteeing a minimum jackpot.

To conclude, we note that in this section we have focused on the case of a single lottery. When there are multiple lotteries (e.g., Powerball and Mega Millions, which we study empirically in Section 5), the arguments immediately extend. Specifically, each supply function is independent of the other since the rules of the game do not involve any interaction between different lotteries. On the demand side, there may be interactions between lotteries (we estimate them to be small), in which case the arguments in this section apply with "residual demand" replacing "demand."

4 Application 1: Lottery Odds Across Nations

We now explore the model's dynamic theoretical predictions. Specifically, we provide a first informal test of the model by shedding light on a known empirical regularity: lottery odds tend to vary in proportion to the population of the country or state. For instance, Powerball and Mega Millions each have jackpot odds of around 300 million to one, while Canada's Lotto Max has odds of about 33 million to one. We show that our model predicts this regularity under the additional assumption that lottery authorities maximize revenues, and that it holds broadly across different datasets.

4.1 Odds and Population: a Theoretical Analysis

We now assume that ticket demand, as formulated in Section 3, scales with the population. Namely, there is a fixed function $\underline{\Lambda}$ such that inverse demand is $\Lambda_N(Q) = \underline{\Lambda}(Q/N)$ for a country that is N times as populous. For example, if country A is twice as populous as country B, then A has double the lottery ticket demand of B, for each expected loss. For supply, if the lottery win chance scales to α/N, then the probability that someone wins $1 - e^{-(\alpha/N)(NQ)}$ is unchanged at each loss — since ticket sales rise by a factor N. All told, the equilibrium quantity function $\mathcal{Q}_N(\cdot|\alpha)$ for a country N times more populous than the country solving (4) obeys $\mathcal{Q}_N(\cdot|\alpha/N) \equiv N\mathcal{Q}(\cdot|\alpha)$.[19]

Now, assume a new lottery cycle starts. If no one wins in periods $0, 1, \ldots k-1$, the kth positive jackpot $J_k(\alpha)$ adds the untaxed portion of all past lottery ticket sales: $J_k(\alpha) = (1-\tau)p\mathcal{Q}(J_0(\alpha)|\alpha) + \cdots + (1-\tau)p\mathcal{Q}(J_{k-1}(\alpha)|\alpha)$. By independence of lottery draws, the probability that a new cycle starts with $k+1$ rollovers is the product:

$$e^{-\alpha\mathcal{Q}(J_0(\alpha)|\alpha)} \cdots e^{-\alpha\mathcal{Q}(J_k(\alpha)|\alpha)} = e^{-\alpha J_k(\alpha)/[p(1-\tau)]}.$$

[19] Indeed, we see that $L(\mathcal{Q}_N(J)|J, \alpha/N) \equiv \Lambda_N(\mathcal{Q}_N(J))$ holds for any $N > 0$ if (4) holds for $N = 1$.

This yields the following formula for the time undiscounted average lottery revenue:[20]

$$V(\alpha) = p\tau \frac{\mathcal{Q}(J_0(\alpha)|\alpha) + e^{-\alpha\frac{J_0(\alpha)}{p(1-\tau)}}\mathcal{Q}(J_1(\alpha)|\alpha) + e^{-\alpha\frac{J_1(\alpha)}{p(1-\tau)}}\mathcal{Q}(J_2(\alpha)|\alpha) + \cdots}{1 + e^{-\alpha\frac{J_0(\alpha)}{p(1-\tau)}} + e^{-\alpha\frac{J_1(\alpha)}{p(1-\tau)}} + \cdots}. \tag{9}$$

This formula accounts for the dynamic trade-off the lottery authority faces when devising their odds. A lower win chance α initially raises the loss and lowers demand, thus reducing early lottery revenues. On the other hand, eventually, the jackpot is more likely to grow very large, thus yielding a low expected loss and high demand (and revenues). The formula allows us to deduce that the optimal lottery win chance scale with the population, and so countries share the stochastic process of jackpot win chances.[21]

Theorem 5 *Assume that ticket demand scales proportionately in the population N, so that inverse demand is $\underline{\Lambda}(Q/N)$, for some fixed function $\underline{\Lambda}$. If the win chance α is optimal, the win chance α/N is optimal in a region with population N times higher.*

Proof: For a country N times as populous, write (9) as $V_N(\alpha) \equiv p\tau B_N(\alpha)/C_N(\alpha)$, where $B_N(\alpha) \equiv N B_1(N\alpha)$ and $C_N(\alpha) \equiv C_1(N\alpha)$. Maximizing $V_N(\alpha)$, the FOC is:

$$B'_N(\alpha)C_N(\alpha) = B_N(\alpha)C'_N(\alpha) \quad \Leftrightarrow \quad B'_1(N\alpha)C_1(N\alpha) = B_1(N\alpha)C'_1(N\alpha).$$

So α^*/N is optimal for N iff α^* is optimal for $N = 1$, as asserted. □

4.2 Lottery Data and Empirical Analysis

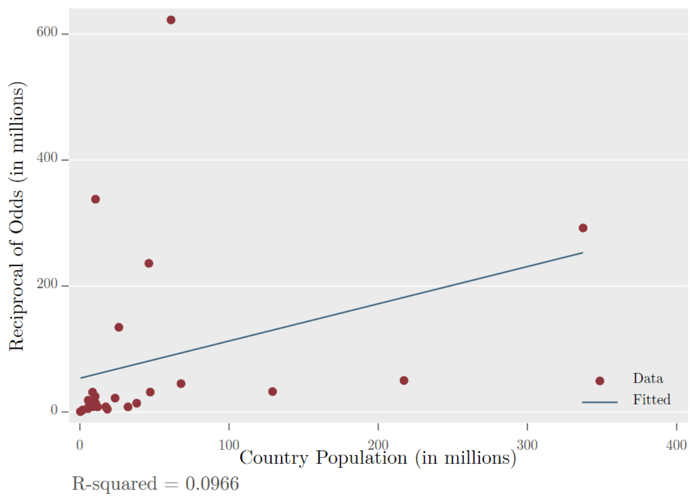
We introduce two datasets that we use for testing the predictions of Theorem 5: a dataset of rollover lotteries across countries and one of U.S. state lotteries.

First, we construct a database of large rollover lotteries from OECD countries. For each country, we obtain population data from the Census Bureau. We identify the largest (by sales) rollover lottery in each country, collect information on lottery rules, including odds, and exclude lotteries that substantially alter the rollover mechanism (e.g., by capping the maximum number of rollovers, or by capping the jackpot). This yields a database of twenty-four countries, each with a corresponding rollover lottery.

[20] Given a stochastic process of rewards z_n in periods $n = 0, 1, 2, \ldots$ that eventually may stop, and continues in period n with chance p_n, the mean reward is $(z_0 + p_0 z_1 + p_1 z_2 + \cdots)/(1 + p_0 + p_1 + \cdots)$.

[21] Our basic rollover lottery mechanism may be slightly altered. See Appendix C for a more detailed description of the rules of U.S. national rollover lotteries, and how these affect expected revenue calculations.

(a) Odds across countries



(b) Odds across U.S. states

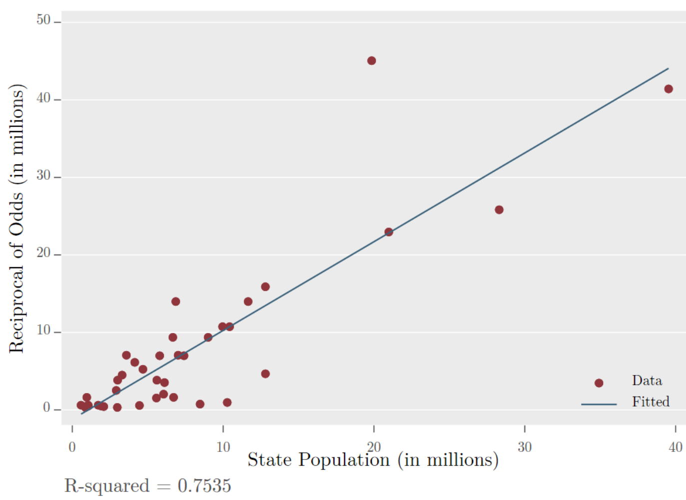


Figure 3: **Lottery odds and population.** The two panels show scatter plots of rollover lottery odds (in millions) and country (in panel a) or U.S. state (in panel b) populations (in millions). Each panel also shows a simple regression line. Slope estimates ($t-$statistics in parentheses) are 0.59 (1.53) and 1.14 (10.49), respectively.

We now turn to the U.S. Almost all states participate in Powerball and Mega Millions. We consider detailed data from these lotteries in Section 5, where we construct an empirical model of the national lottery market. Instead, here we focus on smaller *state-level* rollover lotteries. These lotteries are run by state agencies and are regulated by state legislatures, with state laws setting lottery rules including odds.

Compared to national lotteries, state rollover lotteries have similar rules, but much smaller jackpots. For each state, we collect population data, and obtain data on the lottery's odds from the websites of state lottery agencies. Overall, forty states offer rollover lotteries with rules that match our model.

By Theorem 5, as long as demand for rollover lotteries scales proportionally with population, revenue-maximizing lottery authorities would set odds in a way that also scales with population. We seek evidence on this by computing the relation between population and lottery odds in each of our datasets (across countries for the first and across states for the second). We report scatter plots and regression lines in Figure 3.

Across nations (panel (a)), we observe a positive relationship between odds and population, as the model predicts, but with large outliers. A potential explanation for this result is that, across very different nations, the demand functions vary substantially, which violates the assumption from Theorem 5 that demand scales proportionally with population. On the other hand, across U.S. states, the demand functions for state lotteries is likely to be much more homogeneous. Panel (b) of the figure supports this theory: a simple regression of state-level odds on population now has an R^2 of 0.76. Thus, in a more homogeneous context where demand may scale with population, our theory of optimal lottery odds is supported by the data. We further note that, even in a homogeneous context, we would not expect a perfect correlation between lottery odds and population if the lottery authorities' objective function depends on other factors besides revenues.

In sum, the data broadly supports the relationship between lottery odds and market size predicted by our model, under the additional assumption that lottery authorities maximize revenue and that demand scales with population. Although these strong assumptions are needed to obtain immediate testable implications from the model, our framework can also be used as a basis for quantification exercises when combined with credible estimates of demand, which we pursue in the next section.

5 Application 2: The National U.S. Lottery Market

We now use our theoretical framework from Section 3 to construct and estimate an empirical model of the U.S. national lottery market.

5.1 Lottery Data

We obtain data on draw-level prizes and sales for the two U.S. national lotteries, Powerball and Mega Millions, scraped from official lottery worksheets. We also collect

	Mega Millions	Powerball	
Start date	Oct 19, 2013	Jan 15, 2012	Oct 7, 2015
Ticket price ($)	1	2	2
Format	5/75 + 1/15	5/59 + 1/35	5/69 + 1/26
Jackpot (avg., $ million)	98	105	176
Reset value ($ million)	16	40	40
Probability of Jackpot Win	$1/258,890,000$	$1/175,223,510$	$1/292,201,338$
Expected loss (avg., $)	0.38	0.60	0.60

Table 1: **Lottery Rules and Main Summary Statistics.** This table reports information on lottery rules and summary statistics. Different columns refer to the two U.S. national lotteries (Mega Millions and Powerball) after the start date indicated. All dollar amounts are in nominal dollars. We convert annuity values to cash values using the discount rates applied by the lottery authority — see Appendix C for more details.

state-draw-level sales data from the website LottoReport.com.[22] We estimate the model on the period from October 19, 2013 to October 4, 2015 (204 draws) and use the period from October 7, 2015 to October 28, 2017 (216 draws) for out-of-sample validation.[23] In this latter period, Powerball changed its rules by substantially lengthening the odds of winning. This change allows us to assess how well our model captures equilibrium outcomes when taken truly out of sample, which corresponds to many counterfactual exercises that are relevant for lottery design. Table 1 shows the main lottery rules and summary statistics in our data.

As shown in Figure 4, national sales of the two lotteries fluctuate across draws, responding to the large fluctuations in the jackpot, which in turn generate large swings in expected loss. To quantify this elasticity, we estimate a flexible demand model in the rest of this section. We also note that sales and expected losses have a cyclical nature, and that the rollover mechanism generates outliers, such as the $1.6 billion Powerball jackpot of January 2016.

5.2 Empirical Model and Identification

Our theoretical model of equilibrium takes as key inputs the demand and supply functions. Supply is fully determined by the lottery rules and therefore requires no estimation. Turning to the demand functions, we choose to model them nonparametrically. This allows us to estimate the functions flexibly, which is important in

[22] Further information on our data is in Appendix C.

[23] Starting on October 28, 2017, Mega Millions changed its ticket price from $1 to $2. Since we offer no model of individual decision-making, we cannot speak to this change. If buyers choose to spend a given amount of money to secure their thrill, then this change should be neutral, since buyers would choose to buy half as many tickets after the price doubles.

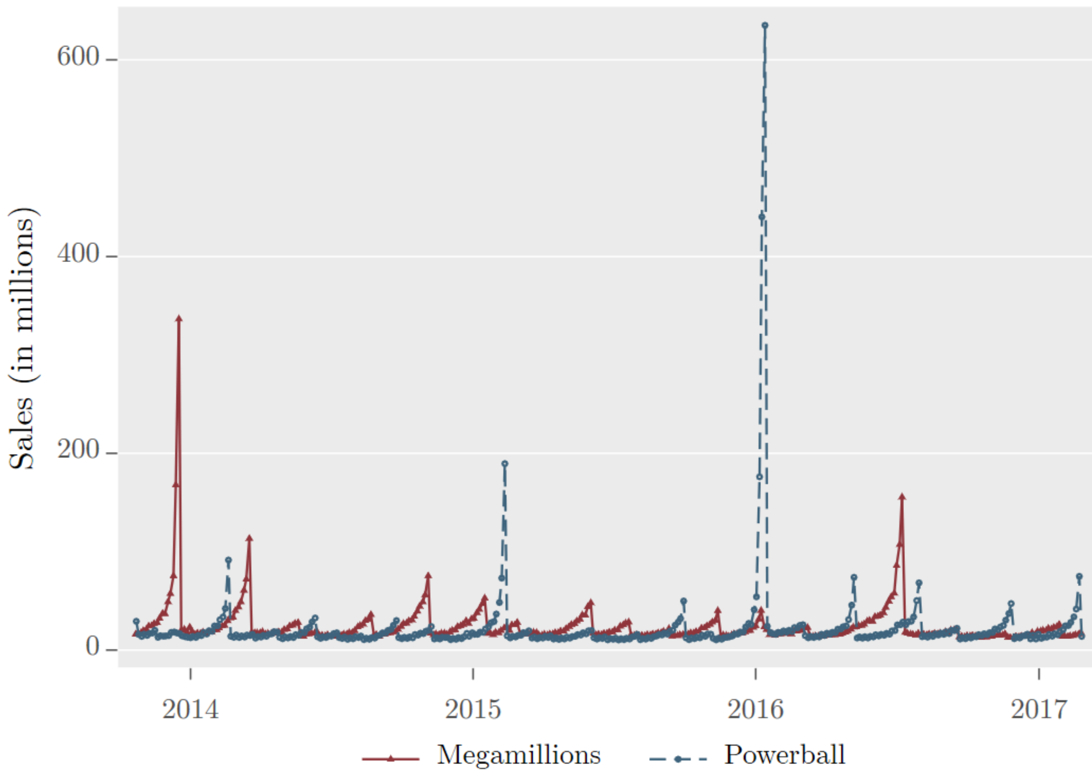


Figure 4: **Mega Millions and Powerball sales.** This shows the time series of sales (in millions of tickets sold nationwide) for Mega Millions and Powerball throughout our sample.

our setting, as the shape of demand — and particularly its curvature — is known to affect market equilibrium.[24] More specifically, we maintain risk neutrality, so that only expected loss matters, but do not impose parametric restrictions on how demand depends on expected loss and thrill (besides intuitive monotonicity restrictions). Because we directly target the demand functions and do not commit to a specific model of individual behavior, our estimates are consistent with a range of micro-foundations.

We model the demand for the two lotteries in draw t and state s as follows:

$$\begin{aligned} q_{MM,s,t} &= \sigma_{MM}(\delta_{MM,s,t}, \delta_{PB,s,t}, \lambda_{MM,t}, \lambda_{PB,t}) \\ q_{PB,s,t} &= \sigma_{PB}(\delta_{PB,s,t}, \delta_{MM,s,t}, \lambda_{PB,t}, \lambda_{MM,t}), \end{aligned} \tag{10}$$

where $q_{j,s,t}$ is the quantity of tickets sold for lottery j, $\lambda_{j,t}$ denotes the expected loss for lottery j, and

$$\delta_{j,s,t} = x'_{j,s,t}\beta + \xi_{j,s,t}$$

[24]While standard models (e.g., mixed logit) can accommodate a wide range of demand curvatures (Miravete et al., 2023), it may be hard to estimate sufficiently flexible specifications with aggregate data.

for observed attributes $x_{j,s,t}$ and unobservable lottery characteristics $\xi_{j,s,t}$ capturing any drivers of lottery demand that vary at the state-draw level (e.g., demand for lotteries may be especially low in Kentucky during the week of the Kentucky Derby since other betting opportunities are particularly salient). The vector $x_{j,s,t}$ consists of the number of years since the lottery was introduced in the state, fixed effects — for state, week, and lottery — and a dummy for whether the draw was in the first or second part of the week. The expected loss $\lambda_{j,t}$ is the same across states s since this is a national lottery, i.e. the sales across all states contribute to the jackpot and thus to the expected loss. Let

$$Q_{j,t} = \sum_s q_{j,s,t}$$

denote the corresponding aggregate demand across all states in a given draw. To connect the empirical model with the theoretical framework, fix $(\delta_{MM,s,t}, \delta_{PB,s,t})$ for all s and the loss for the competing lottery, and invert $Q_{j,t}$ in $\lambda_{j,t}$ to obtain the (residual) inverse demand. This corresponds to the function Λ in Section 3.

We now discuss what variation in the data identifies the model. We focus on the residual demand of either lottery and drop state subscripts. Our point of departure is the standard equilibrium analysis in markets for differentiated products. In the standard model, one typically uses exogenous supply shifts (induced by, e.g., cost shocks) to trace out the demand curve. Similarly, here the rollover mechanism exogenously shifts the supply curve. Thus, the key source of identifying variation is similar.

However, our setting differs from the standard contexts in meaningful ways. First, we note that supply — usually determined by a cost function and a markup function — is trivially identified in our setting, as it is purely a mechanical by-product of the lottery rules. This contrasts with what happens in a standard model, where marginal cost needs to be identified, and firms' markups are determined by an assumed model of strategic interaction. Thus, our empirical context rules out one standard econometric endogeneity concern: while typically firms set prices responding to the full vector of demand and cost unobservables in a market, our supply is known and non-strategic.

But unobserved demand shocks ξ may still be present and make the identification of demand non-trivial. In fact, since the supply curve is not flat, demand shocks will generally be correlated with a lottery's expected loss via the equilibrium mechanism. Note that, depending on the shape of the supply function, this correlation may have a surprising sign. In standard settings, demand shocks tend to be positively correlated with prices, thus attenuating the elasticity estimates towards zero if endogeneity is not taken care of. In our setting, demand shocks are also positively correlated

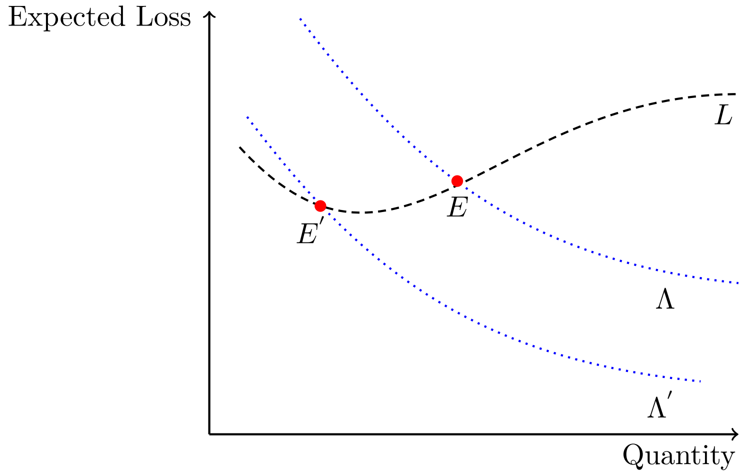


Figure 5: **Negative Lottery Demand Shock.** This figure depicts the equilibrium effects of a negative shock to lottery demand, shifting inverse demand from Λ to Λ'.

with expected loss as long as supply is upward-sloping at the equilibrium. But in areas where supply is downward-sloping, demand shocks and expected loss will be negatively correlated, thus potentially generating a downward bias on loss elasticity estimates. This mechanism is illustrated in Figure 5: a shift in demand from Λ to Λ' corresponding to a negative shock to lottery demand — for example, a big sporting event yielding alternative betting opportunities — results in a decrease in equilibrium expected loss as it occurs along the upward-sloping part of the supply curve. However, a further negative shift in demand would have potentially resulted in an *increase* in expected loss, as demand meets the downward-sloping part of the supply curve.

This analysis underpins our identification strategy: because supply is generically non-flat, we need an instrument to identify the demand curve. The instrument needs to affect the expected loss for the lottery, and be exogenous to unobserved demand shocks. A strong predictor of the level of the jackpot, and thus the expected loss, is whether a rollover occurred in the last lottery draw — which happens if no one won the lottery. But directly using an indicator for whether someone won at time $t-1$, denoted as win_{jt-1}, as an instrument for the expected loss at time t is not a viable strategy. A positive shock to ξ_{jt-1} will result in higher sales Q_{jt-1}, and thus a higher probability that someone wins the lottery at time $t-1$. So ξ_{jt-1} and win_{jt-1} are correlated. If demand shocks ξ_{jt} are serially correlated across time periods, win_{jt-1} is also correlated with ξ_{jt}, violating exogeneity.

Rather than directly using win_{jt-1}, we leverage our knowledge of the supply function to construct an instrument for the expected loss. Denoting by $\mathcal{I}_t$ the time t

information set, we know that

$$E\left[win_{jt-1}|\mathcal{I}_{t-1}\right] = 1 - e^{-\alpha_{jt-1}Q_{jt-1}}.$$

While $E\left[win_{jt-1}|\mathcal{I}_{t-1}\right]$ clearly depends on Q_{jt-1}, which in turn depends on ξ_{jt-1}, we seek to isolate (as a residual) the pure randomness of the draw to generate exogenous variation in expected loss. To this end, we define

$$z_{jt-1} = win_{jt-1} - E\left[win_{jt-1}|\mathcal{I}_{jt-1}\right].$$

By construction, this variable is independent of all variables determined at time $t-1$, including Q_{t-1}, thus making z_{jt-1} a viable instrument even in the presence of serial correlation in the unobservables ξ_{jt-1}. Our identification assumption is then:

$$E\left[\xi_{jt}|\mathbf{z}_{t-1}, \mathbf{x}_t\right] = 0,$$

where $\mathbf{x}_t$ are the observed exogenous characteristics. Consistent with Berry and Haile (2014), the excluded instruments $\mathbf{z}_{t-1}$ provide exogenous variation to tackle the endogeneity of expected losses, whereas the exogenous variables $\mathbf{x}_t$ serve as (included) instruments for quantities, which are also endogenous in equilibrium. When we regress the endogenous variables on exogenous variables, we obtain large F-statistics and coefficient signs that are consistent with economic intuition.[25]

While we maintain that this strategy is credible in our setting, we also mention possible threats to identification. Even when using only the "surprise" element of a lottery win, if unobserved demand shifters ξ_{jt} are set in a way that is dependent on z_{t-1}, our identification strategy is invalid. This could occur for instance if the lottery authority ramps up promotional activity following a jackpot win – which we do not believe is happening in our empirical environment. Alternatively, a big jackpot win could boost demand for the lottery in subsequent draws. In particular, Guryan and Kearney (2008) find that stores selling the winning ticket tend to see increased ticket sales in later draws. This effect is highly localized and thus does not pose a threat to our identification strategy since our data is at the state level. Only the case where demand for the entire state is boosted as a result of a previous win invalidates our instrument, and that is not consistent with the available evidence.

Finally, while in the standard model the causality goes from exogenous ξ to price (a higher demand shock ξ leads to higher prices in equilibrium), here the causality

[25] See Table 4 in Appendix C.4 for the full set of results.

could go in both directions. For example, ξ could represent media buzz that is itself generated by higher jackpots. This is relevant for how one should treat the unobservables ξ in the counterfactuals, but econometrically nothing changes as long as the exogeneity condition is met.

5.3 Nonparametric Estimation of Demand

We estimate demand functions σ_{MM} and σ_{PB} as well as the coefficients β on the exogenous x variables using the sieve-GMM approach proposed in Compiani (2022). Specifically, denoting a given lottery by j and the other lottery by k, we use results in Berry and Haile (2014) to write

$$\delta_{j,s,t} = \sigma_j^{-1}\left(q_{j,s,t}, q_{k,s,t}, \lambda_{j,t}, \lambda_{k,t}\right)$$

and approximate σ_j^{-1} via Bernstein polynomials. Note that $(\sigma_{MM}^{-1}, \sigma_{PB}^{-1})$ is the inverse of the demand system $(\sigma_{MM}, \sigma_{PB})$ in (10) with respect to its first two arguments — the δ indices — while keeping the last two arguments (the expected losses) fixed. A sufficient condition for the inverse to exist is that the two lotteries be weak substitutes, i.e. that as the expected loss of one lottery increases, the number of tickets sold for the other lottery (weakly) increases. We impose this restriction in estimation, but avoid any parametric assumptions on the shape of the demand curves. Specifically, we approximate the function $\sigma_j^{-1}(q_{j,s,t}, q_{k,s,t}, \lambda_{j,t}, \lambda_{k,t})$ using a linear combination of Bernstein polynomials:

$$\hat{\sigma}_j^{-1}(q_{j,s,t}, q_{k,s,t}, \lambda_{j,s,t}, \lambda_{k,s,t}) \equiv \sum_{0\leq v_1,v_2,y_1,y_2\leq m} \theta_{v_1,v_2,y_1,y_2} b_{v_1}^m(q_{j,s,t}) b_{v_2}^m(q_{k,s,t}) b_{y_1}^m(\lambda_{j,t}) b_{y_2}^m(\lambda_{k,t}),$$

where θ denotes coefficients to be estimated and $\{b_{v,m}\}_{v=0}^m$ the univariate Bernstein basis polynomials of degree m[26], so that the overall approximation degree is $4m$. We estimate the coefficients (β, θ) by minimizing a sieve-GMM criterion function obtained by projecting the residuals $\hat{\sigma}^{-1}(q_{j,s,t}, q_{k,s,t}, \lambda_{j,t}, \lambda_{k,t}) - x'_{j,s,t}\beta$ onto the exogenous variables $(\mathbf{x}, \mathbf{z})$. The objective function is a quadratic form in the coefficients (β, θ). Paired with the fact that substitution between the lotteries can be enforced via linear constraints on θ, this yields a well-behaved convex programming problem. We refer the reader to Compiani (2022) for details on the implementation of the estimator.[27]

[26] The Bernstein polynomials of degree m are defined as $b_{v,m}(x) = \binom{m}{v} x^v (1-x)^{m-v}$ for $v = 0, \ldots, m$.

[27] We also constrain the coefficients θ to be the same across the two lotteries. In words, this means that the shape of the two demand functions is assumed to be the same. Of course, this does not mean that the demand *levels* will be the same, since the two functions take different arguments (e.g., in our data the ticket

	λ_{MM}	λ_{PB}
Q_{MM}	-1.70	0.13
	(0.13)	(0.10)
Q_{PB}	0.03	-1.55
	(0.02)	(0.09)

Table 2: **Mean Elasticities in Expected Loss.** This table represents mean elasticities of lottery demand (quantities) to expected loss. Each row corresponds to quantities for one U.S. national lottery, and each column corresponds to the expected loss for one U.S. national lottery. Standard errors are reported in parenthesis below each elasticity figure.

As a comparison, we also consider the following log-log model of demand:

$$\log(q_{j,s,t}) = \gamma_0 + \gamma_{own}\log(J_{j,t}) + \gamma_{other}\log(J_{-j,t}) + \gamma_x x_{j,s,t} + \varepsilon_{j,s,t}, \tag{11}$$

where j is a lottery subscript, $-j$ is the other lottery (e.g., $-j$ is Powerball if j is Mega Millions), $J_{j,t}$ denotes the jackpot of lottery j inherited from draw $t-1$, and $x_{j,s,t}$ is as in the main model of equation (10). While simple and easily interpretable, this model misses that the final jackpot in draw t is a function of ticket sales in that draw and is thus an equilibrium outcome. In this respect, it is similar to models of demand for lotteries estimated in the literature (e.g., Forrest et al., 2000b). In contrast, we achieve that by specifying demand as a function of the expected loss and relating it to the number of tickets sold via the supply function. Also, the log-log model assumes constant demand elasticities, whereas our nonparametric approach learns the shape of the demand curves from the data. As we show next, we will find meaningful violations of the constant elasticity assumption.

5.4 Estimation Results

Panel (a) of Figure 6 plots Powerball's estimated aggregate inverse demand curve (we fix all demand drivers other than the own expected loss at their median values). Table 2 shows the mean elasticities of aggregate demand to own- and cross-expected loss, for the model imposes substitution between the two lotteries. We set $m = 2$, corresponding to a Bernstein approximation of degree 8.[28] The own-loss elasticities

price for Powerball — and thus its expected loss — is higher than for Mega Millions). This restriction is standard in demand estimation as it is implied by the common assumption that the coefficients on product attributes, as well as the distribution of the unobservables, be the same across products. We also estimated a version of the model that relaxes this assumption and found no meaningful differences in the point estimates at the cost of increased standard errors.

[28] The results are similar for $m = 3, 4$, corresponding to respective approximations of degree 12 and 16.

are larger than one in magnitude. Also, the cross elasticities are not statistically different from zero. The pattern is broadly consistent with the finding in Lockwood et al. (2021) of little substitution between the two lotteries. This could be driven by format differences of the lotteries (e.g., the ticket price was lower and the odds were less favorable for Mega Millions relative to Powerball in our estimation sample) as well as differences in the days of the week in which the lottery draws take place. Additionally, habit formation may contribute to this pattern. For instance, long-term Powerball players may not be prone to switching to Mega Millions, even if its expected loss is lower, purely out of habit. (We capture this by including the number of years since each lottery was introduced in a given state in the demand model).

Next, Figure 6, panel (b) shows the relationship between the expected loss and the estimated own-loss elasticity for Powerball across draws. The relationship between loss and elasticity is quite nonlinear, with relatively low elasticities for high values of the expected loss (i.e., when the inherited jackpot is low) and higher elasticities for intermediate values of the expected loss (i.e., when the inherited jackpot is higher). This is consistent with the following intuition: at low jackpots, the lottery tends to attract gamblers that are mostly driven by the thrill motive and are not very responsive to changes in the monetary loss, whereas as the jackpot grows the lottery attracts more gamblers that are mainly motivated by the monetary gain and thus more responsive to changes in the expected loss. This pattern suggests that more restrictive models (such as our benchmark log-log regression with constant elasticity) would likely lead to misspecification.

Finally, we test the assumptions of Theorem 3 that guarantee the existence of a unique equilibrium and find that they are satisfied. Specifically, we verify that at every equilibrium in the data, the derivative of the (residual) inverse demand for each lottery is below the derivative of the inverse supply.

5.5 Model Fit

To assess the fit of the model, we compare the average per-draw revenue observed in the data with the model predictions leveraging again the formula in (9). We do this for two main reasons. First, average revenue is obviously an outcome of interest to the lottery authority. Second, the randomness inherent in the lottery (and thus the potential for outlier jackpots) makes short-run prediction very challenging. By focusing on average revenues we smooth out the randomness inherent in the rollover mechanism, and predict instead a more stable, medium-run target.

Figure 7 shows the results of our prediction exercise. In panel (a), we look at

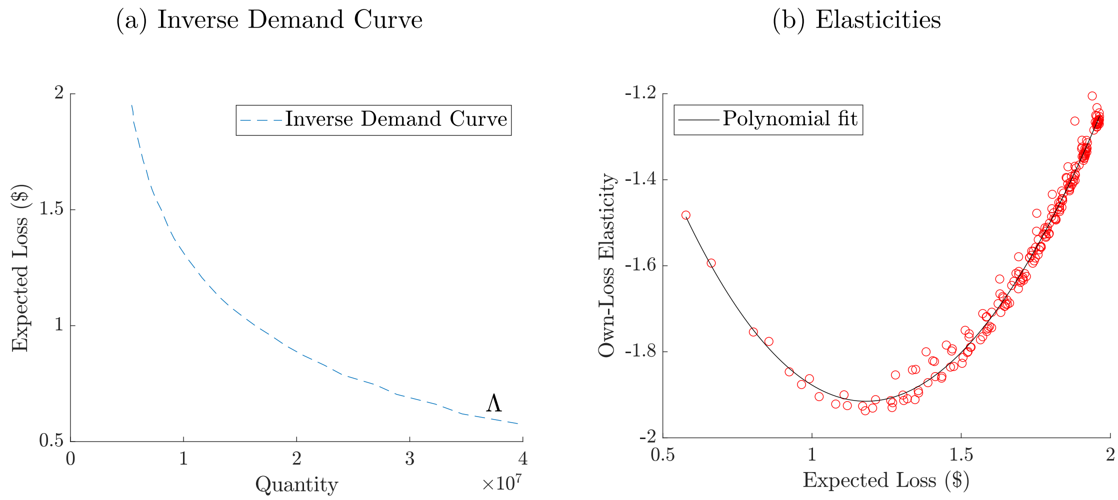


Figure 6: **Inverse Demand Curve and Own-Loss Elasticity for Powerball.** Panel (a) plots the inverse demand curve for Powerball (while fixing all drivers of demand other than the expected loss at their median values). Notably, nowhere in this range is it strictly profitable to play the lottery without a positive thrill. Panel (b) plots the relationship between expected loss and estimated own-loss elasticity across the draws in the sample.

how the model performs in sample, i.e. on the same data used for estimation, and find that our approach captures the average per-draw revenues from Powerball and Mega Millions well. As a comparison, we also compute the predictions from the simpler log-log model in (11).[29] This simpler model ignores that the final jackpot for any given draw is determined in equilibrium (since it is affected by how many tickets are sold during that period) and mimics the specifications commonly used in the literature that focus on the demand side alone. Interestingly, the log-log model tends to underestimate per-draw revenues for both lotteries. The intuition is clear: by ignoring the fact that the final jackpot will be larger than the inherited jackpot due to the concurrent ticket sales, the log-log model predicts lower sales volumes and thus lower revenues.

In panel (b), we repeat the same exercise out of sample. Specifically, we use the same two models to predict revenues in the period from October 7, 2015 to October 28, 2017. Relative to the estimation sample, in this period a few changes took place, notably a worsening of the odds for Powerball from around 1 in 175 million to around 1 in 292 million. Thus, this is a good test of whether our model can capture how revenues adjust to the levers that the lottery authority can pull, which is exactly

[29]Here we show the estimates obtained without instrumenting for inherited jackpots. We also tried a version of the model that uses the same instruments as in our approach and found that the estimates of revenues are very similar to those obtained without instruments, but with larger standard errors.

(a) In sample (b) Out of sample

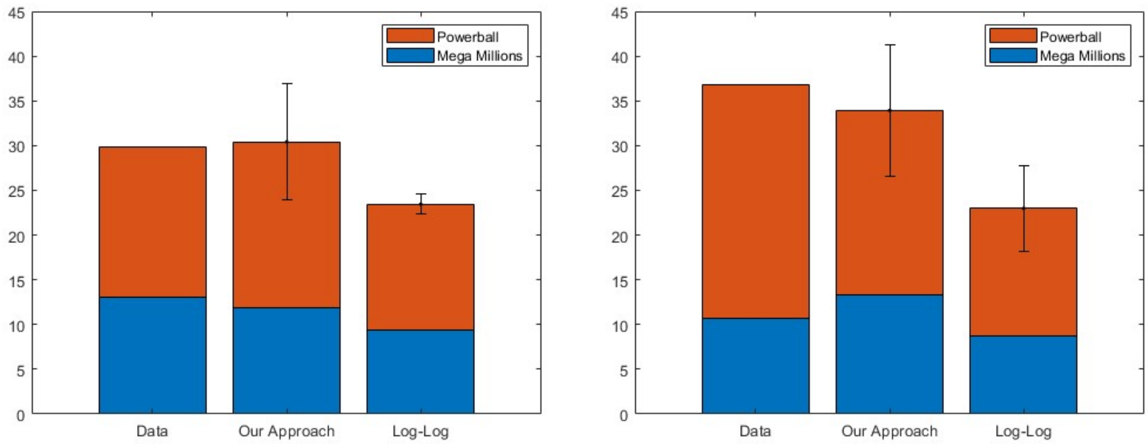


Figure 7: **Average Per-Draw Revenues (in millions of dollars).** We plot cumulative average per-draw revenues for Powerball and Mega Millions. The three bars in each panel are the data (left bar), estimated revenue according to our main demand specification (center bar), and estimated revenue according to a simple log-log model of demand. Panel (a) reports in-sample results, while panel (b) reports out-of-sample predicted values.

the kind of counterfactual exercises one would want to perform to inform lottery design. Similar to the in-sample fit exercise, we find that our approach estimates out-of-sample revenues reasonably well, whereas the simpler log-log model tends to deliver underestimates.

6 Discussion: How Can the Model Be Used?

In this paper, we pursue a positive model of lottery markets in which we characterize outcomes as an equilibrium phenomenon. We highlight two ways in which the model can be used, exemplified by each of our applications. First, the model generates comparative statics and motivates simple rules-of-thumb that can be used to rationalize lottery designs across countries. Second, as in our analysis of the U.S. national lottery market, the model can be brought to data to generate quantitative counterfactual predictions and thus inform lottery design.

Without an empirical model that captures both the demand and the supply response implicit in the rollover mechanism, it is impossible to study counterfactual questions that involve the design of the lottery. For instance, suppose that a researcher estimates demand for lottery tickets, e.g., as a function of jackpots in each draw and other lottery characteristics. The researcher then wishes to predict how

% Change in Take Rate τ	-10%	-5%	Status Quo	+5%	+10%
Per-draw revenues ($million)	30.6	32.5	34.0	35.2	36.2

Table 3: **Counterfactual revenues as take rate varies.** This table shows how the average per-draw revenues from Mega Millions and Powerball combined vary as the take rate τ varies. The calculation is based on the expression in (9) and refers to the period from October 7, 2015 to October 28, 2017. Thus, the lottery authority can increase revenues by locally increasing the take rate τ.

changes in the take rate, ticket prices, odds, or minor prizes of a lottery affect sales and revenues. Such changes do not automatically translate to a change in expected loss: the expected loss is an equilibrium outcome, and can only be determined jointly with sales. Hence, without a full equilibrium model — for example, if one only specifies demand — it is not possible to predict the effects of counterfactual policies. Although supply in this market is governed by lottery mechanics rather than by strategic firm behavior, the equilibrium aspect cannot be ignored.

Our approach can be used to investigate several counterfactual questions relevant to lottery authorities who may want to predict the effects of changes to the lottery rules on revenues. To illustrate this point, we consider one such counterfactual question: do current take rates on U.S. rollover lotteries maximize revenues? To address this, we use the demand model estimated in Section 5 and predict expected revenues for the out-of-sample period while varying the fraction τ of lottery revenues that are taken by the government in each draw (around two thirds for both lotteries in our data). Again applying equation (9), we calculate the average per-draw revenues for both lotteries and find that the overall revenues per draw increase monotonically with τ, from around \$30.6 million when τ is 10% lower than in the status quo to around \$36.2 million when it is 10% higher than in the status quo (Table 3). Thus, at least locally, the lottery authority may increase revenues by increasing take rates.

We emphasize that, despite our ability to address counterfactual questions, our model is positive since it takes no stand on the underlying normative utility. However, it can be used to complement normative analyses. For instance, Lockwood et al. (2021) characterize the optimal levels of the ticket price and jackpot. Since jackpots are equilibrium objects and cannot be directly set by the lottery authority, a natural question is how to reach that desired jackpot level. Our model can be used to characterize the values of the parameters under the control of the lottery authority — i.e., odds and take rate — that would generate those jackpot levels in equilibrium.

7 Conclusion

We have proposed a positive model of rollover lottery markets in which the number of tickets sold and the expected monetary loss are both determined in equilibrium. The equilibrium is given by the intersection of a demand curve — reflecting how gamblers' trade off the expected monetary loss against the thrill of gambling — and a supply curve, determined by the lottery rules. Since the ultimate ticket sales, and thus lottery loss, are not known at the time of purchase, it is a rational expectations equilibrium. Standard comparative statics in our market explain why a rollover is profitable: It allows a special form of stochastic second degree price discrimination.

Rollovers ensure that the supply curve sometimes crosses demand when jackpots are very high, and so losses low. We show how to use exogenous variation induced by the rollover mechanism, paired with recent advances in nonparametric methods, to flexibly identify and estimate the demand curve. The model allows one to run a range of counterfactual exercises. We test the validity of our framework in two empirical applications. First, comparing rollover lotteries across U.S. states as well as different countries, we find that longer odds are positively correlated with the size of the market as measured by population. This is consistent with what the model predicts lottery authorities should do if they wished to maximize expected revenues. Second, for the national U.S. rollover lotteries, we show that the model predicts out-of-sample outcomes reasonably well.

A natural avenue for future research is to apply the proposed framework to address questions on lottery design. For example, one could use the model to compute the revenue-maximizing lottery parameters (e.g., odds or take rate). In this sense, our positive analysis is complementary to existing normative analyses on optimal lotteries that prescribe optimal jackpot levels but do not address how to achieve those jackpots since they do not feature a notion of equilibrium.

References

Aruoba, S. B. and M. S. Kearney (2011): "Do lottery gamblers love risk or overweight small odds?" Tech. rep., Citeseer.

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2013): "The nature of risk preferences: Evidence from insurance choices," *American economic review*, 103, 2499–2529.

Bellhouse, D. R. (1991): "The genoese lottery," *Statistical Science*, 6, 141–148.

Berry, S. and P. Haile (2014): "Identification in differentiated products markets using market level data," *Econometrica*, 82, 1749–1797.

Bradley, R. E. (2001): "Eulerâs analysis of the Genoese Lottery," *Convergence*.

Chiappori, P.-A., B. Salanié, F. Salanié, and A. Gandhi (2019): "From aggregate betting data to individual risk preferences," *Econometrica*, 87, 1–36.

Clotfelter, C. T. and P. J. Cook (1987): "Implicit taxation in lottery finance," *National Tax Journal*, 40, 533–546.

Cohen, A. and L. Einav (2007): "Estimating risk preferences from deductible choice," *American economic review*, 97, 745–788.

Cohen, J. D. (2022): *For a Dollar and a Dream: State Lotteries in Modern America*, Oxford University Press.

Compiani, G. (2022): "Market counterfactuals and the specification of multiproduct demand: A nonparametric approach," *Quantitative Economics*, 13, 545–591.

Cook, P. J. and C. T. Clotfelter (1993): "The Peculiar Scale Economies of Lotto," *The American Economic Review*, 83, 634–643.

Forrest, D., O. David Gulley, and R. Simmons (2000a): "Testing for rational expectations in the UK national lottery," *Applied Economics*, 32, 315–326.

Forrest, D., O. D. Gulley, and R. Simmons (2000b): "Elasticity of demand for UK National Lottery tickets," *National Tax Journal*, 53, 853–863.

Friedman, M. and L. J. Savage (1948): "The utility analysis of choices involving risk," *Journal of Political Economy*, 56, 279–304.

Gandhi, A. and R. Serrano-Padial (2015): "Does belief heterogeneity explain asset prices: The case of the longshot bias," *The Review of Economic Studies*, 82, 156–186.

Gulley, O. D. and F. A. Scott (1993): "The demand for wagering on state-operated lotto games," *National Tax Journal*, 46, 13–22.

Guryan, J. and M. S. Kearney (2008): "Gambling at lucky stores: Empirical evidence from state lottery sales," *American Economic Review*, 98, 458–473.

——— (2010): "Is lottery gambling addictive?" *American Economic Journal: Economic Policy*, 2, 90–110.

Haigh, J. (2008): "The statistics of lotteries," in *Handbook of Sports and Lottery Markets*, Elsevier, 481–502.

Kearney, M. S. (2005): "State lotteries and consumer behavior," *Journal of Public Economics*, 89, 2269–2299.

Knight, B. and N. Schiff (2012): "Spatial competition and cross-border shopping: Evidence from state lotteries," *American Economic Journal: Economic Policy*, 4, 199–229.

Lockwood, B., H. Allcott, D. Taubinsky, and A. Y. Sial (2021): "The Optimal Design of State-Run Lotteries," Tech. rep., National Bureau of Economic Research.

Miravete, E., K. Seim, and J. Thurk (2023): "Elasticity and Curvature of Discrete Choice Demand Models," Tech. rep., Working paper.

Miravete, E. J., K. Seim, and J. Thurk (2018): "Market power and the Laffer curve," *Econometrica*, 86, 1651–1687.

Oster, E. (2004): "Are all lotteries regressive? Evidence from the Powerball," *National Tax Journal*, 57, 179–187.

Rabin, M. (2000): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281–1292.

Rabin, M. and R. H. Thaler (2001): "Anomalies: risk aversion," *Journal of Economic perspectives*, 15, 219–232.

Thaler, R. H. and W. T. Ziemba (1988): "Anomalies: Parimutuel betting markets: Racetracks and lotteries," *Journal of Economic perspectives*, 2, 161–174.

von Neumann, J. and O. Morgenstern (1944): *Theory of games and economic behavior*, Princeton university press.

Walker, I. and J. Young (2001): "An economistâs guide to lottery design," *The Economic Journal*, 111, 700–722.

Appendix A Robustness and Alternative Models

A.1 Risk Preference or Dispreference

We have so far assumed that lottery players are risk neutral and willing to play the lottery even when they expect to lose money from it because they derive a non-monetary thrill from it. This contrasts with the vast literature that follows Friedman and Savage (1948) explaining why people can both buy insurance and play the lottery. (Our answer is that lottery has an extra fixed thrill.) They posited initially risk-averse and eventually risk-loving behavior. Of course, by itself, this cannot explain why individuals would play low-stakes lotteries.[30] Still, it is worth asking: might our lottery players employ a nonlinear utility function, thus deviating from risk neutrality?

Posit Bernoulli utility function $u_r(x) = (1 - e^{-rx})/r$, where $r \neq 0$, and wealth $x > 0$. This is *constant absolute risk aversion* (CARA) if $r > 0$, and *constant absolute risk loving* (CARL) if $r < 0$, and is risk neutral at $r = 0$ via l'Hopital's Rule: $\lim_{r\to 0} u_r(x) = \lim_{r\to 0} xe^{-rx} = x$. *We assume individuals differ by lottery thrill and share a common risk preference parameter r.*

We have not specified the number of tickets needed for the lottery thrill, as it was inessential. (We pursue an aggregate demand curve — and could not identify this number in any event.) But now we must take a stand to derive the supply curve. Assume lottery buyers buy a single ticket. For any initial wealth $I > 0$, the final wealth is $I - p$ plus thrill and winnings.[31] To derive supply, we must focus on the marginal ticket buyer who is indifferent about the purchase, and derives the same expected utility from playing and not playing the lottery. The expected lottery utility of the marginal ticket buyer replaces the thrill by the loss $L(Q|J)$, i.e. $1/r$ times

$$1 - (1-\pi)e^{-r[I+L(Q|J)-p+w-\frac{1}{2}rw\sigma^2]} - \pi e^{-\alpha Q} \sum_{k=0}^{n} C(Q,k)\pi^k e^{-r\left[I+L(Q|J)-p+\frac{[J+p(1-\tau)Q]}{k+1}\right]},$$

given the expected value w and variance σ^2 of smaller prizes. As the utility of not buying is $\frac{1}{r}(1-e^{-rI})$, the indifference equation does not depend on I:

$$(1-\pi)e^{-r[w-\frac{1}{2}rw\sigma^2]} + \pi e^{-\alpha Q} \sum_{k=0}^{n} C(Q,k)\pi^k e^{-r\frac{[J+p(1-\tau)Q]}{k+1}} = e^{-r[p-L(Q|J)]}.$$

We can now solve this equation for L to obtain the loss that must be offered in order

[30] Rabin and Thaler (2001) call this utility function "contrived", and note it has implausible implications.

[31] That is, we impose the assumption that utility is additively separable in loss and thrill. This restriction is not needed for the model with risk neutrality in the main text.

to sell Q tickets. If the loss was higher than this level, fewer than Q tickets would be sold; if the loss was lower, more than Q tickets would be sold.

Theorem 6 *With CARA or CARL utility $u_r(x)$, the inverse supply is:*

$$L^r(Q|J) = p + \frac{1}{r}\log\left((1-\pi)e^{-r[w-\frac{1}{2}rw\sigma^2]} + \pi e^{-\alpha Q}\sum_{k=0}^{Q} C(Q,k)\rho^k e^{-r\frac{[J+p(1-\tau)Q]}{k+1}}\right).$$

For small α, in the risk neutral limit $r \to 0$, the lottery loss $L^r(Q|J)$ tends to (1).

Proof: Equating utility with income I and no ticket to the expected utility with a ticket, and its attendant costs and benefits:

$$\begin{aligned}(1-e^{-rI})/r &= (1-\pi)(1-e^{-r[I-p+L(Q|J)+w-\frac{1}{2}rw\sigma^2]})/r \\ &\quad + \frac{\pi}{r}\sum_{k=0}^{Q} C(Q,k)\pi^k(1-\pi)^{Q-k}\left[1-e^{-r[I-p+L(Q|J)]-r\frac{[J+p(1-\tau)Q]}{k+1}}\right].\end{aligned}$$

Rearranging terms:

$$e^{r[-p+L^r(Q,J)]} = (1-\pi)e^{-r[w-\frac{1}{2}rw\sigma^2]} + \pi\sum_{k=0}^{Q} C(Q,k)\pi^k(1-\pi)^{Q-k}e^{-r\frac{[J+p(1-\tau)Q]}{k+1}}.$$

Take logs, and put $(1-\pi)^Q = e^{-\alpha Q}$ to get the limit. By l'Hopital's rule, $p - L^r(Q,J) \to (1-\pi)w + \pi e^{-\alpha Q}\sum_{k=0}^{Q} C(Q,k)\rho^k\left[\frac{J+p(1-\tau)Q}{k+1}\right]$ as $r \to 0$. Small π gives the limit. $\square$

So Theorem 1 is the risk neutral limit $r \to 0$ of Theorem 6.

We now exploit our very large lottery jackpots to show that unless the absolute Arrow-Pratt measure of risk preference is extremely small, we arrive at implausible implications for losses or counterfactual predictions of ticket sales. For this exercise, we calibrate lottery characteristics to typical values for national U.S. rollover lotteries. Specifically, we assume the lottery price, odds, and take rate are such that $\frac{p(1-\tau)}{\alpha} \approx 100M$, set the inherited jackpot at 50M and the non-jackpot prices w at 0.25.

First, consider any slight risk preference — namely, a fixed negative risk aversion parameter $r < 0$. In the context of the U.S. national lotteries, this leads to implausibly large certainty equivalents for typical, large enough, inherited jackpots. (See Figure 8.) In other words, this forces our inverse supply curve to explode negatively. So, marginal lottery players — who, say, play at a jackpot of \$40M but not \$20M — have a negative thrill. This implies that typical lottery players dislike playing, which is not realistic. Note that this argument holds for any demand function.

On the other hand, slight risk aversion leads to implausible ticket sales predictions

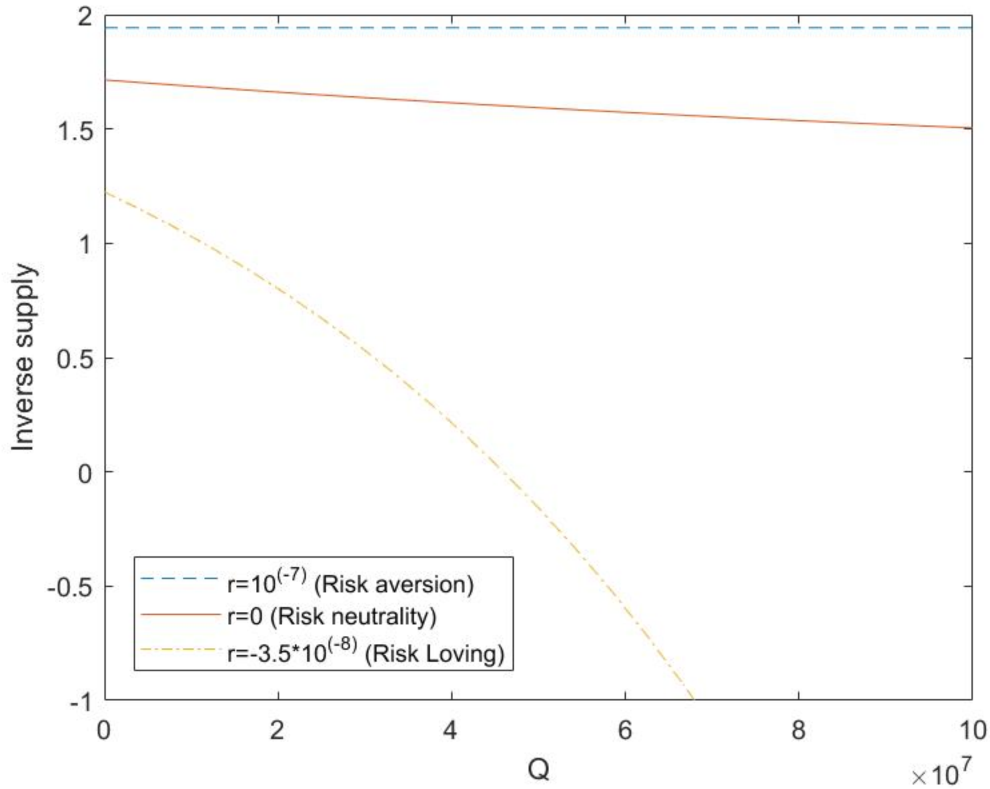


Figure 8: **Inverse Supply Curves with Risk Preference.** Plotted are the supply curves with a typical inherited jackpot of \$67M, against quantity ($\times 10^7$). Inverse supply quickly becomes negative for slight risk loving preference $r \geq -10^{-7}$. Inverse supply is almost perfectly elastic for risk aversion preference $r \leq 10^{-8}$.

(Rabin, 2000). As seen in Figure 8, for $r > 10^{-8}$, the exponent of the last term in (6) does not vary by more than 0.1 if the jackpot grows by \$10M. Since the supply curve is nearly perfectly elastic, large inherited jackpots have essentially no impact on ticket sales for a fixed demand curve. To explain the patterns we see in our data, we would need to posit large unobservable shifts in demand that are systematically correlated with inherited jackpot size, thus requiring a theory on how the shocks are generated.

All told, we find the risk neutral analysis more parsimonious, and we dispense with risk preference as an explanation for the observed behavior.

A.2 Jackpot Levels in Demand

In our lottery market model, demand depends on the expected loss. This allows us to establish an intuitive parallel with standard partial equilibrium analysis, done in the price-quantity space. But different modeling assumptions are possible: for example, demand may depend directly on the jackpot level, as opposed to the expected loss.

This would require choosing a definition of jackpot. Possible candidates are lagged jackpots from the previous draw, rational expectation jackpots for the current draw, or advertised jackpots. In the latter case, performing counterfactual analyses requires a model of how the lottery authority determines advertised jackpots given the available information, which can be problematic if we think that the lottery authority has more information than researchers. Including jackpots alone in demand would also limit the range of counterfactuals we can run: for instance, we could not examine how counterfactual changes in odds affect equilibrium outcomes in this market – something our framework can address given that the expected loss is a function of the odds. Finally, in our specific application to U.S. national lottery markets, we find that including lagged jackpots instead of the expected loss in the demand function leads to similar out-of-sample fit.

Another option is to include in demand both jackpots and expected loss. There are empirical considerations that suggest avoiding this route: jackpots and expected loss, though not identical, are highly correlated, and separately identifying the effects of these variables on demand (especially in a non-parametric context) is very hard. While it is possible to add instruments (e.g., more lags of our instrument) to mechanically obtain estimates for our model, we have not been able to come up with sources of variation that cleanly identify the separate effects of jackpots and expected loss.

Appendix B Omitted Proofs

B.1 Risk Neutral Supply: Proof of Theorem 1

For completeness, we derive the earlier formula directly, considering all the possible numbers of ticket winners — highlighting the positive probability of a shared jackpot. Here, we assume Q is a natural number. The chance of exactly k *other winners* is $C(Q-1,k)\pi^k(1-\pi)^{Q-1-k}$, where $C(n,k) = \frac{n!}{k!(n-k)!}$. So the expected winnings are

$$w + \pi[J + p(1-\tau)Q] \sum_{k=0}^{Q-1} \frac{C(Q-1,k)}{k+1} \pi^k (1-\pi)^{Q-1-k} \tag{12}$$

So winners secure a $1/(k+1)$ share when k others pick the same winning sequence. We use generating functions to show that this collapses to (1). Define $\rho \equiv \pi/(1-\pi)$

and a new function $f(\rho) \equiv \sum_{k=0}^{Q-1} \frac{C(Q-1,k)}{k+1} \rho^{k+1}$. Then expected winnings (12) are:

$$w + [J + p(1-\tau)Q] f(\pi/(1-\pi))(1-\pi)^Q$$

Since $f'(\rho) \equiv (1+\rho)^{Q-1}$ and $f(0) = 0$, integration yields $f(\rho) = \frac{1}{Q}[(1+\rho)^Q - 1]$. As $1 + \rho = 1/(1-\pi)$, and $(1-\pi)^Q = e^{-\alpha Q}$, winnings per ticket (12) imply losses:

$$w + [J + p(1-\tau)Q] \cdot \frac{(1/(1-\pi))^Q - 1}{Q} \cdot (1-\pi)^Q$$

This reduces to the expected gains subtracted in (1) from the price p. $\square$

B.2 Curvature of Inverse Supply: Proof of Theorem 2

We argue that when supply is not monotone, a unique *inverse supply minimum* exists:

$$x = \underline{Q}(J) \quad \Leftrightarrow \quad e^{\alpha x} - 1 - \alpha x = [p(1-\tau)/(\alpha J)]x^2 \tag{13}$$

Since $e^{\alpha x} > 1 + \alpha x + \frac{1}{2}\alpha^2 x^2$, the root $\underline{Q}(J)$ exists if $0 < J < 2p(1-\tau)/\alpha$, and falls in J. Easily, it explodes at small inherited jackpots: $\underline{Q}(J) \uparrow \infty$ as J vanishes.

Lemma 1 *(a) If $J = 0$, then supply is monotonically falling for all Q.*
(b) If $J < 2p(1-\tau)/\alpha$, then supply is falling then rising for $Q \lesseqgtr \underline{Q}(J)$.
(c) At high inherited jackpots $J \geq 2p(1-\tau)/\alpha$, supply is monotonically increasing.
(d) The supply minimum $\underline{Q}(J)$ falls in J, and obeys $\underline{Q}(J) < \frac{3}{\alpha}\left(\frac{2p(1-\tau)}{\alpha J} - 1\right)$.

Proof: The slope of inverse supply in (3) implies:

$$L'(Q|J) = \left(e^{\alpha Q} - 1 - \alpha Q[1 + p(1-\tau)Q/J]\right) J e^{-\alpha Q}/Q^2 \tag{14}$$
$$= \left(\frac{\alpha^2 Q^2}{2}\left(1 - \frac{2p(1-\tau)}{\alpha J}\right) + \sum_{k=3}^{\infty} \tfrac{1}{k!}(\alpha Q)^k\right) J e^{-\alpha Q}/Q^2 \tag{15}$$

Then

$$L'(0|J) = \tfrac{1}{2}\alpha[\alpha J - 2p(1-\tau)]$$

So supply starts falling if $J = 0$, as $L'(0|J) = -\alpha p(1-\tau)$. If $J \geq 2p(1-\tau)/\alpha$, then $L'(Q|J) > 0$ always, by (15). If $J < 2p(1-\tau)/\alpha$, the lead term of (15) is negative: So $L'(Q|J) < 0$ for small $Q > 0$.

Finally, the zero of (14) is obviously $\underline{Q}(J)$, and its properties have been laid out.

Also, tossing aside all but one term in the infinite sum yields:

$$\frac{1}{2}\left(1-\frac{2p(1-\tau)}{\alpha J}\right)+\frac{1}{3!}(\alpha \underline{Q}(J))<0 \qquad \Rightarrow \qquad \underline{Q}(J)<\frac{3}{\alpha}\left(\frac{2p(1-\tau)}{\alpha J}-1\right)$$

Supply starts positive iff $J<(p-w)/\alpha$, and falling iff $\underline{Q}(J)>0$, or iff $J<2p(1-\tau)/\alpha$. Since $p-w>2p(1-\tau)$, supply starts off positive and falling for low J, positive and rising for intermediate J, and negative and rising for high J. □

We next argue that supply is first convex and then concave in Q. Loosely, since higher degree polynomials grow faster, the supply slope L_Q in (15) changes sign at most once, $-$ to $+$, and ends $+$. Inverse supply curvature turns on understanding the root:

$$x=\overline{Q}(J) \quad \Leftrightarrow \quad e^{\alpha x}=1+\alpha x+\alpha^2x^2/2+\alpha^2p(1-\tau)x^3/(2J) \tag{16}$$

Since $e^{\alpha x}>1+\alpha x+\frac{1}{2}\alpha^2x^2+\frac{1}{6}\alpha^3x^3$, a root $\overline{Q}(J)$ exists iff $0<J<3p(1-\tau)/\alpha$.

Lemma 2 *(a) If $\underline{Q}(J)=0$, inverse supply $L(Q|J)$ is concave.*
(b) If $\underline{Q}(J)>0$, inverse supply $L(Q|J)$ is strictly convex in Q on $[0,\overline{Q}(J)]$, and strictly concave in Q on $[\overline{Q}(J),\infty)$, where the roots (13) and (16) obey $\overline{Q}(J)>\underline{Q}(J)$.
(c) The supply curve inflection point $\overline{Q}(J)$ is falling in J.
(d) At inherited jackpots $J\geq 3p(1-\tau)/\alpha$, supply is initially concave in Q.

Proof: Differentiating (3):

$$\begin{aligned}
L''(Q|J) &= -2J/Q^3+e^{-\alpha Q}(\alpha Q+2)J/Q^3+\alpha e^{-\alpha Q}\Big([J/Q+p(1-\tau)]\alpha+J/Q^2\Big)\\
&= \left[-e^{\alpha Q}+1+\alpha Q+\alpha^2Q^2/2+\alpha^2p(1-\tau)Q^3/(2J)\right]2Je^{-\alpha Q}/Q^3\\
&= \left[\alpha^2p(1-\tau)Q^3/(2J)-\sum_{k=3}^{\infty}\tfrac{1}{k!}(\alpha Q)^k\right]2Je^{-\alpha Q}/Q^3
\end{aligned}$$

Then

$$L''(0|J)=\alpha^2[p(1-\tau)-\alpha J/3]>0$$

From (16), $\overline{Q}(J)$ falls in J, and explodes at small inherited jackpots: $\overline{Q}(J)\uparrow\infty$ as $J\downarrow 0$. Now, (19) vanishes at $\overline{Q}(J)$. When $2p(1-\tau)Q/J\leq\alpha$, we have $\overline{Q}(J)>\underline{Q}(J)$. For $L'(Q|J)\leq 0$ implies $L''(Q|J)>0$ when $Q\leq\underline{Q}(J)$, by (3) and (17):

$$\begin{aligned}
L''(Q|J) &= -2L_Q/Q+(Je^{-\alpha Q}/Q^2)\left(2\alpha+2/Q+\alpha^2[Q+p(1-\tau)Q^2/J]\right)\\
&\approx -2L'(Q|J)/Q-(\alpha Je^{-\alpha Q}/Q)\,(2p(1-\tau)Q/J-\alpha)
\end{aligned}$$

Finally, $L''(Q|J)<0$ since $2L'(Q|J)/Q<0$ and we assumed $2p(1-\tau)Q/J<\alpha$. □

B.3 Unique Equilibrium: Proof of Theorem 3

Existence owes to continuity, and the Intermediate Value Theorem: demand exceeds supply at $Q = 0$, and supply exceeds demand at ∞, as $L(\infty, J) = p - w$.

Next, we claim that $\Lambda(Q) - L(Q|J)$ is downcrossing through zero, and so the equilibrium is unique and stable. Indeed, once the supply curve is increasing, it does so forever, by Theorem 2. So, multiple equilibria can only happen when supply slopes down. But the supply curve is convex when it is decreasing (by Theorem 2), and so the supply curve steepens in Q. A second crossing with a falling demand curve is impossible: After one crossing, $\Lambda(Q) - L(Q|J)$ falls in Q. □

Appendix C U.S. Lotteries: Details and Data Construction

C.1 Lottery Rules: the Fine Print

Both Mega Millions and Powerball introduce significant modifications to the basic rollover lottery mechanism.[32] First, they advertise an annuity value for the jackpot, as opposed to cash amounts. The jackpot amounts that are commonly advertised on billboards refer thus to the sum of 30 increasing yearly payments, which grow at a 5% rate every year. As an alternative, winners of the jackpot can choose to receive the full cash amount of the prize; the annuity rates are set by competitive auction. In our analysis, we assume that consumers take into account the cash value of the prize in computing the expected loss from playing. In addition, the advertised jackpot is an estimate of the actual jackpot for the current draw.

Notably, the lottery authority does not commit to paying out the advertised (cash) value of the prize: rather, it will pay out the cash amount of the underlying jackpot prize pool. Exceptions to this are two instances in which advertised jackpots are guaranteed: (i) in the first draw after the jackpot is won, the lottery will start from a set minimum amount (e.g., $40 million in annuity value for Powerball after 2015), and (ii) typically for the first few rolls, a minimum increase of the jackpot is guaranteed (e.g., $10 million in annuity value for Powerball after 2015), to speed up the increase of the jackpot. Therefore, the lottery authority may have to pay jackpots that exceed the value of the jackpot pool during one of the guaranteed draws.

Lastly, lottery authorities seem to actively manage additional reserve accounts, perhaps because in the guaranteed period the jackpot prize pool would otherwise be

[32] Detailed lottery rules can be obtained from state law. See, e.g., the Texas Administrative Code on Powerball rules: https://texreg.sos.state.tx.us/public/readtac$ext.TacPage?sl=R&app=9&p_dir=&p_rloc=&p_tloc=&p_ploc=&pg=1&p_tac=&ti=16&pt=9&ch=401&rl=317 (accessed November 2023).

insufficient to pay out jackpot wins. The laws and regulations are somewhat unclear as to how this is done; for instance, in the Powerball regulations we find the following language: "An amount up to 5% shall be deducted from a Party Lottery's Grand Prize Pool contribution and placed in trust in one or more Powerball prize pool accounts [...] is below the amounts designated by the Product Group."

C.2 Expected Revenue for Powerball and Mega Millions in Finite Data

We create a tractable expected revenue formula for our finite data set with n lottery draws starting at an arbitrary jackpot. To proceed, we approximate it with a constant continuation chance $\delta_n = 1 - 1/n$. This rollover process lasts expected n periods.

In line with our discussion in Section C.1, Powerball and Mega Millions have modified the rollover jackpot process from the simple case described in (9). We show here that a similar formula for expected revenue can be derived in this special case. The actual jackpot process $(\hat{J}_k)$ imposes a minimum jackpot $\underline{J} > 0$ and a minimum jackpot increment $\underline{\Delta} > 0$. Since the first constraint only binds on the first draw, we have $\hat{J}_0(\alpha) = 0$ and for $k = 1, 2, \ldots$:

$$\hat{J}_k(\alpha) = \begin{cases} \max((1-\tau)p\mathcal{Q}(0|\alpha), \underline{J}) & k = 1 \\ \hat{J}_{k-1}(\alpha) + \max((1-\tau)p\mathcal{Q}(\hat{J}_{k-1}(\alpha)|\alpha), \underline{\Delta}) & k > 1 \end{cases}$$

The expected lotto revenue over n periods is:[33]

$$V_n(\underline{J}|\alpha) = p\tau \frac{\mathcal{Q}(\hat{J}_0(\alpha)|\alpha) + \delta_n e^{-\alpha \frac{J_1(\alpha)}{p(1-\tau)}} \mathcal{Q}(\hat{J}_1(\alpha)|\alpha) + \delta_n^2 e^{-\alpha \frac{J_2(\alpha)}{p(1-\tau)}} \mathcal{Q}(\hat{J}_2(\alpha)|\alpha) + \cdots}{1 + \delta_n e^{-\alpha \frac{J_1(\alpha)}{p(1-\tau)}} + \delta_n^2 e^{-\alpha \frac{J_2(\alpha)}{p(1-\tau)}} + \cdots}$$

Finally, if we start at jackpot $J_0 > \underline{J}$, then the first jackpot run-up is shorter than later ones. The expected per period average lotto revenue is

$$\begin{aligned} V_n(J_0|\alpha) &= V_n(\underline{J}|\alpha) + (1-\delta_n)[p\tau\mathcal{Q}(\hat{J}_0(\alpha)) - V_n(\underline{J}|\alpha)] \\ &\quad + \delta_n(1-\delta_n)e^{-\alpha \frac{J_1(\alpha)}{p(1-\tau)}}[p\tau\mathcal{Q}(\hat{J}_1(\alpha)) - V_n(\underline{J}|\alpha)] \\ &\quad + \delta_n^2(1-\delta_n)e^{-\alpha \frac{J_2(\alpha)}{p(1-\tau)}}[p\tau\mathcal{Q}(\hat{J}_2(\alpha)) - V_n(\underline{J}|\alpha)] + \cdots \end{aligned}$$

For the jackpot lasts at least to draw $k = 1, 2, 3, \ldots$ with chance $\delta_n^{k-1} e^{-\alpha \frac{J_k(\alpha)}{p(1-\tau)}}$.

[33] There's an important nuance now. *Actual ticket sales, and not the adjusted jackpot, determine the continuation probabilities.* Thus, the exponents use $J_k(\alpha)$ and not $\hat{J}_k(\alpha)$.

C.3 Data Construction

We construct our data from different sources. We obtain information about lottery rules from official documents,[34] and collect data on ticket price, odds, rollover rules (including take rate), and minor prizes. Using odds and amounts for minor prizes, we can immediately compute the expected value for those—since they do not involve a rollover mechanism and are paid out to all winners irrespective of how many, this step is straightforward. Second, we scrape official lottery worksheets[35] to obtain data on lottery-draw-level advertised jackpots, and actual annuity values of the jackpot prize pool. For each draw, we record the number of winners. We also collect data on annuity rates to convert annuity values into cash values. Although worksheets contain total sales information, they do not contain state-level information. This is available from each state lottery agency; we scrape the data from Lottoreport.com, and validate them by consulting different state lottery agencies, finding no discrepancies.

Finally, we use our data to construct expected losses $\lambda_{j,t}$ for each lottery j and draw t according to Equation (1):

$$\lambda_{j,t} = p_j - w_{j,t} - [J_{j,t}/Q_{j,t} + p_{j,t}(1 - \tau_{j,t})][1 - e^{-\alpha_{j,t}Q_{j,t}}],$$

where $p_{j,t}, w_{j,t}, Q_{j,t}, \tau_{j,t}$ and $\alpha_{j,t}$ are in our data. We compute the (cash value) jackpot $J_{j,t}$ recursively by applying the rollover mechanism, or $J_{j,t} = J_{j,t-1} + p_{j,t}Q_{j,t}(1 - \tau_{j,t})$, where we set $J_{j,t-1} = 0$ if someone won the jackpot in the previous draw.

C.4 First Stage Results

In Table 4 below, we report the results of OLS regressions of the endogenous variables in our model on exogenous variables, including the instruments. The values of the F-statistic suggest a strong association between the exogenous variables and endogenous outcomes. In line with economic intuition, the lagged residual instrument has a strong positive correlation with the lottery's expected loss and a negative correlation with its sales. In line with our findings that there is limited substitution

[34] See for instance `https://hoosierlottery.com/getmedia/8870e03d-8346-427f-8033-261a1beadd06/Powerball-Group-Rules-8-23-21.pdf` (accessed November 2023) for the latest Powerball rules. Previous versions can be recovered with WaybackMachine at various state lotteries' websites.

[35] The worksheets for Powerball are available at `https://www.texaslottery.com/export/sites/lottery/Games/Powerball/Estimated_Jackpot.html` (accessed in November 2023). Similar documents for Mega Millions are available at `https://www.texaslottery.com/export/sites/lottery/Games/Mega_Millions/Estimated_Jackpot.html` (accessed in November 2023).

across lotteries, the correlation between the lagged residual for a given lottery and the expected loss and sales of the competing lottery is much smaller in magnitude, indicating a limited economic effect. Because we include state and week fixed effects, the correlation between the number of years since introduction of a lottery in a state and the endogenous outcomes is overall quite weak.

Table 4: Lottery Rules and Main Summary Statistics

	(1)	(2)	(3)	(4)
Variables	Expected loss		Sales	
	Mega Millions	Powerball	Mega Millions	Powerball
Lagged residual - own lottery	0.208***	0.297***	-654,612***	-285,488***
	(0.00306)	(0.00446)	(46,615)	(20,002)
Lagged residual - other lottery	0.0180***	0.0337***	64,830*	-9,405
	(0.00257)	(0.00521)	(39,082)	(23,372)
Years since introduced - own lottery	-0.0376***	-0.00306	-20,352	-98,285
	(0.00914)	(0.0291)	(141,014)	(130,345)
Years since introduced - other lottery	-0.00185	-0.0733***	54,090	81,719
	(0.00602)	(0.0208)	(101,650)	(93,148)
Constant	0.998***	1.749***	-280,418	489,817
	(0.0308)	(0.0854)	(479,988)	(382,918)
R-squared	0.887	0.875	0.538	0.552
F-statistic	471.8	413.9	69.30	72.99

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table reports results of a regression of exp losses and sales for each lottery on instruments, exogenous variables, and fixed effects for week, week part, and state.