Research Article

Psychological Science
1–13

DOI: 10.1177/0956797615618366
pss.sagepub.com

Unexpected but Incidental Positive Outcomes Predict Real-World Gambling

A. Ross Otto[1], Stephen M. Fleming[1,2], and Paul W. Glimcher[1,3]
[1]Center for Neural Science, New York University; [2]Department of Experimental Psychology, University of Oxford; and [3]Institute for the Interdisciplinary Study of Decision Making, New York University

Abstract
Positive mood can affect a person's tendency to gamble, possibly because positive mood fosters unrealistic optimism. At the same time, unexpected positive outcomes, often called prediction errors, influence mood. However, a linkage between positive prediction errors—the difference between expected and obtained outcomes—and consequent risk taking has yet to be demonstrated. Using a large data set of New York City lottery gambling and a model inspired by computational accounts of reward learning, we found that people gamble more when incidental outcomes in the environment (e.g., local sporting events and sunshine) are better than expected. When local sports teams performed better than expected, or a sunny day followed a streak of cloudy days, residents gambled more. The observed relationship between prediction errors and gambling was ubiquitous across the city's socioeconomically diverse neighborhoods and was specific to sports and weather events occurring locally in New York City. Our results suggest that unexpected but incidental positive outcomes influence risk taking.

Keywords
decision making, risk taking, emotions, open data

Received 7/16/15; Revision accepted 10/29/15

People in a good mood are more willing to gamble—or, more broadly, to accept risky bets (Arkes, Herren, & Isen, 1988; Bassi, Colacito, & Fulghieri, 2013; Isen & Patrick, 1983; Schulreich et al., 2014). According to one important account of this phenomenon, positive mood fosters an *optimism bias*—an unrealistic expectation that positive events will occur in the future (Love, Kopeć, & Guest, 2015; Sharot, 2011; Wright & Bower, 1992). Such optimism in turn influences risk-seeking behaviors such as gambling (Gibson & Sanbonmatsu, 2004; Gilovich, 1983; Kuhnen & Knutson, 2011; Rogers, 1998). However, academic study of the linkage between mood and gambling has been limited because consequential manipulations of mood (i.e., those involving valenced, real-world outcomes) are ethically challenging, and the laboratory gambling tasks used to study risk taking are artificial, small-scale measures by design. Massive urban data sets, so-called big data, may provide a powerful alternative: the opportunity to directly study the real-world phenomena that laboratory experiments are designed to simulate. Day-to-day variations in state lottery gambling measured at the level of neighborhood ticket sales should provide just such an opportunity by elucidating the linkage between real-world risk-taking behavior and events in the environment that may influence mood.

Both laboratory and survey studies clearly indicate that a positive outcome in one's life exerts a nuanced effect on mood: Such an outcome exerts a stronger effect when it is unexpected rather than expected, and this manifests in both affective experience (Eldar & Niv, 2015; Mellers, Schwartz, Ho, & Ritov, 1997; Shepperd & McNulty, 2002) and momentary happiness (Rutledge, Skandali, Dayan, & Dolan, 2014). Neurobiological work conducted in the laboratory over the past decade emphasizes this point, identifying a critical role for *prediction errors* (the difference between expected and obtained rewards) in driving dopamine neuron firing and in learning the

Corresponding Author:
A. Ross Otto, Center for Neural Science, New York University, 4 Washington Pl., New York, NY 10003
E-mail: rotto@nyu.edu

affective structure of the environment (Bayer & Glimcher, 2005; Delgado, Li, Schiller, & Phelps, 2008; Schultz, Dayan, & Montague, 1997). It thus appears that the difference between expected and obtained outcomes drives many important behavioral phenomena inside and outside the lab. Accordingly, we set out to examine whether positive prediction errors resulting from incidental events predict real-world risky decision making by the roughly 8 million people living in New York City.

Sporting events and the weather are a source of incidental phenomena that are thought to influence mood (Cunningham, 1979; Sloan, 1989), laboratory-assessed risk taking (Bassi et al., 2013), criminal behavior (Card & Dahl, 2011), and financial market activity (Edmans, García, & Norli, 2007; Garrett, Kamstra, & Kramer, 2005; Hirshleifer & Shumway, 2003; Kamstra, Kramer, & Levi, 2003; Kamstra, Kramer, Levi, & Wermers, 2015). We hypothesized that prediction errors—which we define here as deviations from recent historical trends—stemming from these incidental, valenced outcomes would influence population-level mood, in turn predicting subtle day-to-day fluctuations in lottery gambling that would be observable at the level of a large city. Because both sporting events and sunshine occur continually and exhibit considerable intrinsic variability, we were able to assess how daily lottery gambling in a large city responds to these kinds of prediction errors. We note that our focus on deviation between outcomes and expectations contrasts with previous work examining how sports outcomes themselves drive economic behavior (e.g., Edmans et al., 2007).

The purchase rates of daily fixed-odds lottery games in New York City have a number of useful properties as an operationalization of risk-taking behavior at the level of a city. First, state-lottery gambling is pervasive: Among the 6.1 million adult residents of New York City, more than $1.3 billion of fixed-odds lottery tickets are sold annually. On average, residents of neighborhoods with notably high lottery-gambling rates spend more than $1.40 per day on these products. Second, the expected values of these lotteries, from the perspective of the decision maker, remain constant because the jackpots do not vary over time or as a function of the number of participants. Thus, there is good reason to believe that the day-to-day fluctuations in purchases of these games (averaged over a year) reflect factors extrinsic to the lottery itself and can function (after important nuisance variables are accounted for) as a proxy for, or even a direct measurement of, the risk attitudes of New York City's gambling population—which are themselves a proxy for unrealistic optimism (Wright & Bower, 1992). Finally, this massive time series of neighborhood-by-neighborhood lottery purchases affords the statistical power necessary to detect subtle changes in risk attitudes that are hypothesized to occur in the real world, given data from existing laboratory studies.

In keeping with the emerging standard for good statistical practice in these large, real-world data sets, we performed a number of exploratory analyses on the $1.3 billion of purchases in 2011. Our statistical conclusions, however, are based on mixed-effects regressions performed on the $1.3 billion of purchases in 2012. Massive data sets of this kind allow use of such highly precise replication-based approaches to avoid false positives under conditions of multiple comparisons—an issue of growing concern in small laboratory studies (Button et al., 2013). To foreshadow our results, we found that large prediction errors resulting from sports or sunshine could shift daily purchase rates up or down by up to 0.5%, and these effects were largely homogeneous across both the city's wealthier and poorer neighborhoods.

Method

Our data sources and statistical procedures are described in detail in the Method section. Information regarding the logic guiding our analyses can be found in the Results section.

New York City lottery data

We acquired daily lottery purchases, sorted by postal code (ZIP code) for the years 2011 (our exploratory data set) and 2012 (our confirmatory data set) from the New York State Gaming Commission via a Freedom-of-Information-Act request. We also obtained aggregated daily sales, across 174 ZIP codes, for all daily, non-jackpot-based lottery games available in New York state ("Numbers," "Win4," "Pick10," "Take5," "QuickDraw"). For each ZIP code, we summed the sales of these games and divided this composite by the ZIP code's adult population to control for population differences across ZIP codes (Oster, 2004) and log-transformed this value. This calculation yielded our dependent measure of purchases per adult (Figs. 1a and 1b). Our exploratory data analyses were conducted on the 2011 data. Hypotheses developed from uncorrected statistical analyses of the 2011 data set were then tested on the 2012 data with a single, prespecified set of regressions.

Sports outcomes

We used custom software to access the ESPN.com Web site and obtain the outcomes (wins, losses, and ties) of regular and postseason games played by New York City–area teams in 2011 (492 games) and 2012 (564 games). These data covered the National Football League, the National Basketball League, the National Hockey League,

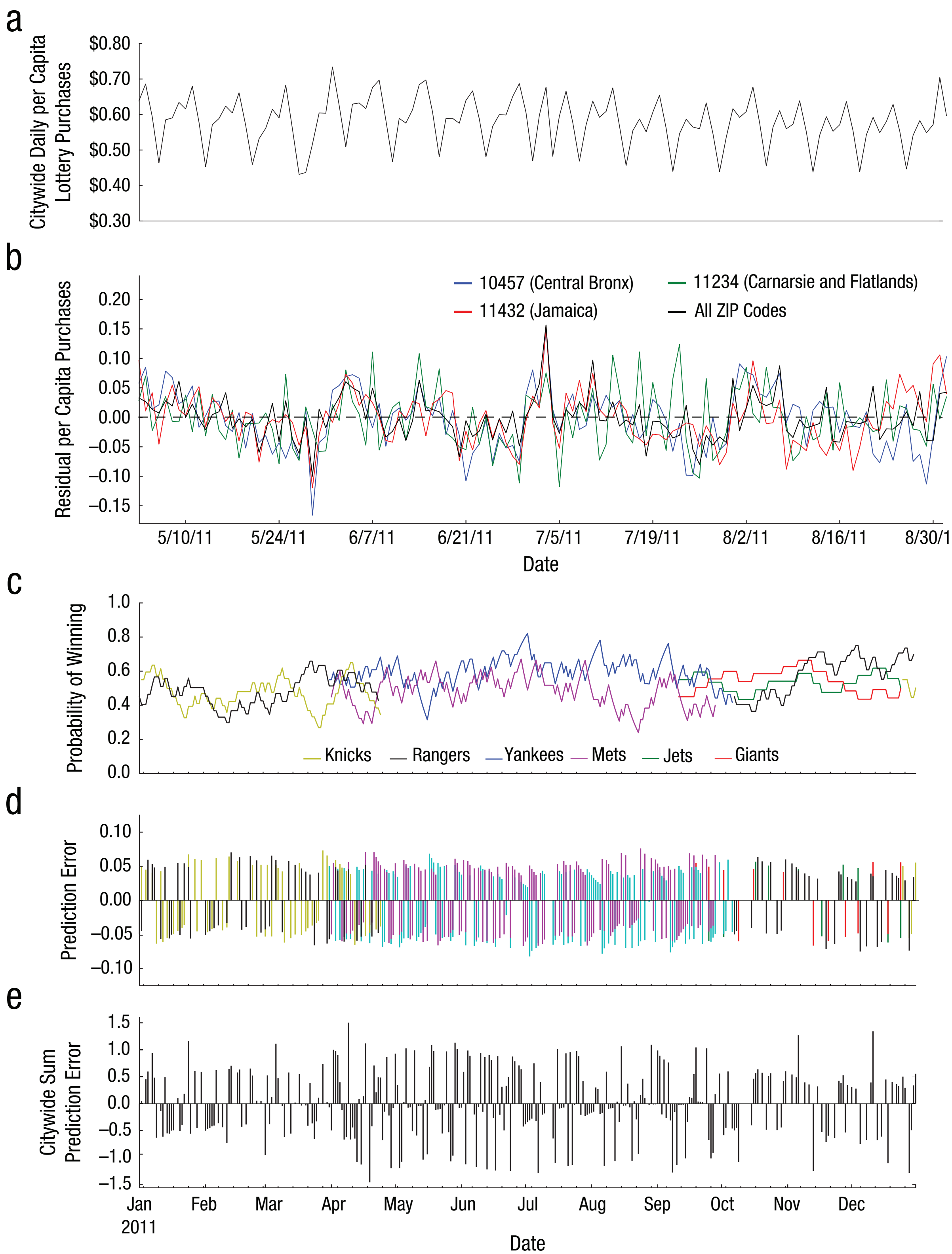

Fig. 1. The time course of lottery purchases, sports teams' probability of winning, and prediction errors based on probability of winning. Citywide per capita (a) and residual per capita (b) purchases of daily lottery tickets are graphed as a function of date for 4 months in 2011. The citywide values are averages of purchases across all ZIP codes in New York City. Residual values were obtained by controlling for a number of cyclical and noncyclical nuisance variables (see the Nuisance Variables section) and are graphed for the city as a whole and separately for three ZIP codes. Exponentially weighted estimates of the probability of winning are graphed over the course of 2011 for each of six teams (c). Each time an outcome (i.e., a win or a loss) occurred, the estimate was updated to adjust the season-based estimate of winning. Prediction errors over the course of 2011 are graphed (d) separately for the six teams and (e) citywide (i.e., summed across the six teams). Prediction errors were calculated as the deviation between the outcome (win or loss) and the estimated probability of winning for that team on that day.

and Major League Baseball. We also obtained sports outcomes for the six U.S. metropolitan regions that have the next largest media-market sizes (Nielsen Company, 2013) and are also home to three or more teams affiliated with the four major sports leagues noted earlier: Los Angeles, Chicago, Philadelphia, Dallas–Fort Worth, the San Francisco Bay Area, and Boston.

For each team, we constructed an exponentially weighted daily average of team success,

$$p_{\text{win}}(t+1) = p_{\text{win}}(t) + \alpha\left[O(t) - p_{\text{win}}(t)\right],$$

where t is the day of the year, $O(t)$ is the outcome (win = 1, loss = 0, tie = .5) on that day, and α is a recency parameter (i.e., learning rate) that makes more recent outcomes more influential than earlier outcomes. This recency parameter was set to a value of .1 for all analyses; there is strong behavioral evidence for this learning rate (Behrens, Woolrich, Walton, & Rushworth, 2007) and that is broadly used in analyses of this kind (Eldar & Niv, 2015). On a day when a team did not play, p_{win} was simply carried forward from the previous day, which made our analysis of prediction error analogous to the trial-based learning algorithms used in the experimental literature (Rutledge et al., 2014). The prediction error (*PE*) for a team on a given day was calculated as the difference between that day's expected outcome, $p_{\text{win}}(t)$—the moving average from the previous day—and actual outcome, $O(t)$:

$$PE(t) = O(t) - p_{\text{win}}(t).$$

On each day, the prediction errors resulting from teams that played on that day were summed to compute a citywide sports prediction error. This exact computation was applied for the six other U.S. metropolitan areas, resulting in citywide sports prediction errors for each area. These prediction errors were then related to gambling behavior on the subsequent day.

Solar irradiance data

We used satellite-derived estimates of direct normal irradiance (DNI), a measure of solar irradiance, in units of watts per square meter, on a surface directly facing the sun (Clean Power Research; http://www.solaranywhere.com). Larger DNI values indicate clearer skies (i.e., more sunshine). These hourly satellite-derived irradiance estimates have been demonstrated to be a reliable source of solar microclimate characterization (Perez et al., 2002). For each day in 2011 and 2012, we computed the mean nonzero DNI (i.e., because estimates for hours between sunset and sunrise yield zero), which served as our daily estimate of solar irradiance. From this quantity, we constructed a daily exponentially weighted average, computed analogously to the sports indices described earlier:

$$\overline{DNI}(t+1) = \overline{DNI}(t) + \alpha\left[DNI(t) - \overline{DNI}(t)\right].$$

Again, α was set to a value of 0.1, and the prediction error for a given day was calculated as the difference between $DNI(t)$ and $\overline{DNI}(t)$. In addition, we collected DNI estimates for four of the six U.S. metropolitan areas listed earlier. Boston and Philadelphia were omitted because their proximity to New York City (342 and 143 km, respectively) would induce spurious correlations into our analyses.

Demographic data

We used custom software to retrieve American Community Survey 5-year estimates from the U.S. Census Bureau (2015). We obtained the number of adult residents, per capita income levels, and highest completed level of education for the population 25 years of age and over. We also obtained the percentage of the adult workforce engaged in "management, business, science and arts" occupations; this major-level classification, defined by the Standard Occupational Classification Policy Committee at the U.S. Bureau of Labor Statistics (2010), is a rough measure of white-collar employment. The information retrieved included data for all 174 New York City ZIP codes in the years 2011 and 2012. Each neighborhood's composite socioeconomic status (SES) was computed as the sum of the z scores of its per capita income, years of education, and proportion of white-collar workers, following the method of Roberts (1997); we refer to this as ZIP-code SES. To calculate the proportion of residential properties in each ZIP code, we obtained lot-level information for 2011 from the New York City Department of City Planning (2013).

Nuisance variables

Because of the cyclicality inherent in data series of this sort, we specified a number of dummy variables to control for day-of-week effects, holidays, common paycheck cycles, and severe weather events. We constructed individual dummy-coded regressors for all days of the week and months of the year. Following the method of prior work (Evans & Moore, 2011), we dummy-coded regressors for U.S. national holidays that fell on Mondays or Fridays (Presidents' Day, Martin Luther King Jr. Day, Memorial Day, Labor Day, Columbus Day, Veterans' Day), as well as for New Year's Day, January 2, Easter Sunday, Independence Day, Thanksgiving Day, and Christmas Day. Separate dummy-coded regressors were constructed for common days on which people are paid, the 1st and the 15th days of each month (if either of these fell on a weekend, the immediately preceding weekday was used). We constructed a separate regressor coding for Hurricane Irene (August 27–29, 2011) and Hurricane Sandy (October 29–November 1, 2012), as well as a regressor coding for blizzards, defined as days on which snow occurred and average visibility was below 5 miles (there were 2 such days in 2012 and 5 in 2011; we obtained data on the seven cities and the two years in question from WeatherUnderground.com).

Data analysis

To mitigate the concern that lottery gambling in primarily nonresidential neighborhoods could be driven by nonresidents (e.g., people who purchase lottery tickets near their places of employment), we excluded data from 11 ZIP codes in which less than 10% of the properties were classified as residential (relative to other zoning classifications). Because Hurricane Sandy exerted lasting effects on population levels and commercial activity in a number of coastal neighborhoods (which was evident in per capita lottery purchases), we excluded data from 6 ZIP codes for which mean per capita purchases after Sandy (November 1–30, 2012) fell below 75% of their average levels before Sandy (January 1–October 27, 2012). These ZIP codes were removed only from the 2012 (confirmatory) data set; 163 and 157 ZIP codes remained in the 2011 and 2012 analyses, respectively.

Linear regressions were conducted as mixed-effects models, performed using the lme4 package (Pinheiro & Bates, 2000) in the R software environment (Version 3.2.2; R Development Core Team, 2015). Contrasts were performed using the esticon function in the doBy package for R (Højsgaard & Halekoh, 2009). The linear model included the dummy-coded nuisance regressors specified earlier. The predictors of interest were entered into the regression as z scores. Estimates and statistics are reported as fixed effects at the population level except as noted. In the Results section, for the graphs with lottery purchases as a dependent measure, we used the residual purchases calculated from the mixed-effects regression with only the nuisance variables described earlier as predictor variables. To examine the specificity of New York City and aggregated other-city effects, we performed a bootstrap test to compare the two distributions of effect sizes, comparing the observed Kolmogorov-Smirnov distance with those of 10,000 shuffled effect-size distributions (the null distribution). We note that our key hypotheses regarding prediction errors were developed in an analyses of the 2011 data, which served as a discovery data set. Our ex ante key hypothesis was then tested in the 2012 data set using a single-shot analysis. The supplemental outcome-based analyses (reported in Tables S2 and S4 in Supplemental Figure and Tables in the Supplemental Material available online) were subsequently carried out post hoc on the 2012 data set (in response to reviewer suggestions) without additional corrections for multiple comparisons. These subsequent analyses did not lead to a revision of the main conclusion derived from the single-shot analysis.

Results

Overall characteristics of city lottery gambling

We obtained daily lottery sales data, as described in the Method section. We averaged these data across 174 New York City ZIP codes. For each ZIP code, we aggregated the dollar sales of each of the daily lottery games and normed those numbers by the ZIP code's population, producing a composite per-capita index of lottery gambling, depicted in Figure 1a. We removed the influence of the day of week, the month of year, and a number of other nuisance variables using a mixed-effects regression, which yielded residual time courses of lottery gambling for each postal code (Fig. 1b). The temporal patterns in the amount of lottery purchases were correlated between neighborhoods (mean r = .23 across all ZIP codes over 2011), which suggests that common causes—unexplained by cyclicality or seasonality—might influence citywide gambling behavior.

Sports-based prediction errors and gambling

We obtained the outcomes of games played by the seven major professional sports teams based in New York City. From each team's day-to-day outcomes, we calculated p_{win} from an exponentially weighted average of past outcomes (Fig. 1c). For each day on which a given team played, we calculated the prediction error as the difference between the game's outcome and the expected p_{win}, yielding a time course of prediction errors that spanned the team's playing season (Fig. 1d). Finally, we calculated a citywide prediction-error measure covering an entire calendar year (Fig. 1e), as described in the Method section. We adopted this model because of its strong neurobiological (O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003) and psychological (Eldar & Niv, 2015; Rutledge et al., 2014) support and because exponentially weighted averaging makes minimal assumptions about the source of these expectations. The prediction error was positive when, on average, the city's teams performed better than expected, and it was negative when the teams performed worse than expected.

We hypothesized that when New York City–based teams performed unexpectedly well—that is, when prediction error was positive—citywide lottery gambling would increase on the next day. When teams performed worse than expected, gambling would decrease. We reasoned that the effects of sporting outcomes would be most salient on the following day. In addition, to correctly link outcomes of games played on a given day to lottery gambling occurring later on that same day (Edmans et al., 2007), hour-by-hour purchasing data would be needed, and such data were not available. Strikingly, the predictive relationship between prediction error and residual per capita lottery purchases was positive and close to linear (Fig. 2a): New York City residents gambled more on the lottery the greater their prediction error on the previous day.

A mixed-effects linear regression revealed a significant predictive effect of citywide sports prediction error, $\beta = 0.0029$, $p < .0001$ (see Table S1 in Supplemental Figure and Tables). Considered separately, positive and negative prediction errors exerted significant positive and negative effects, respectively; the negative prediction error was larger in magnitude, linear contrast, $p < .0001$ (see Table S8 in Supplemental Figure and Tables).

According to the logic of our hypothesis, only sports outcomes consequential to New York City residents should predict gambling behavior. To test this hypothesis, we computed citywide sports prediction errors for six additional U.S. metropolitan areas, as described in the Method section, and estimated each city's predictive effect on New York City lottery gambling. To facilitate comparison of prediction-error effect sizes across the metropolitan areas considered, we individually estimated prediction-error effect sizes for each city's teams at the New York City ZIP-code level and aggregated the effect sizes for the non–New York City teams. This yielded two distributions of neighborhood-level effect sizes (Fig. 2b).[1] The distributions of New York City's and the other cities' prediction-error effect sizes revealed that prediction-error effects on gambling were specific to New York City's sports teams (bootstrap test on effect sizes, $p < .0001$).

We also investigated whether absolute outcomes rather than prediction errors exerted similar effects on gambling. In particular, we examined whether the proportion of teams that won (Fig. 3a) and the aggregate number of wins (Fig. 3b) on the previous day were related to lottery purchases. In both cases, we found weaker and markedly less consistent relationships than those observed in the prediction-error based analysis (Fig. 2a), which suggests that the influence of incidental sporting outcomes on gambling is mediated by prediction error. Statistically, the day-to-day composite of sports outcomes (i.e., citywide wins minus citywide losses) was not significantly predictive of gambling (see Table S2 in

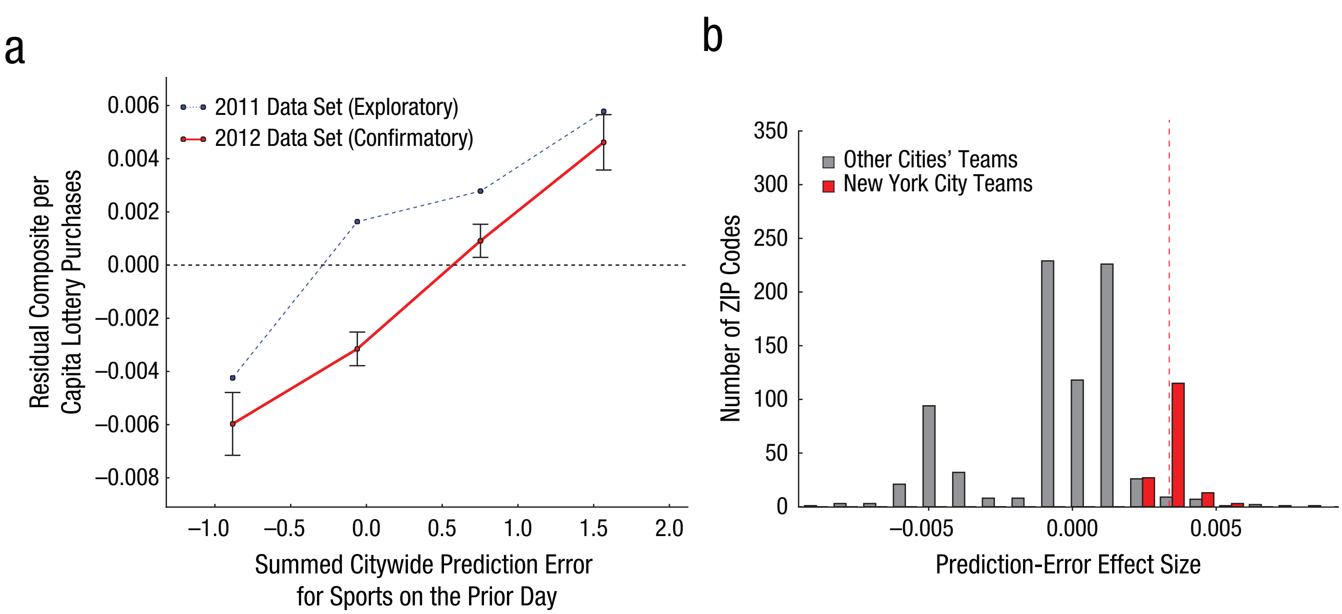

Fig. 2. Effect of prediction errors for sports on per capita lottery purchases. The graph in (a) shows the relationship between residual composite per capita lottery purchases and the summed citywide prediction error for sports on the prior day, separately for 2011 and 2012 data sets. Residual values were obtained by controlling for a number of cyclical and noncyclical nuisance variables. Error bars represent ±1 *SE*. The dashed black line indicates a zero-magnitude effect, interpreted as the baseline. The graph in (b) shows the distribution of prediction-error effect sizes for New York City teams and for teams from the other six cities combined (from the main mixed-effects regression). The dashed red line indicates the population-level New York City prediction-error effect size for sports.

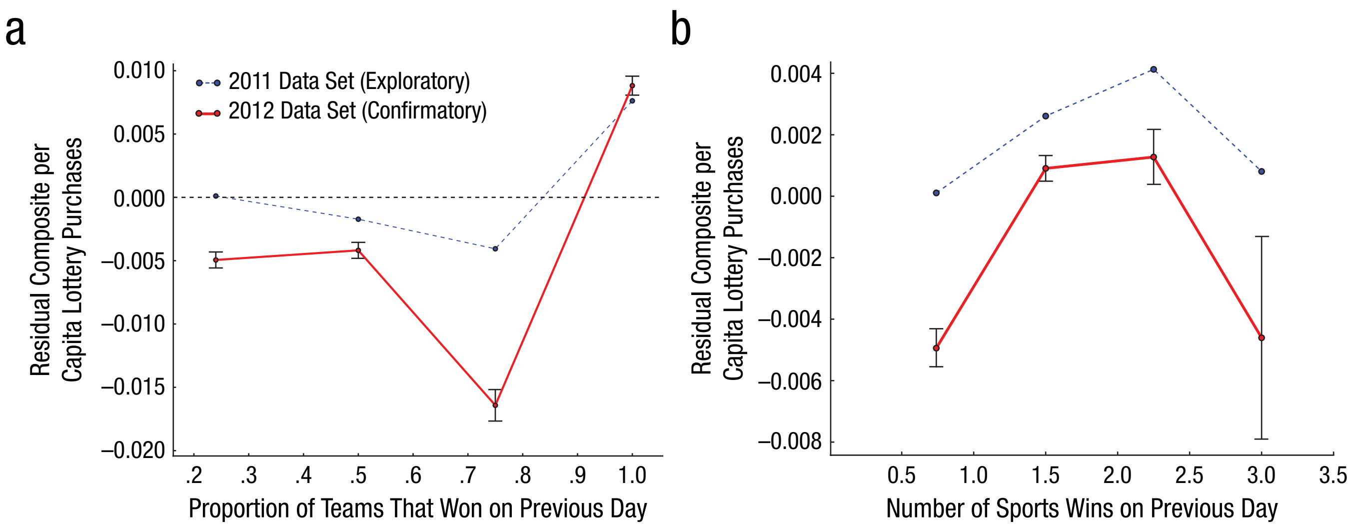

Fig. 3. Residual composite per capita lottery purchases as a function of (a) the proportion of teams that won on the previous day and (b) the number of wins on the previous day, separately for the 2011 and 2012 data sets. Residual values were obtained by controlling for a number of cyclical and noncyclical nuisance variables. Error bars represent ±1 *SE*. The dashed black line in (a) indicates a zero-magnitude effect, interpreted as the baseline.

Supplemental Figure and Tables), β = 0.0015, p = .108. Moreover, this outcomes-based model explained significantly less variance than the prediction-error-based model, likelihood ratio test, p < .01.

Sunshine-based prediction errors and gambling

The observed effect of sports prediction errors suggested that incidental good fortune affected gambling behavior. To ensure that this effect was not specific to sports and would generalize to other incidental outcomes, we performed a similar analysis of fluctuations in local sunshine intensity. On the basis of day-to-day average measurements of solar irradiance in New York City, we calculated an irradiance prediction error (in units of watts per square meter), which quantified how each day's sunshine level deviated from the exponentially weighted expectation calculated from recent history (see Fig. S1 in Supplemental Figure and Tables). These deviations positively predicted same-day lottery gambling (Fig. 4a), which indicated that positive changes in sunshine—for example, a sunny day after a prolonged period of cloudiness—predicted increased gambling behavior (see Table S3 in Supplemental Figure and Tables), β = 0.0035, p < .0001. Individually, positive and negative prediction errors exerted significant positive and negative effects, respectively, but in the case of irradiance (unlike for sports wins), the effect of positive prediction errors was larger in magnitude, linear contrast, p < .0001 (see Table S9 in Supplemental Figure and Tables).

Irradiance prediction errors calculated for a subset of the six other cities (omitting Boston and Philadelphia, which are close to New York City and have similar climates) again revealed that the effect of incidental good fortune on lottery gambling is specific to events in New York City (Fig. 4b), bootstrap test, p < .0001. Following our sports-based analyses, we also examined whether absolute levels of sunshine, rather than prediction errors, predicted gambling; again, these effects were less pronounced (Fig. 5). Sunshine itself was not significantly predictive of gambling, β = −0.0009, p = .389 (see Table S4 in Supplemental Figure and Tables), and the outcomes-based model again yielded a significantly poorer model fit than the model based on prediction error, likelihood ratio test, p < .05.

Effects of SES

Because our data set spanned 174 socioeconomically diverse ZIP codes, we were well positioned to examine whether the predictive effects of unexpected incidental outcomes varied with SES. When we divided ZIP codes into low- and high-SES subgroups via a median split, we found that sports prediction errors exerted nearly identical effects on lottery gambling (Fig. 6a). Indeed, the Citywide Sports Prediction Error × ZIP-Code SES interaction was not significant, β = −0.0001, p = .769 (see Table S5 in Supplemental Figure and Tables). Likewise, solar irradiance prediction errors were not moderated by SES (Fig. 6b). The Irradiance Prediction Error × ZIP-Code SES interaction was not significant, β = −0.0001, p = .407 (see Table S6 in Supplemental Figure and Tables). Thus, wealth level appears not to moderate the effect of incidental positive outcomes on risk attitudes as measured by fluctuations in lottery gambling.

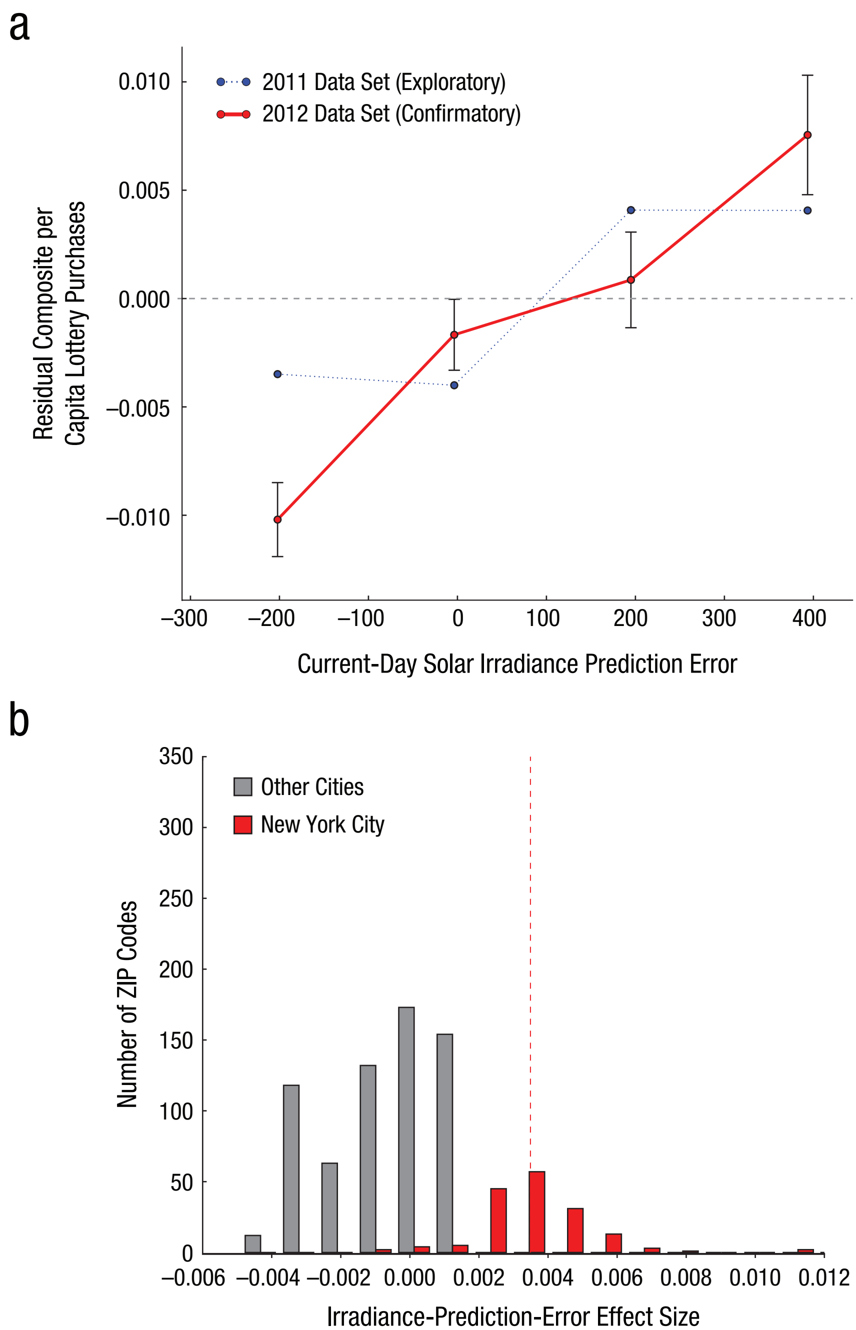

Fig. 4. Effect of prediction errors for irradiance on per capita lottery purchases. The graph in (a) shows the relationship between residual composite per capita lottery purchases and same-day irradiance prediction error, separately for the 2011 and 2012 data sets. The dashed black line in (a) indicates a zero-magnitude effect, interpreted as the baseline. Residual values were obtained by controlling for a number of cyclical and noncyclical nuisance variables. The graph in (b) shows the distribution of prediction-error effect sizes for New York City and for four cities that were geographically distant from New York combined. Error bars represent ±1 *SE*. The dashed red line in (b) indicates the population-level New York City effect size.

However, SES significantly (and negatively) predicted overall lottery gambling (Fig. 7a), $\beta = -0.1284$, $p < .001$ (see Table S7 in Supplemental Figure and Tables): People in low-SES neighborhoods bought more tickets per

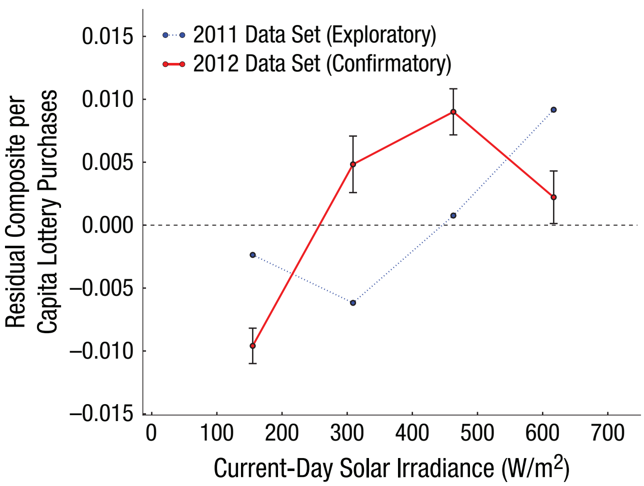

Fig. 5. Residual composite per capita lottery purchases as a function of current-day direct normal irradiance, separately for the 2011 and 2012 data sets. The dashed black line indicates a zero-magnitude effect, interpreted as the baseline. Error bars represent ±1 *SE*. Residual values were obtained by controlling for a number of cyclical and noncyclical nuisance variables.

capita in absolute terms than did people in high-SES neighborhoods. This result corroborates the well-documented finding that low-income individuals purchase more lottery tickets than high-income individuals (Barnes, Welte, Tidwell, & Hoffman, 2011; Oster, 2004; Rubenstein, Scafidi, & Rubinstein, 2002). Perhaps even more striking, residents of low-SES neighborhoods spent much larger fractions of their income on lottery tickets (Fig. 7b).

In some low-SES neighborhoods, people spent as much as 2% of their annual income on lottery tickets. That we recovered this relationship in our New York City–based lottery data supports our intuition that the composite behavior of people in individual ZIP codes by and large reflects the behavior of its residents rather than the behavior of people commuting to that ZIP code (11 low-residency exceptions were excluded from our analysis, as noted in the Method section).

Discussion

We analyzed a large urban data set using a simple model inspired by computational accounts of reward learning, and our results revealed that people gamble more when incidental outcomes in the environment are better than expected.

Indeed, we found compelling relationships between positive prediction errors and lottery gambling across two disparate domains—professional sports wins and sunshine levels—suggesting that the malleability of gambling attitudes may generalize, more broadly, to other positive or negative outcomes that occur outside of residents' control. Taken together, our results reveal a remarkable malleability to human risk taking. Furthermore, our results underscore the (so far largely unrealized) utility of massive real-world data sets to elucidate psychological phenomena, both to externally validate existing hypotheses developed in the laboratory and to circumvent the limitations of traditional laboratory study (Coviello et al., 2014; Eichstaedt et al., 2015).

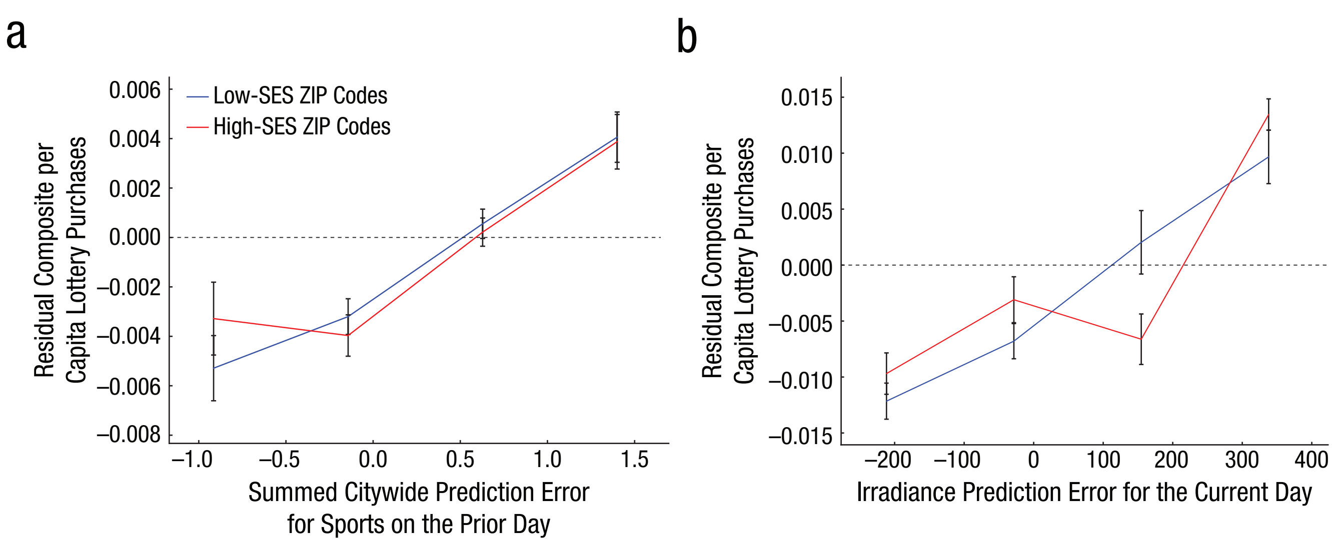

Fig. 6. Residual composite per capita lottery purchases as a function of (a) summed citywide prediction error for sports on the prior day and (b) solar irradiance prediction error for the current day, separately for high- and low-socioeconomic-status (SES) ZIP codes. The dashed black lines indicate zero-magnitude effects, interpreted as the baseline. Error bars represent ±1 *SE*. Residual values were obtained by controlling for a number of cyclical and noncyclical nuisance variables.

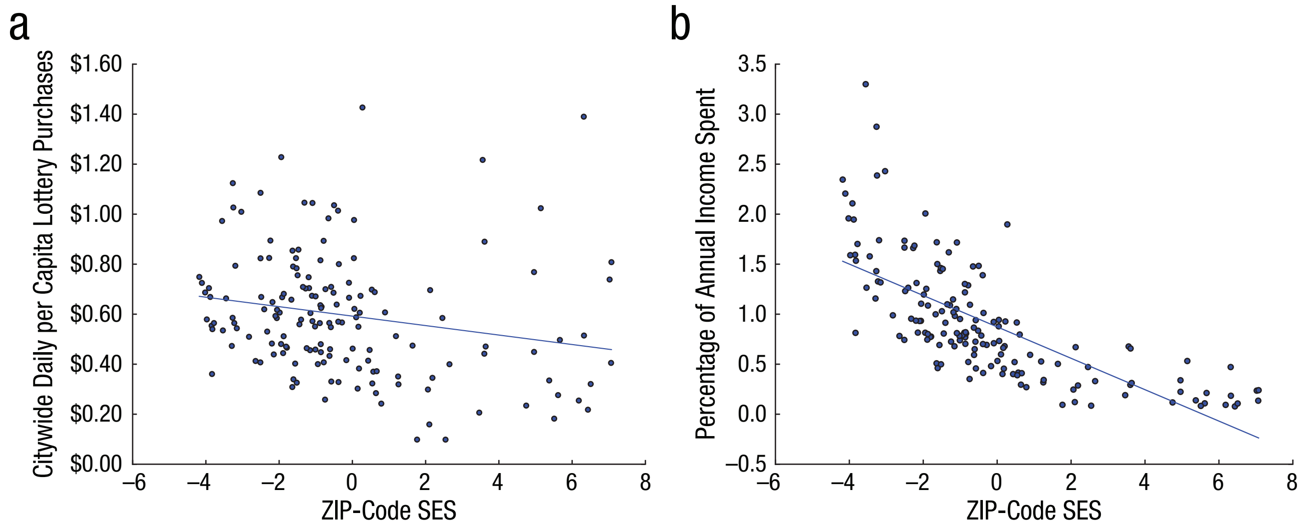

Fig. 7. Scatterplots (with best-fitting regression lines) of the relations between ZIP-code socioeconomic status (SES) and (a) mean per capita lottery purchases and (b) percentage of annual income spent on the lottery per year.

Our data set is powerful enough to demonstrate that the effect of prediction errors is both very specific (i.e., it was observed only after local outcomes that would be germane to New York City residents) and very general (i.e., neighborhoods differing in wealth level exhibited similar responses to incidental prediction errors). Although these effects appear small at the level of individual residents, the power of large-scale data sets can elucidate subtle psychological effects that loom larger at the level of a city. Indeed, on a day in which multiple sports teams unexpectedly win, approximately $160,000 more is spent on lottery gambling relative to an average day. In both sports and weather, positive and negative prediction errors were predictive of increased and decreased gambling, respectively. It is noteworthy that in sports, negative prediction errors exerted larger effects, although this is perhaps not surprising given that previous work examining professional sports outcomes and financial markets found that only losses could account for changes in market returns (Edmans et al., 2007). Conversely, the influence of sunshine itself has been demonstrated to exert a positive effect on risk tolerance (Bassi et al., 2013), which dovetails with our finding that irradiance prediction errors exerted larger positive effects than negative effects. We note that previous work has examined the effects of longer-term fluctuations in seasonality on risk attitudes (Kamstra, Kramer, & Levi, 2014). A relevant question for future work is whether a prediction error on weather should be viewed as reflecting expectations based on a weighted average of the recent past or on foreseeable seasonal variation.

One possible psychological-level explanation of our effects is that positive deviations from expectation may foster positive moods (Rutledge et al., 2014); these moods, in turn, are known to drive risk-taking behavior (Arkes et al., 1988; Bassi et al., 2013; Schulreich et al., 2014). However, it is important to consider that changes in affective state (e.g., mood) may not be necessary to explain how positive prediction errors foster risky decision making, and alternative mechanisms may merit consideration. For instance, neurocomputational accounts of learning emphasize a critical role for dopamine in signaling positive prediction errors (Bayer & Glimcher, 2005; Schultz et al., 1997); at the same time, dopamine levels modulate risk taking in animals and humans (Rutledge, Skandali, Dayan, & Dolan, 2015; St. Onge & Floresco, 2008)—and neither of these explanations necessarily involves mood. Thus, incidental prediction errors could drive fluctuations in risk taking through dopaminergic action without changing affective state. Alternatively, individuals may overgeneralize from their experience with incidental outcomes, which may lead them to believe that other low-likelihood outcomes are possible. Finally, a disproportionate sensitivity to positive prediction errors or a shift in attention to positive outcomes (or both) may lead to systematic overconfidence (Johnson & Fowler, 2011; Sharot, Korn, & Dolan, 2011). Clearly, future work is needed to evaluate these alternative hypotheses and to assess more definitively the role of mood state in the relationship between prediction errors and gambling.

The effect of prediction errors was similar in both low- and high-SES neighborhoods. Because low SES is associated with reduced willingness to take risks (Haushofer & Fehr, 2014) and with negative mood (Galea et al., 2007), it is perhaps surprising that unexpected positive outcomes would engender increased risk-taking behavior. At the same time, we found that low-SES neighborhoods exhibited higher lottery gambling levels overall. According to one account in the economics literature, credit constraints cause risk-averse individuals to play the lottery to increase their chances of crossing a threshold for purchasing a costly, indivisible good (Crossley, Low, & Smith, 2013). In this view, unexpected positive outcomes might engender increased risk taking of this sort.

Taken with caution, these findings may also have significance for policy. As demonstrated in previous work (Clotfelter & Cook, 1987; Oster, 2004) and in the present study (Figs. 7a and 7b), the daily lottery games examined are disproportionately consumed by lower-SES individuals. At the same time, lottery-linked savings accounts—financial products that provide randomized returns to depositors through periodic lottery drawings—have successfully induced saving behavior among low-income populations (Pfiffelmann, 2013). The ability to forecast when people will engage in this sort of gambling on the basis of incidental outcomes could inform how and when to market lottery-linked savings accounts, which could foster increased financial security.

Over a century ago, psychologists began to turn to the laboratory to isolate and analyze psychological phenomena. This powerful approach offered many advantages, but it also had the effect of shifting psychological inquiry away from the study of natural human behavior in the real world. Although correlational studies are limited in their ability to reveal causation, the growth of large-scale data sets now offers psychologists the opportunity to return to the field with newfound statistical power and precision. Our contribution to this endeavor is to show that unexpected outcomes predict meaningful behavior in the environment.

Author Contributions

A. R. Otto, S. M. Fleming, and P. W. Glimcher developed the research questions and analyses and wrote the manuscript. A. R. Otto performed the data analyses.

Acknowledgments

The authors acknowledge Agnieszka Tymula and Robb Rutledge for helpful conversations and Clean Power Research for providing public access to SolarAnywhere® data for use in this study.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://pss.sagepub.com/content/by/supplemental-data

Open Practices

All data have been made publicly available via Open Science Framework and can be accessed at https://osf.io/4bwq3/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received the badge for Open Data. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/1.%20View%20the%20Badges/ and http://pss.sagepub.com/content/25/1/3.full.

Note

1. Teams from other cities sometimes compete against teams based in New York City; as a result, certain negative outcomes for these other teams could positively have driven gambling in New York City.

References

Arkes, H. R., Herren, L. T., & Isen, A. M. (1988). The role of potential loss in the influence of affect on risk-taking behavior. *Organizational Behavior and Human Decision Processes*, *42*, 181–193. doi:10.1016/0749-5978(88)90011-8

Barnes, G. M., Welte, J. W., Tidwell, M.-C. O., & Hoffman, J. H. (2011). Gambling on the lottery: Sociodemographic correlates across the lifespan. *Journal of Gambling Studies*, *27*, 575–586. doi:10.1007/s10899-010-9228-7

Bassi, A., Colacito, R., & Fulghieri, P. (2013). ’O Sole Mio: An experimental analysis of weather and risk attitudes in financial decisions. *The Review of Financial Studies*, *26*, 1824–1852. doi:10.1093/rfs/hht004

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*, 129–141. doi:10.1016/j.neuron.2005.05.020

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214–1221. doi:10.1038/nn1954

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*, 365–376. doi:10.1038/nrn3475

Card, D., & Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, *126*, 103–143. doi:10.1093/qje/qjr001

Clotfelter, C. T., & Cook, P. J. (1987). *Implicit taxation in lottery finance* (Working Paper No. 2246). Retrieved from the National Bureau of Economic Research Web site: http://www.nber.org/papers/w2246

Coviello, L., Sohn, Y., Kramer, A. D. I., Marlow, C., Franceschetti, M., Christakis, N. A., & Fowler, J. H. (2014). Detecting emotional contagion in massive social networks. *PLoS ONE*, *9*(3), Article e90315. doi:10.1371/journal.pone.0090315

Crossley, T. F., Low, H., & Smith, S. (2013). *Do consumers gamble to convexify?* (Koç University-TÜSİAD Economic

Research Forum Working Paper Series No. 1314). Retrieved from http://www.econstor.eu/handle/10419/108633

Cunningham, M. R. (1979). Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan. *Journal of Personality and Social Psychology*, *37*, 1947–1956. doi:10.1037/0022-3514.37.11.1947

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 3787–3800. doi:10.1098/rstb.2008.0161

Edmans, A., García, D., & Norli, Ø. (2007). Sports sentiment and stock returns. *The Journal of Finance*, *62*, 1967–1998. doi:10.1111/j.1540-6261.2007.01262.x

Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., . . . Seligman, M. E. P. (2015). Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science*, *26*, 159–169. doi:10.1177/0956797614557867

Eldar, E., & Niv, Y. (2015). Interaction between emotional state and learning underlies mood instability. *Nature Communications*, *6*, Article 6149. doi:10.1038/ncomms7149

Evans, W. N., & Moore, T. J. (2011). The short-term mortality consequences of income receipt. *Journal of Public Economics*, *95*, 1410–1424. doi:10.1016/j.jpubeco.2011.05.010

Galea, S., Ahern, J., Nandi, A., Tracy, M., Beard, J., & Vlahov, D. (2007). Urban neighborhood poverty and the incidence of depression in a population-based cohort study. *Annals of Epidemiology*, *17*, 171–179. doi:10.1016/j.annepidem.2006.07.008

Garrett, I., Kamstra, M. J., & Kramer, L. A. (2005). Winter blues and time variation in the price of risk. *Journal of Empirical Finance*, *12*, 291–316. doi:10.1016/j.jempfin.2004.01.002

Gibson, B., & Sanbonmatsu, D. M. (2004). Optimism, pessimism, and gambling: The downside of optimism. *Personality and Social Psychology Bulletin*, *30*, 149–160. doi:10.1177/0146167203259929

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology*, *44*, 1110–1126.

Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *Science*, *344*, 862–867. doi:10.1126/science.1232491

Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, *58*, 1009–1032. doi:10.1111/1540-6261.00556

Højsgaard, S., & Halekoh, U. (2009). doBy: Groupwise statistics, LSmeans, linear contrasts, utilities (Version 4.5-13) [Software]. Retrieved from http://CRAN.R-project.org/package=doBy

Isen, A. M., & Patrick, R. (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance*, *31*, 194–202. doi:10.1016/0030-5073(83)90120-4

Johnson, D. D. P., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, *477*, 317–320. doi:10.1038/nature10384

Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2003). Winter blues: A SAD stock market cycle. *The American Economic Review*, *93*, 324–343.

Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2014). *Seasonal variation in Treasury returns*. Retrieved from Social Science Research Network Web site: http://papers.ssrn.com/abstract=1076644

Kamstra, M. J., Kramer, L. A., Levi, M. D., & Wermers, R. (2015). *Seasonal asset allocation: Evidence from mutual fund flows*. Retrieved from Social Science Research Network Web site: http://ssrn.com/abstract=1907904

Kuhnen, C. M., & Knutson, B. (2011). The influence of affect on beliefs, preferences, and financial decisions. *Journal of Financial and Quantitative Analysis*, *46*, 605–626. doi:10.1017/S0022109011000123

Love, B. C., Kopeć, Ł., & Guest, O. (2015). Optimism bias in fans and sports reporters. *PLoS ONE*, *10*(9), Article e0137685. doi:10.1371/journal.pone.0137685

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, *8*, 423–429. doi:10.2307/40063228

New York City Department of City Planning. (2013). *Primary land use tax lot output* [Data set]. Retrieved from http://www.nyc.gov/html/dcp/html/bytes/dwn_pluto_mappluto.shtml

Nielsen Company. (2013). *Local television market universe estimates: Comparisons of 2012-13 and 2013-14 market ranks*. Retrieved from http://www.nielsen.com/content/dam/corporate/us/en/docs/solutions/measurement/television/2013-2014-DMA-Ranks.pdf

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*, 329–337. doi:10.1016/S0896-6273(03)00169-7

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

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., & Vignola, F. (2002). A new operational model for satellite-derived irradiances: Description and validation. *Solar Energy*, *73*, 307–317. doi:10.1016/S0038-092X(02)00122-6

Pfiffelmann, M. (2013). What is the optimal design for lottery-linked savings programmes? *Applied Economics*, *45*, 4861–4871. doi:10.1080/00036846.2013.804168

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York, NY: Springer.

R Development Core Team. (2015). R: A language and environment for statistical computing (Version 3.2.2) [Computer software]. Retrieved from https://www.r-project.org/index.html

Roberts, E. M. (1997). Neighborhood social environments and the distribution of low birthweight in Chicago. *American Journal of Public Health*, *87*, 597–603.

Rogers, P. (1998). The cognitive psychology of lottery gambling: A theoretical review. *Journal of Gambling Studies*, *14*, 111–134. doi:10.1023/A:1023042708217

Rubenstein, R., Scafidi, B., & Rubinstein, R. (2002). Who pays and who benefits? Examining the distributional consequences of the Georgia lottery for education. *National Tax Journal*, *55*, 223–238.

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National*

Academy of Sciences, USA, *111*, 12252–12257. doi:10.1073/pnas.1407535111

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2015). Dopaminergic modulation of decision making and subjective well-being. *The Journal of Neuroscience*, *35*, 9811–9822. doi:10.1523/JNEUROSCI.0702-15.2015

Schulreich, S., Heussen, Y. G., Gerhardt, H., Mohr, P. N., Binkofski, F. C., Koelsch, S., & Heekeren, H. R. (2014). Music-evoked incidental happiness modulates probability weighting during risky lottery choices. *Frontiers in Psychology*, *4*, Article 981. doi:10.3389/fpsyg.2013.00981

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599. doi:10.1126/science.275.5306.1593

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*, R941–R945. doi:10.1016/j.cub.2011.10.030

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*, 1475–1479. doi.org/10.1038/nn.2949

Shepperd, J. A., & McNulty, J. K. (2002). The affective consequences of expected and unexpected outcomes. *Psychological Science*, *13*, 85–88. doi:10.1111/1467-9280.00416

Sloan, L. (1989). The motives of sports fans. In J. H. Goldstein (Ed.), *Sports, games, and play: Social and psychological viewpoints* (2nd ed., pp. 175–240). Hillsdale, NJ: Erlbaum.

St. Onge, J. R., & Floresco, S. B. (2008). Dopaminergic modulation of risk-based decision making. *Neuropsychopharmacology*, *34*, 681–697. doi:10.1038/npp.2008.121

U.S. Bureau of Labor Statistics. (2010). *2010 standard occupational classification*. Retrieved from http://www.bls.gov/soc/soc_structure_2010.pdf

U.S. Census Bureau. (2015). *American Community Survey* [Data set]. Retrieved from https://www.census.gov/programs-surveys/acs/data.html

Wright, W. F., & Bower, G. H. (1992). Mood effects on subjective probability assessment. *Organizational Behavior and Human Decision Processes*, *52*, 276–291. doi:10.1016/0749-5978(92)90039-A