An Equilibrium Model of Rollover Lotteries[*]

Giovanni Compiani[†] Lorenzo Magnolfi[‡] Lones Smith[§]

December 5, 2024

Abstract

We develop a novel model of rollover lottery ticket sales, assuming preferences over thrill and money won. Treating the monetary loss on tickets as an implicit price, lottery rules imply an inverse supply curve. Growing jackpots shift the inverse supply down, and help identify the falling demand curve arising from thrill heterogeneity. We nonparametrically estimate the demand for Powerball.

Our model allows risk aversion or risk loving preferences, but we show that even slight deviations from risk neutrality deviate from the data tremendously. This is a high stakes empirical test (based on 160 million gamblers) of Rabin's (2000) calibration theorem that low stakes risk aversion yields implausible larger stakes implications. While ticket buyers are risk neutral, Powerball acts as a risk loving gambler for rollovers up to \$540M, but should cap the jackpot at \$920M.

Aside from the excellent model fit, we check risk neutrality in two ways. First, we characterize how log ticket sales should convexly grow in the log jackpot at least up to \$409M — which we verify in Powerball data. Next, lottery odds should scale linearly in the population — which we verify in a regression across forty state rollover lotteries.

JEL codes: D81, L10, L83.

[*]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).

[†]University of Chicago - Booth School of Business, giovanni.compiani@chicagobooth.edu

[‡]Department of Economics, University of Wisconsin, Madison, WI 53706, magnolfi@wisc.edu

[§]Department of Economics, University of Wisconsin, Madison, WI 53706, www.loonessmith.com.

1 Introduction

Rollover lotteries are the dominant lottery genre since the 1980s. They build on a Genoese lottery, where ticket buyers pick their own number combinations, and all tickets matching a randomly drawn winning combination equally share a jackpot prize. But if no one wins, the jackpot rolls over to the next draw, seeding a higher stakes lottery. The lottery authority chooses the odds — about three hundred million to one for Powerball and Mega Millions — as well as the share of ticket sales rolled over to the next jackpot.

American consumers spend over \$100 billion on state lotteries — exceeding combined sales of movie tickets, video streaming, concert tickets, and books. An eighth of U.S. adults play the lottery weekly and 50% do so annually,[1] and as rollover jackpots can rise very high, over half play at least once a year (Cohen, 2022). Lotteries rival sports gambling as the largest economic footprint of gambling. Despite their social costs as regressive taxation, governments use them because they constitute a major revenue source — \$31 billion in 2021.[2]

Gambling was an early use of expected utility. Friedman and Savage (1948) asked why the same people gamble and insure. As Rabin and Thaler (2001) wryly summarized, "Friedman and Savage (1948) contrived a convex/concave utility-of-wealth function . . . to reconcile the existence of both risk loving gambling and risk-averse insurance preferences in an individual". In this paper, we build and test an equilibrium lottery model assuming people gamble not only because of monetary gain, but because they derive a fixed intrinsic gain (*thrill*) from it. Buyers purchase tickets when the expected utility of winnings plus gambling thrill exceeds the abstention utility. We then argue strongly for risk neutrality, where our model is most tractable — one buys a set number of tickets when the thrill exceeds the ticket price minus the expected winnings.

Our model is best thought of as an "implicit market", since the price is *not the stated ticket price but instead the expected monetary loss on a lottery ticket.* For this implicit price falls in demand (lower thrill buyers will buy tickets at a lower loss), and on the implied (inverse) supply loss falls in the jackpot. Thrill heterogeneity generates aggregate demand, while lottery rules imply an (inverse) supply function mapping ticket sales to the expected money loss, given

[1] See www.fool.com/money/research/lottery-statistics/

[2] See www.statista.com/statistics/249128/us-state-and-local-lottery-revenue

Jackpot ($ million)	50	150	250	350	450	550	650
Ticket sales (million)	13	24	61	131	245	409	633
Sales increment		11	37	70	114	164	224

Table 1: **Powerball Jackpots and Tickets Sales.** Average Powerball Ticket Sales Rise in the Jackpot at an Increasing Rate. (We convert all jackpots in the paper to their lump sum present values; see Section 4 for details on our data.)

the jackpot inherited from past draws plus a fixed tax on new ticket sales. The crossing of supply and rationally forecast demand determines the new ticket sales and expected loss (Theorems 1–3). Rollovers reduce expected losses, and so shift inverse supply down, helping identify this model. This implicit market reduces all essential lottery structure to its inverse supply curve (e.g. Figure 1).

The response of the ticket sales at large Powerball jackpots is incompatible with all but risk neutrality given how responsive ticket sales are to jackpots. Indeed, even slight deviation from risk neutrality cannot possibly explain ticket sales at high jackpots. For consider interpolated ticket sales in Table 1 at Powerball lottery draws. Even after jackpots have ascended into the hundreds of millions, ticket sales not only continue to rise, but at an increasing rate. For the literature's smallest risk aversion parameters, a jackpot rise from $\$250M$ to $\$350M$ should not impact the desirability of ticket sales. And for even slight risk loving parameters, marginal ticket buyers must derive a *hugely negative thrill*. This offers an empirical test of Rabin's (2000) calibration theorem that risk aversion in low stakes gambles yields implausible implications at larger stakes — where the stakes are in the hundreds of millions of dollars, and the numbers of gamblers reflected in our data exceeds 160 million.

We then explore a structural empirical implementation of the model, and nonparametrically estimate demand for Mega Millions and Powerball. In a typical differentiated product oligopoly, supply involves strategic interaction among firms. But here inverse supply is pinned down by lottery rules. We exploit the randomness inherent in the rollover mechanism to estimate the demand functions — akin to the classic idea of using cost shocks to identify demand. Notably, demand falls from elastic to inelastic. We use new nonparametric methods (Compiani, 2022) to estimate the demand system as a flexible function of expected losses and thrill components. This does not require committing to a specification for consumer utility. Justifying our focus on the single lottery problem, we find little cross-substitution among Powerball and Mega Millions, consistent with existing literature. When we apply the estimates to predict equilibrium outcomes, we find our approach fits the data well.

We apply our implicit market model to underscore that lottery ticket buyers are indeed risk neutral. Theorem 4 proves that ticket sales are increasing and convex in jackpots at least up to \$400M, consistent with Table 1. We use the relative risk loving coefficient of the lottery owner, whom we call Lotto, to prove more strongly that log ticket sales are convex in log jackpot; the fit of our model to the ticket sales data in this log-log space is excellent. See Figure 4.

We explore one predictive implication for rollover design. Lotto faces a dynamic gamble since one can view the jackpot as Lotto's money invested to increase ticket sales, at the cost of a jackpot won: As noted, Lotto is initially strictly risk loving — his expected ticket sales revenue is strictly convex in the jackpot. But after he tips risk averse, he is still willing to continue the rollover, since the gambles are favorable. We predict that at a jackpot around \$920M, Lotto would ideally wish to cap the rollover, since the chance of someone winning is monotonically increasing to one as the jackpot rises (Theorem 5 and Figure 5). In fact, Canada's Lotto Max caps its rollover at \$80M, and EuroMillions caps its rollover at \$250M. Caps are the rule outside the USA.

For our final lens on risk neutrality, we consider the set of all rollover lotteries across forty U.S. states. These rollover lotteries vary in their odds — longer odds for larger population states. Cook and Clotfelter (1993) first noted this correlation. They proposed a behavioral explanation, venturing that players infer lottery odds from the frequency of jackpot wins. Our rational theory by contrast derives a stronger linear relation of optimal ticket odds and population, assuming that lottery ticket demand scales linearly in the population.

To see this, note that the rollover lottery is dynamic: A lower win chance raises the lottery ticket loss for early rounds, reducing early lottery revenues. But it raises the chance that the jackpot grows larger, thus increasing demand and revenues later on. By scaling the odds directly to the population, Lotto maximizes revenues when buyers care linearly about the lotto prize or win chance. This ensures that marginal costs and benefits of changing the odds remain balanced. Loosely, ticket buyers in a state twice as populous respond equally if the doubled jackpot prize is compensated by a halved win chance.

Across U.S. states — where we expect similar demand conditions — we find an R^2 coefficient above 0.84, suggesting that the linear model explains the relationship between odds and market size in the data very well.

PRIOR LITERATURE. While we have not found any market models of rollover lotteries, its predictions are evocative of well-studied economic models.

Firstly, given our implicit price, the rollover lottery with its falling supply curve randomly offers increasing discounts. But it wholly differs from the logic of stochastic sales models, such as Varian (1980) — least of all as because it does not require incomplete information or search costs (Sobel, 1984).

Next, slow falls and fast price rises is reminiscent of the durable goods problem with demand inflow (Sobel, 1991; Conlisk et al., 1984). But that economic environment and logic is unrelated: "By lowering its price over time, the firm can make extra sales to low-valuation agents; however, such a strategy will lead some high-valuation agents to postpone their purchases" (Board, 2008).

Finally, storability also induces cycles in prices (Hendel et al., 2014).

A wealth of papers address risk preference implication in real world gambles. For instance, Post et al. (2008) found risk aversion on the TV show "Deal or No Deal" for prizes up to €$5M$. Our stakes are 200 times higher, and our winning chances accordingly *much* smaller. So while our goal of inferring risk aversion from naturally observed choices is surely not new, we hope a market inference from over 160 million lottery ticket buyers makes it worthy of consideration.

Lastly, there are structural papers on rollover lotteries. Our overlap is negligible: We pursue questions arising in a rational market model not possible in a non-equilibrium behavioral structural setting (Lockwood et al., 2021).

OVERVIEW. We briefly sketch the history of this popular rollover lottery market in Section 2. Section 3 develops our model and equilibrium analysis. In Section 4, we develop our estimate of the demand curve, whose elasticity is key to our theory. We focus on Lotto's risk preference over market sales in Section 5, and prove Lotto should cap large rollovers. Insights about Lotto explain how ticket sales grow at high jackpots. Section 6 shows that best lottery odds linearly scale with population, and then finds this in a forty state regression. We adapt our model to allow for risk preference or aversion in Appendix A. Appendix B finishes all proofs. Appendix C concludes our empirical work.

2 Background: The Emergence of Rollover Lotteries

Early lotteries resembled modern fundraising raffles, where tickets come with a receipt identifying the buyer, and a single winning ticket is randomly drawn.[3] In the 1530s, in Genoa, Italy, a betting pool arose on which five public officials candidates would be randomly picked. Formally, in this *Genoese lottery*, buyers pick their own numbers trying to match numbered balls drawn without replacement from an urn. So zero, one, or many bettors may win the prize.

Lotteries grew widespread in the late 20th century, using retail locations linked to computerized networks. In 1975, New Jersey created the first Genoese lottery: Pick-It.[4] A worldwide lottery sea change happened starting in 1980, when New Jersey added Pick-6 Lotto, the first rollover lottery. Buyers picked six numbers 10–46, and unwon lotteries rolled over. In 1982, Canada replaced its national raffle lottery with Lotto 6/49, the first national rollover lottery.

National U.S. lotteries Powerball (1992–) and Mega Millions (2002–) have since emerged. In Powerball, one picks five numbers from 1–70 and another from 1–25; in Mega Millions, ranges are 1–70 and 1–26. Both lotteries had biweekly draws, until Powerball introduced a third in 2021. They provide more revenues than corporate income taxes in many states.[5] Transnational rollover lotteries EuroMillions began in 2004 in west Europe, and Eurojackpot in 2012.

3 Model and Equilibrium

3.1 Supply and Demand

We explore a large lottery run by an agency whom we call Lotto. Ticket sales reflect an equilibrium of demand and supply in a suitable market. We assume a continuous ticket quantity $Q \geq 0$. We use integers when we empirically estimate lottery demand, or make probabilistic statements about tie events.

Buyers trade off a monetary loss — computed for a risk averse, neutral, or risk loving utility function — and a non-monetary lottery real-valued *thrill*. We assume risk neutrality, and argue that fixed slight risk loving or aversion

[3] 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).

[4] See `https://www.njlottery.com/en-us/aboutus/history.html`.

[5] Source: `https://www.census.gov/topics/public-sector/government-finances/data.html`.

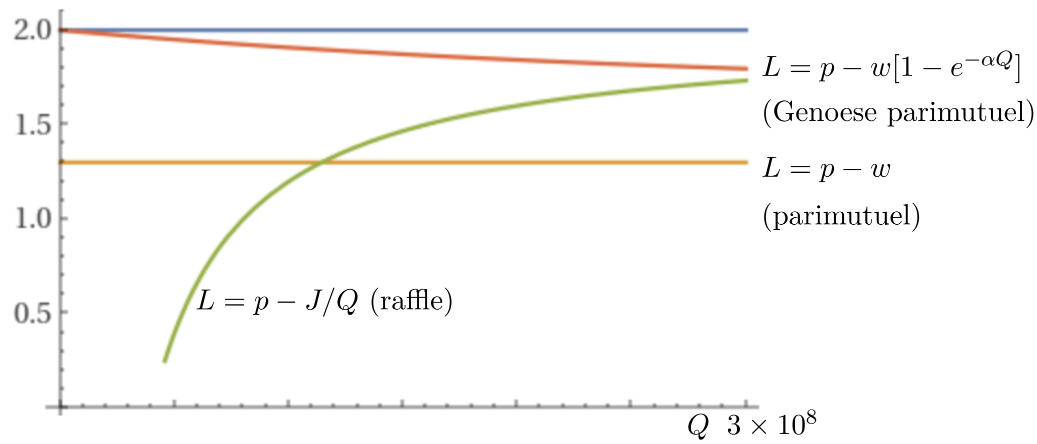

Figure 1: **Inverse Supply Curves of Classic Lotteries.** Given $p = \$2$ tickets, we plot the loss $L(Q)$ for the increasing, infinitely elastic, and falling inverse supply curves of a raffle lottery (with jackpot $J = \$80M$), a parimutuel lottery with pay rate $w = 0.32$, and a parimutuel Genoese lottery with win chance α (respectively).

are incompatible with the data (Appendix A).[6] The thrill is secured by buying a single ticket, for simplicity.[7] Those enjoying a positive thrill will play for a low enough expected monetary *loss* (ticket price less expected lotto winnings). Integrating over all potential buyers' thrill values yields a standard downward-sloping inverse demand function. A mass Q of tickets comes from buyers with thrill at least $\Lambda(Q)$. Assume Λ is twice differentiable, and can be negative.

Next, a supply function summarizes lottery ticket purchase opportunities. It does not reflect active choices but rather embeds the lottery rules. The *inverse supply curve* $L(Q|J)$ maps the quantity Q of tickets sold in a given draw to the expected monetary loss implied by the laws of probability and a jackpot $J \geq 0$. 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.[8]

Supply need not slope up, but it does for a *raffle lottery* which pays a fixed jackpot J to exactly one buyer. Its inverse supply $L(Q|J) = p - J/Q$ is

[6] We develop the model with risk loving or risk aversion in Appendix A.

[7] The same aggregate demand arises if a consumer acts as any sum of one ticket buyers. Buyers may purchase more tickets at higher jackpots — like a limit demand schedule. We only need demand driven by linear expected payoffs. We cannot identify finer preferences over ticket gambles.

[8] These are fixed prizes awarded for matching 3–5 of the six winning numbers. Focusing on risk neutrality, we don't precisely model these prizes or parimutuel gambles than by their expected value here, but in the appendix, properly account for variance without risk neutrality.

monotonically rising, since everyone's chance of winning falls in the sales. A raffle loses money if fewer than $Q = J/p$ tickets are sold. A *parimutuel* lottery (like a horserace) pays a fixed share w of ticket sales and so never loses money. This yields an infinitely elastic inverse supply $L(Q|J) = (1-w)p$ in Figure 1.

In a *Genoese lottery*, ticket buyers pick their own number combinations, and all winners equally share the prize — a jackpot or a parimutuel share $w \leq p$ of ticket revenue. If each ticket wins with probability π, no one wins with chance $(1-\pi)^Q = e^{-\alpha Q}$, where $\alpha = -\log(1-\pi) \approx \pi$.[9] We call α the *win chance*, and $1/\alpha$ the *odds*. A *Genoese parimutuel* has a falling supply curve $L(Q|J) = p - w[1-e^{-\alpha Q}]$, as wQ is not won with chance $e^{-\alpha Q}$ (Figure 1).[10]

A fraction τ (the *take rate*) of ticket revenues is withheld for lesser prizes, lottery expenses, and revenue. A *rollover lottery* builds on a Genoese lottery, adding the untaxed ticket sales to the next draw as an *inherited jackpot*. The *final jackpot* is the inherited jackpot J plus a share $1-\tau$ of the current draw revenues not withheld by the lottery, or $J + (1-\tau)pQ$. The expected winnings depend on the final jackpot accounting for a possible tie. The jackpot resets to $\underline{J} > 0$ if someone wins the jackpot. No lottery draw loses money if $p\tau \geq w$.[11]

Consumers act myopically, as they cannot individually impact the future trajectory of jackpots. Only the inherited jackpot impacts current ticket sales.

Theorem 0 (Risk Neutrality) *The inverse supply for a rollover lottery is:*

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

For an intuition, 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, divided by the quantity Q of tickets sold.[12,13]

[9] The approximation error $O(\pi^2)$ is negligible, since $\pi = O(10^{-8})$.

[10] A *Genoese raffle* has a rising supply curve: $L(Q|J) = p - (J/Q)[1-e^{-\alpha Q}]$.

[11] In our recent Powerball and Mega Millions data, tickets sell for $p = \$2$, the tax rate is $\tau = 0.66$ of revenues, but some of this is used to pay for secondary prizes amounting to $w = 0.32$. The Powerball odds are $1/\alpha_P = 292{,}201{,}338$ to one, and Mega Millions odds are $1/\alpha_M = 302{,}575{,}350$ to one.

[12] Cook and Clotfelter (1993) derive this formula in the actual integer tickets world accounting for all ways the jackpot can be multiply shared among 2,3,4,... winners. Amazingly, Euler first published related calculations in 1862 (Bellhouse, 1991). See our Addendum for our proof.

[13] We assume that players choose numbers at random. Although there is evidence that this is not the case (see, e.g., Thaler and Ziemba, 1988), this is unlikely to matter quantitatively: 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.

Each rollover draw simply sums a Genoese parimutuel and a Genoese raffle lottery. Supply depends on quantity Q via two channels. Higher Q inflates prize money, as the chance $(1+\alpha Q)e^{-\alpha Q}$ of a shared prize rises to one.[14] The inverse supply curve falls in Q at smaller jackpots J when the first force dominates, and rises in Q when a shared prize is more likely at bigger J (Figure 2).

Shared prizes plays a critical role in Theorem 0. If ticket buyers blithely ignore the multiple winner prospect, they think with chance α they win a jackpot $J + p(1-\tau)Q$. The rollover lottery would thus have a *naive loss*[15]

$$\hat{L}(Q|J) = p - w - \alpha[J + p(1-\tau)Q] \tag{2}$$

Naively ignoring multiple wins leads one to understate the loss, or $\hat{L} < L$, since $1 - e^{-\alpha Q} < \alpha Q$.[16] The deviation grows with quantity, since the conditional chance that a winner is the only one equals $\alpha Q e^{-\alpha Q}/(1-e^{-\alpha Q})$, by Bayes rule. This falls to zero in Q, and for instance equals $1/(e-1)$ at $Q = 1/\alpha_P \approx 292M$ for Powerball. So a naive Powerball user acts as if the jackpot is inflated by 78%. We statistically reject that buyers envision a naive loss function in §4.6.

Inverse supply shifts down as the inherited jackpot J rises. For the expected loss falls as J rises at each Q. In particular, supply is initially positive for small J and negative for larger J. Also, as the jackpot J rises, 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, inverse supply tends to $p\tau - w$ as $Q \to \infty$, since a large enough number of tickets sold for a draw swamps any jackpot J. See Figure 2.

Theorem 1 (Inverse Supply Changes in Quantity and Jackpot)
(*a*) *Inverse supply is positive at $Q = 0$ iff the inherited jackpot is $J < (p-w)/\alpha$.*
(*b*) *For any inherited jackpot J, supply tends to $p\tau - w$ as $Q \to \infty$.*
(*c*) *At inherited jackpot $J = 0$, supply is both globally falling and strictly convex.*
(*d*) *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.*
(*e*) *If $J > 2p(1-\tau)/\alpha$, then supply is rising and strictly concave.*

[14] Namely, one minus the chance of zero or one winner, each computed by the Poisson distribution.
[15] This is reminiscent of naivety in Eyster and Rabin (2005, 2010).
[16] Clear for $\alpha Q > 1$. If not, invert $e^x = 1 + x + x^2/2 + \cdots < 1 + x + x^2 + \cdots = 1/(1-x)$ for $|x| < 1$.

Proof of Theorem 1: For the vertical intercept of supply in part (a), consider:

$$L(0|J) = p - w - [J/Q + p(1-\tau)][1 - e^{-\alpha Q}]\Big|_{Q=0} = p - w - J\alpha. \tag{3}$$

Part (b) follows by inspecting (1). Proofs of (c)–(e) rely on the derivative of (1):

$$L'(Q|J) = J/Q^2 - e^{-\alpha Q}\left([J/Q + p(1-\tau)]\alpha + J/Q^2\right). \tag{4}$$

Using Taylor series, in Appendix B.1, we find thresholds $\overline{Q}(J) \geq \underline{Q}(J) \geq 0$, such that supply shifts from falling to rising at $\underline{Q}(J)$, and has an inflection point $\overline{Q}(J)$. Also, $0 < \underline{Q}(J) < \overline{Q}(J)$ for small jackpots J, $0 = \underline{Q}(J) < \overline{Q}(J)$ for larger jackpots, and $\underline{Q}(J) = \overline{Q}(J) = 0$ for the largest jackpots. □

3.2 Market Equilibrium and Comparative Statics

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

$$L(\mathcal{Q}(J)|J) \equiv \Lambda(\mathcal{Q}(J)). \tag{5}$$

Since a buyer must correctly forecast how many tickets sell to compute the ticket loss and thus to know if he should buy, a lottery equilibrium is a rational expectations equilibrium.[17] A buyer who ignored current ticket sales would think the jackpot is lower, and so the marginal ticket buyer would not purchase.

We now provide conditions under which the equilibrium is unique and *stable* — so 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 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, discouraging marginal ticket purchases. So buyers should learn the equilibrium.

Theorem 2 *If inverse demand $\Lambda(Q)$ exceeds inverse supply (3) at $Q=0$, and is below $p-w$ for large Q, there is a unique equilibrium $\mathcal{Q}(J)$, and it is stable.*

[17] This is supported by the good fit of the model in our empirical applications (Sections 4–6). Forrest et al. (2000) find empirical evidence for rational expectations for the UK national lottery.

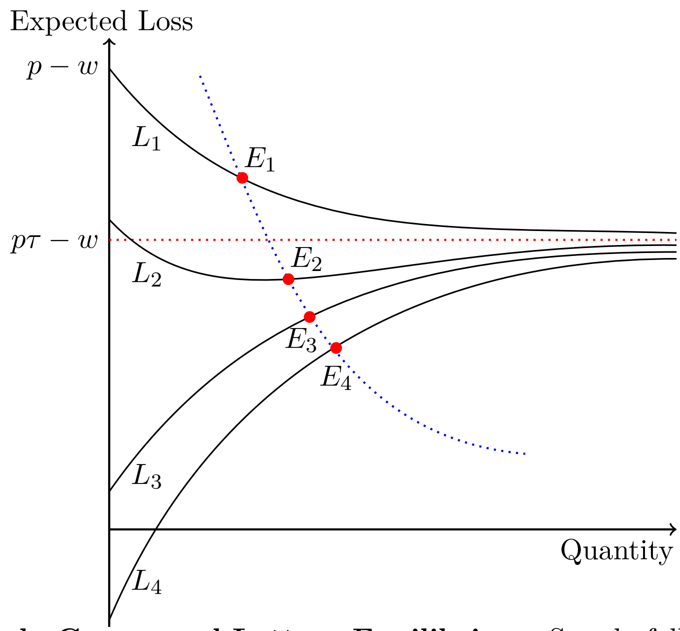

Figure 2: **Supply Curves and Lottery Equilibrium.** Supply falls as the jackpot J rises. Schematic inverse supply curves L_i for $J_1 < J_2 < J_3 < J_4$, in Theorem 1. For $J < 2p(1-\tau)/\alpha$, supply is convex then concave, and decreasing then increasing; supply is concavely increasing for $J > 2p(1-\tau)/\alpha$. Equilibrium is unique and stable as demand is steeper than any falling supply curve (see E_1, E_2), as in Theorem 2.

The proof in Section B.2 uses the intermediate value theorem, a falling demand curve, and a unique supply inflection point. We will see empirically that the assumptions in Theorem 2 are met by Powerball and Mega Millions.

Uniqueness ensures meaningful comparative statics, and helps estimation.

Theorem 3 *The equilibrium quantity $\mathcal{Q}(J|\tau, \alpha)$ falls*[18] *in the take rate τ, and rises in the inherited jackpot J, and win chance α.*

As Samuelson's correspondence principle implies, stability delivers intuitive comparative statics predictions. Theorem 3 also motivates our identification strategy in Section 4.4: exogenous jackpot increments shift the inverse supply down and thus help trace out the demand curve, as seen in Figure 2.

Proof of Theorem 3: Easily, from (1), the supply loss falls in the inherited

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

jackpot J:

$$L_J(Q|J) = -[1 - e^{-\alpha Q}]/Q < 0. \tag{6}$$

Differentiate (5) in J, using (1). Since $L_J < 0$ and $\Lambda'(\mathcal{Q}(J)) - L'(\mathcal{Q}(J)|J) < 0$ by stability:

$$\mathcal{Q}'(J) = \frac{L_J(\mathcal{Q}(J)|J)}{\Lambda'(\mathcal{Q}(J)) - L'(\mathcal{Q}(J)|J)} > 0. \tag{7}$$

We can proceed likewise, and compute supply derivatives in α, p, and τ from (1), and deduce formulas for $\mathcal{Q}_\alpha > 0 > \mathcal{Q}_\tau$ at any stable equilibrium. □

We highlight that this market theory does not merely sign the partial effects of parameters. Rather, it gives precise closed forms for all derivatives. This helps us in Sections 5 and 6 to perform normative and positive analysis.

4 Estimation of Lottery Demand

We now use our model to construct and estimate an empirical model of the two national US rollover lotteries, Powerball and Mega Millions. Because these two lotteries could be substitutes for buyers, we model them jointly in this section.

4.1 Data

We obtain data on draw-level prizes and sales for Powerball and Mega Millions, scraped from official lottery worksheets. We also collect state-draw-level sales data from the website LottoReport.com. We estimate the model on the period from October 19, 2013, to December 26, 2020 (748 draws). Some major changes occurred to lottery rules during this period: for instance, after October 2015 Powerball changed its rules by considerably lengthening the odds of winning. Moreover, 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. See Appendix C for more details on data construction, lottery rules, and summary statistics.

4.2 Empirical Model

Our equilibrium model takes the demand and supply functions as key inputs. Supply is fully determined by the lottery rules and so requires no estimation. We model demand nonparametrically. This allows us to estimate the functions flexibly, which is important in our setting, as the shape of demand — and particularly its curvature — is known to affect market equilibrium. 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).

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

$$\begin{aligned} q_{MM,s,t} &= \sigma_{MM}(\delta_{MM,s,t}, \delta_{PB,s,t}, \lambda_{MM,t}, \lambda_{PB,t}) \\ q_{PB,s,t} &= \sigma_{PB}(\delta_{PB,s,t}, \delta_{MM,s,t}, \lambda_{PB,t}, \lambda_{MM,t}), \end{aligned} \tag{8}$$

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

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

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 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.

4.3 Identification

We now discuss what variation in the data identifies the model. We focus on the residual demand of Powerball, and drop state subscripts.[19] Our point of departure is the standard empirical analysis of market equilibrium. In this context, 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.

Due to the nature of our empirical context, where supply is a mechanical by-product of lottery rules, one standard econometric endogeneity concern is ruled out: 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 render the identification of demand non-trivial. Since the supply curve is not flat, demand shocks will be correlated with a lottery's expected loss via the equilibrium mechanism.

To tackle this, we use instrumental variables. The instrument must impact the expected loss and be exogenous to unobserved demand shocks. A strong predictor of the level of the jackpot, and so the expected loss, is whether a rollover just happened. But directly using an indicator $win_{j,t-1}$ for whether someone won at draw $t-1$ as an instrument for the expected loss at time t is not a valid strategy. A positive shock to $\xi_{j,t-1}$ increases sales $Q_{j,t-1}$, and thus the chance that someone wins the lottery at draw $t-1$. So $\xi_{j,t-1}$ and $win_{j,t-1}$ are correlated. If demand shocks $\xi_{j,t}$ are serially correlated across time periods, $win_{j,t-1}$ is also correlated with $\xi_{j,t}$, violating exogeneity.

Rather than directly use $win_{j,t-1}$, we leverage our knowledge of supply to construct an instrument for the expected loss. Let $\mathcal{I}_t$ denote the time t information set:

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

While $E\left[win_{j,t-1}|\mathcal{I}_{t-1}\right]$ clearly depends on $Q_{j,t-1}$, which in turn depends on $\xi_{j,t-1}$, 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

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

[19] We use one product intuition although there are two products in the market we study; all of our discussion goes through in the two-product setting when interpreting demand as residual demand.

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

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

where $\mathbf{x}_t$ are observed exogenous characteristics. Consistent with Berry and Haile (2014), 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 consistent with economic intuition. (See Table 5 in Appendix C.4 for the full set of results.)

4.4 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 (8) 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:

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

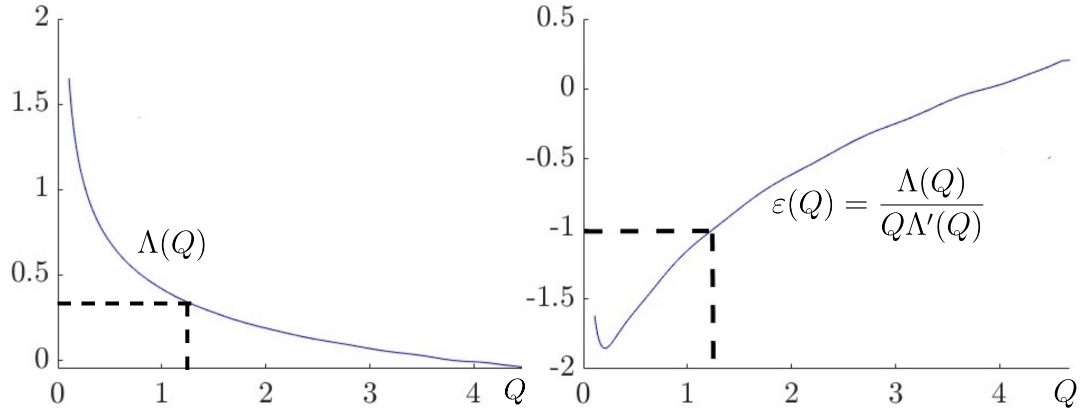

Figure 3: **Inverse Demand Curve and Own-Loss Elasticity for Powerball.** Left: the estimated inverse demand curve for Powerball (fixing all demand drivers other than the expected loss at median values). Right: the signed own-loss elasticity against expected loss. *Absolute demand elasticity* $|\varepsilon|$ *is falling in the quantities above* $20M$. Quantities on the x-axis are expressed in hundreds of millions.

where θ are the coefficients to be estimated, and where $\{b_{v,m}\}_{v=0}^{m}$ are univariate Bernstein basis polynomials of degree m.[20] The overall approximation degree is thus $4m$. We estimate (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 (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. See Compiani (2022) for details on the implementation of the estimator.[21]

4.5 Estimation Results

Panel (a) of Figure 3 plots Powerball's estimated aggregate inverse demand curve (fixing demand drivers other than the own expected loss at their median values). Table 2 shows the mean elasticities of aggregate demand to own- and

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

[21] 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 equal. 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 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.

	λ_{MM}	λ_{PB}
Q_{MM}	-1.54	0.01
	(0.09)	(0.04)
Q_{PB}	0.01	-1.50
	(0.01)	(0.14)

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 the expected loss for one U.S. national lottery. Standard errors are in parenthesis below each elasticity.

cross-expected loss, for the model imposes substitution between the lotteries. We set $m = 2$, corresponding to a Bernstein approximation of degree 8.[22]

The absolute own-loss elasticities exceed one. 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, panel (b) in Figure 3 shows the relationship between tickets sold and the estimated own-loss elasticity for Powerball across draws. The relationship between sales and elasticity is nonlinear, with absolute elasticity growing for sales level up to $20M$, then decreasing in sales. This pattern indicates that more restrictive models (such as a standard log-log regression with constant elasticity) would likely lead to misspecification.

Finally, we verify Theorem 2's premise that guarantees existence of a unique equilibrium: At every equilibrium in the data, the derivative of the (residual) inverse demand for each draw is below the derivative of inverse supply.

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

	J ($ million)			
$\mathcal{Q}(J)$ (million)	100	200	300	400
Rational	16	28	60	130
Naive	17	28	44	75
Risk aversion	13	15	15	16
Data	(16, 17)	(28, 35)	(48, 89)	(85, 222)

Table 3: **Comparison between Model Predictions and Data.** In the first three rows, we report the equilibrium quantities $\mathcal{Q}(J)$ implied by different models. The data row reports 95% confidence intervals around the Nadaraya-Watson kernel regression estimate; we chose the bandwidth parameter via leave-one-out cross validation. Risk aversion is very slight: the Arrow-Pratt risk aversion coefficient is 10^{-8}.

4.6 Model Fit

To assess fit, we calculate the equilibrium sales levels predicted by the model at different inherited jackpots and compare them with the data (Table 3). Our model with risk neutral and rational consumers fits the data very well.

We also consider two alternative models. First, we assume that gamblers do not account for the possibility of ties, giving rise to the naive inverse loss function $\hat{L}$ in (2). To assess the empirical performance of this model, we re-estimate demand as a function of the naive loss and use it along with the naive supply to predict equilibrium outcomes. We find that for low jackpot values, the rational and naive models yield similar predictions in line with the data. But the naive model underestimates the equilibrium quantities for larger jackpots relative to the data, while the predictions of the rational model are within the data confidence intervals. We repeat an analogous exercise for a model with slightly risk averse (and rational) gamblers. Because in this case the inverse supply is essentially flat (we will return to this in Figure 4), the model predicts close to no change in sales as the inherited jackpot grows. In sum, the model with risk neutral and rational consumers better fits the patterns in the data.

4.7 Does it Ever Pay to Play Powerball?

One might wonder if the jackpot ever rises so high that it is profitable to play the lottery even with zero thrill. Of course, this requires that supply go negative, for which it suffices that it starts negative, or the jackpot exceeds $2p(1-\tau)/\alpha$

by Theorem 1. Given an empirical context that gives us specific values of lotto parameters and demand, we can check these conditions. In the case of the US rollover lottery Powerball, we have $p(1-\tau)/\alpha = \$400M$ (recalling $\tau = 0.65$ in footnote 11). Given an estimated demand curve, such as the one for Powerball shown in Figure 3, we can see that nowhere in the range of past jackpots it was profitable to play the lottery without a positive thrill.

5 Lotto as the Gambler

Governments run lotteries to raise revenue, despite the social costs. We now turn tables, and consider the profit objective of the lottery owner Lotto. We focus on data from Powerball. We explore how market quantity responds to the inherited jackpot,[23] and how this impacts Lotto's rollover design.

5.1 The Rollover Lottery Makes Lotto a Risk Loving Gambler

We first analyze the random rollover process as a gambling opportunity for Lotto. In principle, Lotto could gamble using jackpot money for a single period — e.g. flipping a fair coin and either adding or subtracting a dollar for a single draw. In this case, $(p-w)\mathcal{Q}(J)$ acts like Lotto's utility function. We prove that Lotto is risk loving until the jackpot gets high enough after many rollovers.

Theorem 4 (Lotto Loves Rollover Risk) *Lotto is risk loving if demand is convex, the inherited jackpot is* $J < 2p(1-\tau)/\alpha$*, and quantity is* $Q < 2/\alpha$*.*

Consistent with Table 1, Theorem 4 gives three sufficient conditions for ticket sales to increase at an increasing rate: convex demand, $J < 2p(1-\tau)/\alpha = \$400M$ and $Q < 2/\alpha = 580M$. This is summarized in the MR_J plot in Figure 5.

Proof of Theorem 4: We consider the *relative risk aversion* measure $\mathcal{Q}''(J)/\mathcal{Q}'(J)$:

$$\frac{J\mathcal{Q}''(J)}{\mathcal{Q}'(J)} - \frac{J\mathcal{Q}'(J)}{\mathcal{Q}} > 0 \tag{9}$$

Abbreviate inverse supply less inverse demand $\Gamma(Q|J) = L(Q|J) - \Lambda(Q)$. Write (7) as $\mathcal{Q}'(J) = -L_J/\Gamma'(\mathcal{Q}(J)|J) > 0$. Differentiate it in J using $L_{JJ} = 0$,

[23] Ours is a large unexplored human-designed Markovian market (Rosen et al., 1994), where the inherited jackpot is the state variable.

from (6):

$$\mathcal{Q}''(J) = L_J \frac{\mathcal{Q}'(J)\Gamma''(\mathcal{Q}(J)|J) + \Gamma'_J(\mathcal{Q}(J)|J)}{\Gamma'(\mathcal{Q}(J)|J)^2} \tag{10}$$

Substitute expressions (7) and (10) for $\mathcal{Q}'(J)$ and $\mathcal{Q}''(J)$, using $\Gamma_J = L_J$:

$$\frac{J\mathcal{Q}''(J)}{\mathcal{Q}'(J)} - \frac{J\mathcal{Q}'(J)}{\mathcal{Q}} = \frac{J[L_J\Gamma''(\mathcal{Q}(J)|J) - \Gamma'_J(\mathcal{Q}(J)|J)\Gamma'(\mathcal{Q}(J)|J)]}{\Gamma'(\mathcal{Q}(J)|J)^2} + \frac{JL_J(\mathcal{Q}|J)}{\mathcal{Q}\Gamma'(\mathcal{Q}|J)}$$

Differentiate $L_J(Q|J) = -[1 - e^{-\alpha Q}]/Q < 0$ from (6) in Q to get $\Gamma'_J(Q|J) = 1/Q^2 - e^{-\alpha Q}(\alpha/Q + 1/Q^2)$, as $\Lambda(Q)$ is independent of J. So (9) is true iff

$$[L'(Q|J) + QL''(Q|J)] - [\Lambda'(Q) + Q\Lambda''(Q)] < 1 + \frac{\alpha Q}{e^{\alpha Q} - 1} \tag{11}$$

Now, the jackpot bound gives $L'(Q|J) + QL''(Q|J) < 0$ (see Claim B.1). Also, $\Lambda''(Q) > 0$ if demand is convex. The demand slope $\Lambda'(Q) > 0$ but vanishing in Q if $|\varepsilon| > 0$ is falling.[24] Lastly, the right side of (11) exceeds $3 - \alpha Q \geq 1$.[25] □

A compelling case for risk neutrality comes at large jackpots. Intuitively, even slightly risk averse buyers should eventually care little about increments of immense jackpots (Rabin, 2000). The risk aversion coefficient 10^{-8} almost erases the difference between jackpots \$100M and \$500M. Slight risk preference, conversely, explodes the impact of large jackpot rises. See Appendix A.

But ticket sales increase at an increasing rate. The log-log plot in Figure 4 is the best functional summary of how fast ticket sales rise in the jackpot. We prove in Appendix Claim B.2 that it is also convex under Theorem 4's premise.

5.2 Should Lotto Cap the Rollovers?

Powerball commits to rolling over the jackpot without a winner. Is this wise? Define a *\$ rollover* as one dollar added to the jackpot J for a single draw and then withdrawn if not won. The rollover chance $e^{-\alpha\mathcal{Q}(J)}$ is large at low market quantities, and thus at low jackpots J, by Theorem 3. The cost of a dollar added to the jackpot J for one draw is $MC_J = 1 - e^{-\alpha\mathcal{Q}(J)}$, namely, the chance that the jackpot is won. Sales rise with a larger jackpot: Given per ticket

[24] Given $-Q\Lambda'(Q) = -\Lambda(Q)/\varepsilon(Q) < 2/|\bar{\varepsilon}|$ and the loss $\Lambda(Q) \leq p = 2$, we have $-\Lambda'(Q) < 2/(|\bar{\varepsilon}|Q)$ on $[0, \bar{Q}]$, where $|\bar{\varepsilon}| > 0$ is the minimum elasticity on $[0, \bar{Q}]$.

[25] For if $x = \alpha Q < 2$, then $e^x = 1 + x + x^2/2 + x^3/3! + \cdots < 1 + x(1 + (x/2) + (x/2)^2 + \cdots) = 1 + x/(1 - x/2)$. Substituting, we get $1 + \alpha Q/(e^{\alpha Q} - 1) > 1 + 1 - \alpha Q/2$

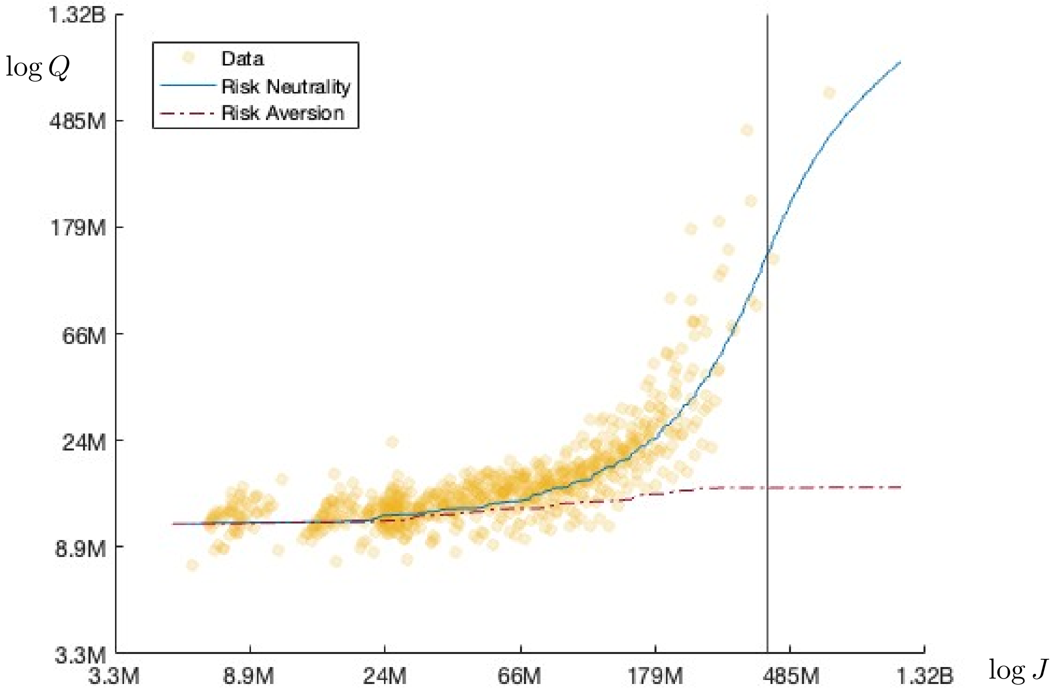

Figure 4: **Jackpot Runups as Seen in Log-Log Space.** We plot the log of our theoretical lottery sales, and the scatter of actual log sales, against log jackpots. The dashed curve is the prediction for a small risk aversion coefficient 10^{-8}. The fit with risk neutral preferences is excellent. It is convex left of the vertical line at \$409M.

profits $\tau p - w$, new revenue from that dollar is:

$$MR_J = (p\tau - w)\mathcal{Q}'(J) = \frac{(\tau p - w)L_J(\mathcal{Q}(J)|J)}{\Lambda'(\mathcal{Q}(J)) - L'(\mathcal{Q}(J)|J)} = \frac{(\tau p - w)(1 - e^{-\alpha \mathcal{Q}(J)})}{1/\eta - 1/\varepsilon} \tag{12}$$

from (7) and (6). Once a rollover is unprofitable (i.e. $MR_J < MC_J$), it remains so at higher jackpots if marginal revenue falls in J. To see that (12) falls in J:

- demand elasticity $|\epsilon|$ falls in $\mathcal{Q}$ and so in J (true in Figure 3 at $Q > 20M$)
- supply elasticity η rises in the jackpot J, which holds when $Q > 20M$.[26]

In other words, the \$-rollover grows less favorable as J rises, as $MC_J = 1 - e^{-\alpha \mathcal{Q}(J)}$ monotonically rises, *and* Lotto shifts to risk averse (MR_J falls).

[26]The denominator of (12) is positive, by stability (Theorem 2). Next, by Theorem 0, inverse supply $L(Q|J)$ clearly falls in J in (1), while $L'(Q|J)$ rises in J since by (4), $\frac{\partial}{\partial J} L'(Q|J) = 1/Q^2 - e^{-\alpha Q}\left(\alpha/Q + 1/Q^2\right) \propto e^{\alpha Q} - 1 - \alpha Q > 0$. Finally, $1/\eta = QL'(Q|J)/L(Q|J)$.

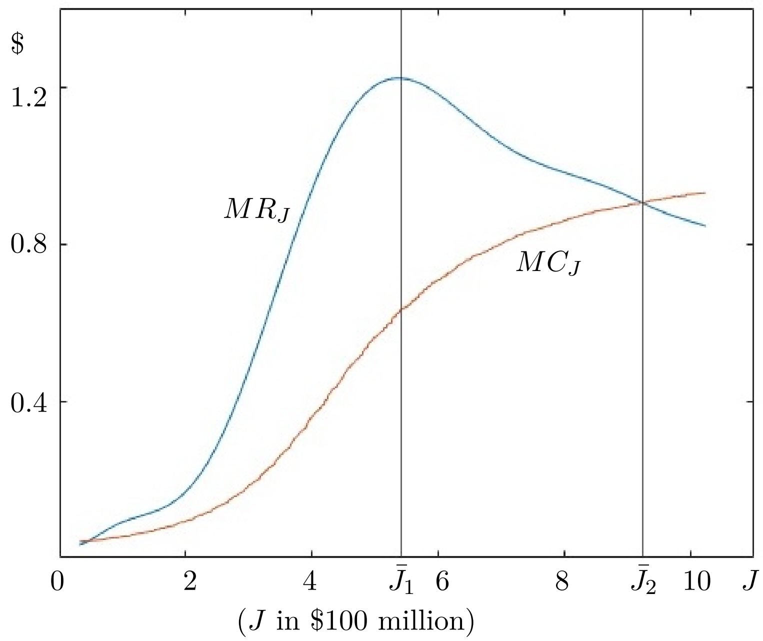

Figure 5: **Optimally Capping Powerball.** We plot the marginal revenue $MR_J = (\tau p - w)\mathcal{Q}'(J) \approx [(0.66)(2) - 0.32]\mathcal{Q}'(J)$ — the increment from the new ticket sales — and marginal cost of a \$ rollover (i.e. the chance the jackpot is won). Consistent with Theorem 4, marginal revenue is rising through $\bar{J}_1 \approx \$540M$. The rollover termination chance $MC_J = 1 - e^{-\alpha \mathcal{Q}(J)}$ rises from 5% to 92% as the jackpot rises to $\$1B$, since from Figure 4, demand is about $720M$ tickets then. Lotto finds the gains from a dollar rollover less than the costs up to $\bar{J}_2 \approx \$920M$ (Theorem 5).

Theorem 5 *A \$ rollover earns profits iff* $MR_J \geq MC_J$ *iff* $\tau p - w > 1/\eta - 1/\varepsilon$.

So while the optimal one-shot lottery happens at unit demand elasticity, this condition also turns on supply and the lottery structure. Demand is unit elastic at quantity $120M$ (Figure 3) and thus a jackpot about $\$400M$ (Figure 4), but the dollar rollover remains profitable up to the jackpot around $\$920M$. Notably, Powerball and Mega Millions have no rollover caps, Powerball could raise its average revenue if it capped rollovers. For instance, Canada's Lotto Max rollover caps out at $\$80M$, after which new funds lead to extra million dollar prizes; formally, this pushes it towards a parimutuel. Eurojackpot caps out at €$120M$, and EuroMillions' cap is €$250M$.

In brief, as the jackpot rises, Lotto shifts from a risk loving gambler playing positive expected value gambles, to a risk averse gambler playing positive expected value gambles, until he hits the point where he should end the rollover.

6 Lottery Odds Across Jurisdictions

We introduce more evidence for risk neutrality from Lotto's optimization. A widely discussed design feature of a rollover lottery is its choice of jackpot odds. For instance, after Powerball lengthened the odds in October 2017 from 175 million to one, up to 292 million to one,[27] average per-draw revenues increased from \$16.8M million to over \$26M. This was Powerball's sixth odds increase since 1992, when it had started out at 55 million to one.

All told, forty states and many countries have rollover lotteries with varying odds. For instance, EuroMillions is 140 million to one, and Canada's Lotto Max is 33 million to one. More populous states / countries have longer odds. Cook and Clotfelter (1993) first noted this correlation, but our take is unrelated.[28] Under a regularity assumption that lottery ticket demand is proportional to the population, we argue more strongly for a linear relationship, and not just a positive correlation: Optimal lottery odds are proportional to the population. For if betting behavior responds to expected payoffs, then Lotto should act as if ticket buyers care linearly about the lotto prize or win chance.[29]

For if country A is twice as populous as country B, assume that A has double the lottery ticket demand of B, for each expected loss. Precisely, inverse demand is $\Lambda_N(Q) = \underline{\Lambda}(Q/N)$ given population N, for some fixed function $\underline{\Lambda}$. For then, if the population doubles, the new win chance is $\alpha/2$, and the chance that someone wins $1-e^{-(\alpha/2)(2Q)}$ is unchanged at each loss if ticket sales double. This intuitively yields the same random rollover process for all states of all sizes.

First, the inverse supply formula (1) is $L(Q|J,\alpha) = L(NQ|NJ,\alpha/N)$. Hence, the market quantity $\mathcal{Q}_N(\cdot|\alpha)$ solving (5) obeys $\mathcal{Q}_N(\cdot|\alpha/N) \equiv N\mathcal{Q}(\cdot|\alpha)$.

We compute the expected revenue in a jackpot run-up. If no one wins in draws $0, 1, \ldots, k-1$, the next jackpot $J_k(\alpha)$ adds the untaxed past ticket sales:

$$J_k(\alpha) = (1-\tau)p\mathcal{Q}(J_0(\alpha)|\alpha) + \cdots + (1-\tau)p\mathcal{Q}(J_{k-1}(\alpha)|\alpha)$$

[27] "How Powerball manipulated the odds to create a \$1.5 billion jackpot" (Washington Post).

[28] They offered a behavioral story: "If players tend to judge the likelihood of winning based on the frequency with which someone wins, then a larger state can offer a game at longer odds but with the same perceived probability of winning as a smaller state." Our theory is, by contrast, is based on a standard rationally understanding of the gambles.

[29] Loosely, if ticket buyers were risk averse, a positive correlation might arise, but a state twice as populous with twice the lottery jackpot would not require double the odds to hold mimick incentives.

By independence of lottery draws, the probability that a new cycle starts with $k+1$ rollovers is the product of first $k+1$ rollover chances:

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

This yields a simple formula for the average lottery revenue in each run-up:[30]

$$V(\alpha) = p\tau \frac{\mathcal{Q}(J_0(\alpha)|\alpha) + e^{-\alpha \frac{J_0(\alpha)}{p(1-\tau)}} \mathcal{Q}(J_1(\alpha)|\alpha) + e^{-\alpha \frac{J_1(\alpha)}{p(1-\tau)}} \mathcal{Q}(J_2(\alpha)|\alpha) + \cdots}{1 + e^{-\alpha \frac{J_0(\alpha)}{p(1-\tau)}} + e^{-\alpha \frac{J_1(\alpha)}{p(1-\tau)}} + \cdots} \tag{13}$$

In devising jackpot odds $1/\alpha$, Lotto faces the tradeoff that that longer odds reduces immediate lottery ticket sales (by Theorem 3), but raises the chance one lands in the continuation game with a larger jackpot. If lottery ticket demand is proportionately similar across states or countries, (13) allows us to deduce that the optimal lottery win chance scales with the population:

Theorem 6 *Lotto's optimal odds* $1/\alpha$ *linearly scale with the population* N*.*

Proof: For a population N country, write (13) 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, 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. □

In other words, it is optimal for Lotto to induce the same stochastic process of rollover lotteries of different demand magnitudes.

We test the predictions of Theorem 6 from a dataset of rollover lotteries across U.S. state lotteries. Almost all states participate in Powerball and Mega Millions. These lotteries are run by state agencies and 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 obtained 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 6, as long as demand for rollover

[30] 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 $(z_0 + p_0 z_1 + p_1 z_2 + \cdots)/(1 + p_0 + p_1 + \cdots)$.

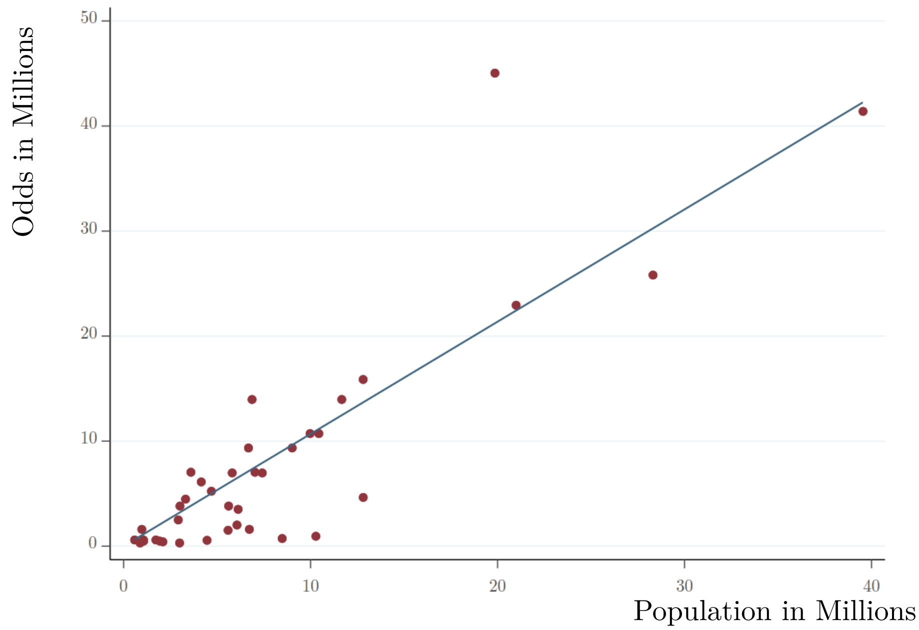

Figure 6: **Lottery Odds and Population across States.** We depict the scatter plot of rollover lottery odds against U.S. state populations. The regression line's estimated slope is 1.07 (with standard error 0.08). This is consistent with our theory that odds should be proportional to population, assuming lottery demand is.

lotteries scales proportionally with population, Lotto in each state would set odds that scale with population. We seek evidence on this by computing the relation between population and lottery odds. We report the scatter plot and a simple regression of state-level odds on population in Figure 6. It has an R^2 of 0.84. As we predict, its slope is not statistically different from unity.

In sum, the data broadly supports the linear relationship between lottery odds and market size predicted with unit slope through the origin. This is evidence that state lotteries around the US act as if guided by our implicit market model premised on risk neutrality.

7 Conclusion

Markets are economists' go-to tool for understanding transactions, since they allow separate exploration of forces impacting supply and demand. We have created a simple and novel rational market model of the dominant lottery format in the world: rollover lotteries. We don't use the lottery ticket price as the market-clearing price, but instead adopt the expected loss on tickets as an implicit price of lottery thrill. This yields a unified rational tractable model of lottery markets that can be estimated for rollover lotteries. The model fit to the data here is excellent. Ticket sales are increasing and convex in the rising jackpots, with a convex log-log relationship to them up to \$400M.

For more evidence for our model, we predict two key lottery design features. First, in its dynamic optimization, Lotto finds the rollover gamble turn sour as the chance of a jackpot win rises, and rollovers around the world have adopted caps. Powerball too should adopt a jackpot cap. Secondly, rollover odds should rise linearly in the population. The fit across forty states is excellent.

Our model is premised on a simple answer to Friedman and Savage's puzzle of why people both gamble and insure: people who gamble derive an individual-specific thrill *at the extensive margin* from the experience. (Might risk aversion over losses likewise formally act as a negative thrill at the extensive margin?)

Our model does not require risk neutrality, but this is the most tractable case. And we focused on this case, after showing that even the slightest risk preference or aversion in the literature is completely at odds with the data. The model therefore provides new evidence from 160 million Americans playing Powerball for Rabin (2000)'s arguments against risk aversion in gambling.

We don't speak to preferences any deeper than linearity in utility for money gambles *at the intensive margin*. Buyers might well have falling thrill thresholds for more tickets, as this would yield the same aggregate demand. With our model, data cannot possibly offer finer insight into preferences over gambles.

Our model invites lottery design inquiry and can predict changes in sales and revenues when, say, the tax rate or secondary prizes adjust. But towards a focused paper, we avoid applications that speak to studied lottery questions.

We are unaware of implicit market models similar to ours. Our idea of using a transaction loss as a market-clearing price might suggest parsimonious models of other economic settings without an explicit market, such as crime.

References

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.

Board, S. (2008): "Durable-Goods Monopoly with Varying Demand," *Review of Economic Studies*, 75, 91–413.

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 multi-product demand: A nonparametric approach," *Quantitative Economics*, 13, 545–591.

Conlisk, J., E. Gerstner, and J. Sobel (1984): "Cyclic Pricing by a Durable Goods Monopolist," *Quarterly Journal of Economics*, 99, 489–505.

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

Eyster, E. and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623–1672.

——— (2010): "Naive Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2, 221–243.

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

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

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

Hendel, I., A. Lizzeri, and N. Roketskiy (2014): "Nonlinear Pricing of Storable Goods," *American Economic Journal: Microeconomics*, 6, 1–34.

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.

Post, T., M. J. van den Assem, G. Baltussen, and R. H. Thaler (2008): "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show," *The American Economic Review*, 98, 338–71.

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.

Rosen, S., K. M. Murphy, and J. A. Scheinkman (1994): "Cattle Cycles," *Journal of Political Economy*, 102, 468–492.

Sobel, J. (1984): "The Timing of Sales," *Review of Economic Studies*, 51, 353–368.

——— (1991): "Durable Goods Monopoly with Entry of New Consumers," *Econometrica*, 59, 1455–85.

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

Varian, H. R. (1980): "A Model of Sales," *The American Economic Review*, 70, 651–59.

Appendix A Risk Preference or Dispreference

We now flesh out a larger model with risk preference or aversion, with our risk neutral model as a special case. We have so far not specified how many tickets delivered the lottery thrill, as it was inessential. But now to derive the supply curve, assume lottery buyers buy a single ticket. Assume a Bernoulli utility function of the money value of thrill plus money x. For any initial wealth $I > 0$, the final wealth is $I - p$ plus thrill and winnings.

By positing the same small gamble at all wealth levels, Rabin (2000) *de facto* assumes *constant absolute risk aversion* (CARA). For coherence and mostly for simplicity, we posit utility functions of the form $u_r(x) = (1 - e^{-rx})/r$, where $r \neq 0$. This as usual captures CARA if $r > 0$, *constant absolute risk loving* (CARL) if $r < 0$, and is risk neutral at the $r = 0$ limit 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.

To derive supply, we focus on the marginal ticket buyer indifferent about buying a ticket that wins with chance π. As usual, let α solve $(1-\pi)^Q = e^{-\alpha Q}$. Assume smaller prizes have expected value w and variance σ^2. Equate his expected utility from buying and not buying. The expected utility of the marginal ticket buyer replaces the thrill by the loss $L(Q|J)$, i.e. $1/r$ times

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

As the utility of not buying is $\frac{1}{r}(1 - e^{-rI})$, indifference does not depend on I:

$$e^{-r[p-L(Q|J)]} = (1-\pi)e^{-r[w-\frac{1}{2}rw\sigma^2]} + \pi e^{-\alpha Q}\sum_{k=0}^{n} C(Q,k)\pi^k e^{-r\frac{[J+p(1-\tau)Q]}{k+1}} \qquad (14)$$

We now solve this for the loss (thrill) $L(Q|J)$ needed to sell Q tickets.

Theorem 7 *With CARA or CARL utility* $u_r(x)$, *inverse supply is:*

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

Also, in the limit $r \to 0$, *the lottery loss* $L^r(Q|J)$ *tends to the risk neutral loss formula* $L(Q|J) = p - w - [J/Q + p(1-\tau)][1 - e^{-\alpha Q}]$ *in Theorem 0.*

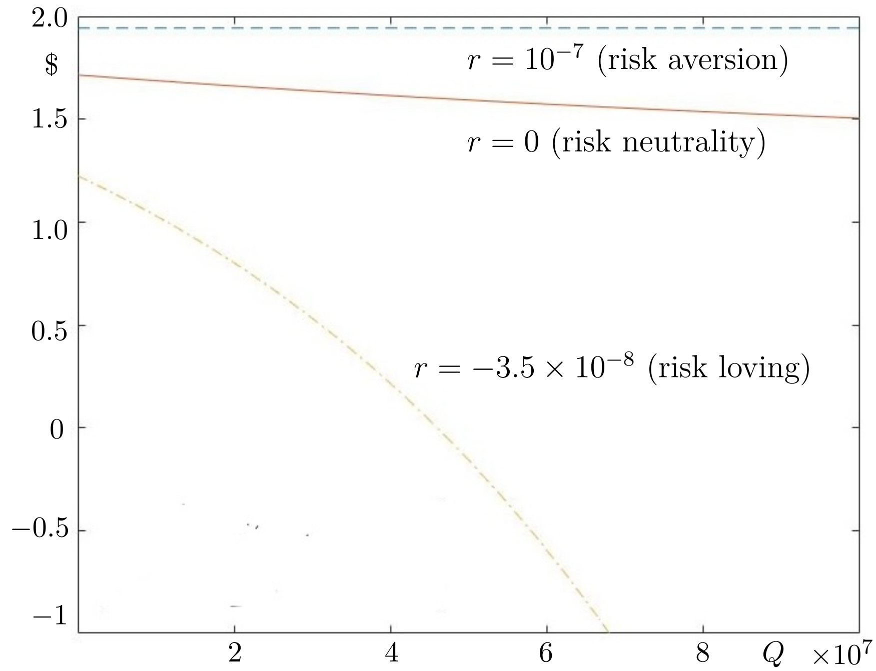

Figure 7: **Inverse Supply Curves with Risk Preference.** Plotted are supply curves for a typical inherited jackpot \$67M against quantity ($\times 10^7$). Inverse supply quickly becomes negative for slight risk loving preference parameters $r \geq -10^{-7}$, and is almost perfectly elastic for risk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

The limit ensures that this model subsumes our risk neutral one with $r = 0$.

Proof of Theorem 7: Take a Taylor series of (14), subtract 1, and divide by r, to get $p - L^r(Q, J) = (1 - \pi)w + \pi e^{-\alpha Q} \sum_{k=0}^{Q} C(Q, k)\rho^k \left[\frac{J+p(1-\tau)Q}{k+1}\right] + O(r)$. □

Given the enormous lottery jackpots, even the smallest absolute Arrow-Pratt risk preference coefficients in the literature yield implausible implications for losses or counterfactual predictions of ticket sales. For an illustrative exercise, we calibrate lottery to the average Powerball values. We pick lottery price, odds, and take rate so that $p(1 - \tau)/\alpha \approx 100M$. Set the inherited jackpot at $J = \$50M$ and the non-jackpot prizes $w = 0.25$.

First, consider risk preference, as captured by a fixed *negative* risk aversion parameter $r < 0$. For Powerball, this implies implausibly large certainty equivalents for typical large inherited jackpots (Figure 7). Specifically, the new inverse supply curve explodes negatively. As a result, marginal lottery players — who, say, play at a jackpot of \$40M but not \$20M — only do so with a negative thrill. So the typical lottery players immensely dislike playing, which is intuitively implausible. This argument holds for any demand function.

On the other hand, even slight risk aversion leads to implausible ticket sales predictions — echoing Rabin's (2000) conclusion, but with different logic. As seen in Figure 7, for $r > 10^{-8}$, the exponent of the last term in (7) does not vary by more than 0.1 if the jackpot grows by \$10M. Since supply is nearly perfectly elastic, large inherited jackpots have essentially no impact on ticket sales for a fixed demand curve. To explain the 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 can dispense with risk preference as an explanation for the observed behavior.

Appendix B Omitted Proofs

B.1 Curvature of Inverse Supply: Proof of Theorem 1

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

$$x = \underline{Q}(J) \quad \Leftrightarrow \quad e^{\alpha x} - 1 - \alpha x = [p(1-\tau)/(\alpha J)]x^2 \tag{15}$$

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 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 monotonically increases.
(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 (4) implies:

$$L'(Q|J) = \left(e^{\alpha Q} - 1 - \alpha Q[1 + p(1-\tau)Q/J]\right) Je^{-\alpha Q}/Q^2 \tag{16}$$

$$= \left(\frac{\alpha^2 Q^2}{2}\left(1 - \frac{2p(1-\tau)}{\alpha J}\right) + \sum_{k=3}^{\infty} \tfrac{1}{k!}(\alpha Q)^k\right) Je^{-\alpha Q}/Q^2 \tag{17}$$

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 (17). If $J < 2p(1-\tau)/\alpha$, the lead term of (17) is negative: So $L'(Q|J) < 0$ for small $Q > 0$.

Finally, the zero of (16) 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 $J < 2p(1-\tau)/\alpha$. As $p - w > 2p(1-\tau)$, supply starts positive/falling for low J, positive/rising for intermediate J, and negative/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 $L'(Q|J)$ in (17) changes sign at most once, − to +, and ends +. Inverse supply curvature turns on understanding the following root:

$$x = \overline{Q}(J) \quad \Leftrightarrow \quad e^{\alpha x} = 1 + \alpha x + \alpha^2 x^2/2 + \alpha^2 p(1-\tau)x^3/(2J) \tag{18}$$

Since $e^{\alpha x} > 1 + \alpha x + \frac{1}{2}\alpha^2 x^2 + \frac{1}{6}\alpha^3 x^3$, 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 on $[0, \overline{Q}(J)]$, and strictly concave on $[\overline{Q}(J), \infty)$, where the roots (15) and (18) 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 (4):

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

Then

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

From (18), $\overline{Q}(J)$ falls in J, and explodes at small jackpots: $\overline{Q}(J) \uparrow \infty$ as $J \downarrow 0$. Now, (21) 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 (4) and (19):

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

B.2 Unique Equilibrium: Proof of Theorem 2

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. For once inverse supply increases, it does so forever, by Theorem 1. So multiple equilibria can only happen when supply slopes down. But inverse supply is convex when it is decreasing (by Theorem 1), and so it steepens in Q. A second crossing with a falling demand curve is impossible: After one crossing, $\Lambda(Q) - L(Q|J)$ falls in Q. □

B.3 Lotto as the Gambler: Omitted Results for Section 5.1

Claim B.1 *If* $\alpha J < 2p(1-\tau)$ *then* $L' + QL'' < 0$.

Proof: Differentiate $L'(Q|J)$ from (4) to get $L''(Q|J)$. Gathering the many terms, we see that $L'(Q|J) + QL''(Q|J)$ equals $Je^{-\alpha Q}/Q$ times

$$-e^{\alpha Q} + 1 + \alpha Q + [1 - p(1-\tau)/(\alpha J)](\alpha Q)^2 + p(1-\tau)(\alpha Q)^3/(\alpha J)$$

This is negative if $1 - p(1-\tau)/(\alpha J) < 1/2$. In other words, $\alpha J < 2p(1-\tau)$.□

Claim B.2 *Given* (9), $\log \mathcal{Q}(J)$ *is an increasing and convex function of* $\log J$.

Proof: Let $j = \log(J)$ and $q = \log(\mathcal{Q})$. Put $\mathcal{J}(j) \equiv e^j$. Then $\mathcal{J}'(j) = e^j = \mathcal{J}(j)$. As all maps are increasing, the function $\mathcal{Q}(J)$ implies a function $\mathfrak{q}(j)$ with derivative

$$\mathfrak{q}'(j) = \frac{d\log\mathcal{Q}}{d\mathcal{Q}}\frac{d\mathcal{Q}}{dJ}\mathcal{J}'(j) = J\frac{\mathcal{Q}'(J)}{\mathcal{Q}} > 0 \tag{24}$$

Differentiate $\log \mathfrak{q}'(\mathfrak{j}(J)) = \log J + \log \mathcal{Q}'(J) - \log \mathcal{Q}$ in (24) in J, with $\mathfrak{j}'(J) = 1/J$:

$$\frac{\mathfrak{q}''(j)}{\mathfrak{q}'(j)} = 1 + \left[\frac{J\mathcal{Q}''(J)}{\mathcal{Q}'(J)} - \frac{J\mathcal{Q}'(J)}{\mathcal{Q}(J)}\right]$$

Since we proved in (9) that the bracketed term is positive, $\mathfrak{q}''(j)/\mathfrak{q}'(j) > 1$. $\square$

Appendix C U.S. Lotteries: Data Construction

C.1 Lottery Rules: the Fine Print

Both Mega Millions and Powerball significantly modify the basic rollover lottery mechanism.[31] For one, they advertise an annuity value for the jackpot, as opposed to cash amounts. The jackpot amounts commonly advertised on billboards 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.

The lottery authority does not commit to paying out the advertised (cash) value of the prize: rather, it pays the cash amount of the jackpot prize pool. Exceptions are two times when advertised jackpots are guaranteed: (i) in the first draw after the jackpot is won, the lottery starts 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 the jackpot rise. 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 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."

[31] 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).

C.2 Data Construction

We construct our data from different sources. We obtain information about lottery rules from official documents,[32] 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[33] 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 (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 $J_{j,t} = J_{j,t-1} + p_{j,t}Q_{j,t}(1 - \tau_{j,t})$, setting $J_{j,t-1} = 0$ if the jackpot was won in the previous draw.

C.3 Lottery Rules and Data Patterns

Table 4 shows the main lottery rules and summary statistics in our data. We show in Figure 8 the fluctuations in national sales of Powerball and Mega Millions over our sample period.

[32] See e.g. 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.

[33] 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).

	Mega Millions		Powerball	
Start date	Oct. 19, 2013	Oct. 28, 2017	Jan. 15, 2012	Oct. 7, 2015
Ticket price ($)	1	2	2	2
Format	5/75 + 1/15	5/70 + 1/25	5/59 + 1/35	5/69 + 1/26
Jackpot (avg., $ million)	60	106	59	100
Reset value ($ million)	16	40	40	40
Odds of Jackpot Win	258,890,000	302,575,350	175,223,510	292,201,338
Expected loss (avg., $)	0.61	1.42	1.33	1.36

Table 4: **Lottery Rules and Main Summary Statistics.** This table reports information on lottery rules and summary statistics. In the columns are Mega Millions and Powerball after the start date indicated. We convert annuity values to cash values using the discount rates of the lottery authority (see Appendix C for details).

C.4 First Stage Results

In Table 5 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 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.

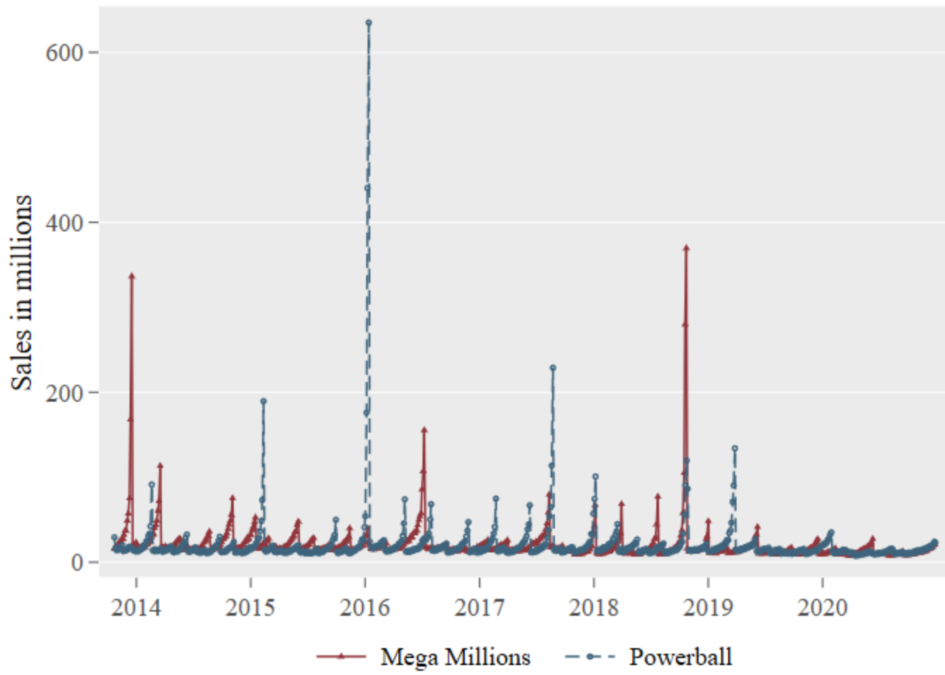

Figure 8: **Mega Millions and Powerball sales.** This shows the time series of sales (in millions of tickets sold nationwide) for Mega Millions and Powerball throughout our sample.

	(1)	(2)	(3)	(4)
Variables	Expected loss		Sales	
	Mega Millions	Powerball	Mega Millions	Powerball
Lagged residual - own lottery	0.235***	0.307***	-486,119***	-563,829***
	(0.00206)	(0.00242)	(24,902)	(28,343)
Lagged residual - other lottery	0.0139***	-0.0309***	33,175	272,966***
	(0.00185)	(0.00270)	(22,264)	(31,695)
Years since introduced - own lottery	-0.0285***	-0.0485***	-38,552	105,601
	(0.00760)	(0.0154)	(93,431)	(180,979)
Years since introduced - other lottery	-0.00200	-0.0269***	40,704	14,244
	(0.00638)	(0.00883)	(84,861)	(103,496)
Constant	0.978***	1.793***	-69,988	-58,132
	(0.0272)	(0.0524)	(338,435)	(614,176)
R-squared	0.983	0.898	0.482	0.366
F-statistic	4467	694.5	72.82	45.30

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

Table 5: **First-Stage Regressions.** We report results of a regression of losses and sales for each lottery on instruments, exogenous variables, and fixed effects for week, week part, and state. $n = 33,602$.