

strong prior motivation as a return predictor is market-wide investor sentiment. At least as early as Keynes (1936), numerous authors have considered the possibility that a significant presence of sentiment-driven investors can cause prices to depart from fundamental values, thereby creating a component of future returns that corrects such mispricing. Baker and Wurgler (2006) and Stambaugh, Yu, and Yuan (2012), among others, find that investor sentiment and/or consumer confidence exhibits an ability to predict returns on various classes of stocks and investment strategies.⁴ These studies also refine the prior motivation of investor sentiment as a predictor. For example, Baker and Wurgler (2006) argue that sentiment should play a stronger role among stocks that are more difficult to value. In support of that hypothesis, they find sentiment exhibits greater ability to predict returns on small stocks, young stocks, high volatility stocks, unprofitable stocks, non-dividend-paying stocks, extreme growth stocks, and distressed stocks. Stambaugh, Yu, and Yuan (2012) hypothesize that when market-wide sentiment is combined with Miller (1977, 's) argument about the effects of short-sale impediments, overpricing due to high sentiment is more likely than underpricing due to low sentiment. Their results support that argument, in that sentiment predicts profits on the short legs of a large set of anomaly-based long-short strategies, whereas sentiment exhibits no ability to predict long-leg profits.

Despite the prior motivation for the properties that investor sentiment exhibits empirically as a predictor of anomaly returns, one might nevertheless be concerned that sentiment is simply a spurious predictor. Such a concern might be prompted, for example, by the results of Novy-Marx (2014), who reports that returns on various subsets of anomalies can apparently be predicted by seemingly unlikely variables such as sunspots and planetary positions.⁵ This study assesses the odds that investor sentiment's observed ability as a predictor can be achieved by a spurious regressor. We focus on the role of consistency across multiple return series and hypotheses. To understand the value of consistency, suppose the true expected returns across a number of portfolios possess some independent variation, but each expected return's true correlation with investor sentiment has the same sign. The greater the number of portfolios, the more difficult it becomes to find a spurious regressor that will exhibit finite-sample predictive ability consistently across portfolios comparable to that of investor sentiment. Our setting for exploring the role of consistency is that of Stambaugh, Yu, and Yuan (2012). That study examines 11

different anomalies and finds consistent results across those anomalies in support of three hypotheses: (i) a positive relation between current sentiment and future long-short return spreads, (ii) a negative relation between current sentiment and future short-leg returns, and (iii) no relation between current sentiment and future long-leg returns. We simply ask how likely it is that such hypotheses are supported as strongly by a randomly generated spurious regressor used in place of investor sentiment.

Out of 200 million simulated regressors, we find one that jointly support the three hypotheses in Stambaugh, Yu, and Yuan (2012) as strongly as investor sentiment. The odds are still quite long if one looks at just one of the three hypotheses. For example, comparably strong and consistent support for the first hypothesis—a positive relation between sentiment and the long-short return spread—occurs once in every 28,500 simulated regressors. For the second hypothesis—a negative relation between sentiment and short-leg returns—comparable support occurs once in every 105,000 regressors. If one sets aside any consideration of strength (*t*-statistics) and simply looks at the signs of regression coefficients dictated by the first two hypotheses, even then only one in every 43 simulated regressors achieves the consistency exhibited with investor sentiment.

2. Empirical setting and simulation results

The empirical setting we analyze here focuses on the main set of regression results reported by Stambaugh, Yu, and Yuan (2012), hereafter SYY. That study estimates the regression,

$$R_{i,t} = a + bS_{t-1} + cMKT_t + dSMB_t + eHML_t + u_t, \quad (1)$$

where $R_{i,t}$ is the excess return in month t on an anomaly strategy's long leg, short leg, or the difference, S_{t-1} is the level of the investor-sentiment index of Baker and Wurgler (2006) at the end of month $t - 1$, and MKT_t , SMB_t , and HML_t are the returns on month t on the three stock-market factors defined by Fama and French (1993). SYY examine 11 anomalies documented previously in the literature:

1. Failure probability (Campbell, Hilscher, and Szilagyi, 2008)
2. Distress (Ohlson, 1980)
3. Net stock issues (Ritter, 1991; Loughran and Ritter, 1995)
4. Composite equity issues (Daniel and Titman (2006))
5. Total accruals (Sloan, 1996)
6. Net operating assets (Hirshleifer, Hou, Teoh, and Zhang, 2004)
7. Momentum (Jegadeesh and Titman, 1993)
8. Gross profitability (Novy-Marx, 2013)
9. Asset growth (Cooper, Gulen, and Schill, 2008)
10. Return on assets (Fama and French, 2006; Chen, Novy-Marx, and Zhang, 2010; Wang and Yu, 2010)
11. Investment-to-assets (Titman, Wei, and Xie, 2004; Xing, 2008)

As in SYY, the sample period is from August 1965 through January 2008 for all but anomaly (1), whose data begin in

(footnote continued)

not the regressor predicts future return. Such a correlation is especially likely for a regressor that is a valuation ratio, such as dividend yield. The finite-sample bias that arises in such a setting is analyzed by Stambaugh (1999).

⁴ Other studies that document the ability of sentiment measures to predict returns include Brown and Cliff (2004, 2005), Lemmon and Portniaguina (2006), Baker and Wurgler (2007, 2012), Livnat and Petrovits (2009), Baker, Wurgler, and Yuan (2012), Antoniou, Doukas, and Subrahmanyam (2013), Stambaugh, Yu, and Yuan (2013), and Yu (2013).

⁵ Indeed, a preliminary version of that study presented such results in the context of spurious regressors.

December 1974, and anomalies (2) and (10), whose data begin in January 1972. For each anomaly, SYY examine the long-short strategy using deciles 1 and 10 of a sort based on the anomaly variable, with the long leg being the decile with the highest average return. SYY also examine a combination strategy that takes equal positions across the long-short strategies constructed in any given month.

The coefficient of interest in Eq. (1) is b . SYY (cf. Table 5) report results of estimating b for each of the 11 anomalies, as well as the combination strategy, in three sets of regressions that relate to the three hypotheses explored in that study. For the first hypothesis, $R_{i,t}$ is the long-short return difference, and the estimate \hat{b} has the predicted positive sign for all 11 anomalies. The t -statistic for \hat{b} , based on the heteroskedasticity-consistent standard error of [White \(1980\)](#), ranges from 0.22 to 3.38 across the individual anomalies and equals 2.98 for the combination strategy. For the second hypothesis, $R_{i,t}$ is the short-leg return, and \hat{b} has the predicted negative sign for all 11

exercise to just the single regressions using the combination strategy, we still find that only one spurious regressor out of every 6,580 simultaneously supports each of the three hypotheses as strongly as investor sentiment.

2.2. Joint-comparison benchmarks

As the above analysis illustrates, the consistency of results across multiple anomalies and hypotheses makes it especially unlikely that such results are produced by a spurious regressor. While simultaneous joint comparisons reveal the importance of consistency, they can also make interpreting the strength of the results less straightforward. Each number in [Table 1](#) essentially gives the reciprocal of the probability under the “null hypothesis”—a spurious predictor—of obtaining a sample outcome at least as extreme as the one actually observed using the sentiment series S_t . However, when the comparison involves a vector of statistics, as opposed to a single statistic, the corresponding probability can be fairly low even if the sample outcome is considerably less extreme than the sample outcome that was actually observed. If considerably less extreme outcomes also have low probabilities under the null, then it becomes difficult to interpret the

low probability associated with outcomes more extreme than the actual outcome.⁶

Interpreting the values in [Table 1](#) becomes easier in the presence of benchmark values that reflect what one expects the values in [Table 1](#) to be when the actual sentiment series S_t is replaced by a truly spurious predictor. [Table 2](#) contains such benchmark values, computed by replacing the t -statistics based on the sentiment series S_t with t -statistics based on a spurious regressor y_t . That is, rather than tabulating how often a spurious regressor x_t supports the SYY hypothesis as well as the actual series S_t , we tabulate how often a spurious regressor x_t does as well as another spurious regressor y_t . A new series y_t is drawn for each draw of the series x_t .

Table 2

Benchmark number of randomly generated predictors required to obtain one predictor that produces results as strong as another random predictor.

The table reports the reciprocal of the frequency with which a randomly generated predictor x_t produces results as strong as another randomly generated predictor y_t when x_t and y_t replace S_t in the regression,

$$R_{i,t} = a + bS_{t-1} + cMKT_t + dSMB_t + eHML_t + u_t,$$

where $R_{i,t}$ is the excess return in month t on an anomaly's long leg, short leg, or the difference, S_t is the level of the investor-sentiment index of Baker and Wurgler (2006), and MKT_t , SMB_t , and HML_t are the three stock-market factors defined in Fama and French (1993). The predictors x_t and y_t are generated as a first-order autoregression with autocorrelation equal to 0.988, the bias-corrected estimate of the autocorrelation of S_t .

Let \bar{t}_i^y denote the i -th highest t -statistic for \hat{b} (the estimate of b) among the 11 anomalies when y_t is used, and let \bar{t}_i^x denote the i -th highest t -statistic when x_t is used. Let \underline{t}_i^y denote the i -th lowest t -statistic for \hat{b} when y_t is used, and let \underline{t}_i^x denote the i -th lowest t -statistic when x_t is used. Let $|t_i^y|$ denote the i -th smallest t -statistic in absolute value when y_t is used, and let $|t_i^x|$ denote the i -th smallest t -statistic in absolute value when x_t is used. The row for anomalies reflects the frequency with which the following conditions are satisfied:

$$\bar{t}_i^x \geq \bar{t}_i^y \text{ occurred at least } j \text{ times among } i = 1, \dots, 11, \text{ in the long-short column.}$$

$$\underline{t}_i^x \leq \underline{t}_i^y \text{ occurred at least } j \text{ times among } i = 1, \dots, 11, \text{ in the short-leg column.}$$

$$|t_i^x| \leq |t_i^y| \text{ occurred at least } j \text{ times among } i = 1, \dots, 11, \text{ in the long-leg column.}$$

The "combination" row reflects the frequencies with which a simulated predictor produces t -statistics satisfying the above inequalities when $R_{i,t}$ is an equally weighted combination of the 11 anomaly strategies. The final row reflects the frequencies with which the above inequalities are satisfied for 11 anomalies as well as the combination strategy. The last two columns reflect the frequencies with which the inequalities are satisfied jointly across the previous columns.

Comparisons	(1) Long-short	(2) Short leg	(3) Long leg	(1) and (2)	(1), (2), and (3)
1 Anomaly	1.4	1.4	1.1	–	–
2 Anomalies	1.6	1.5	1.2	–	–
3 Anomalies	1.7	1.7	1.3	–	–
4 Anomalies	1.8	1.8	1.5	–	–
5 Anomalies	1.9	1.9	1.7	–	–
6 Anomalies	2.0	2.0	2.0	–	–
7 Anomalies	2.1	2.1	2.4	–	–
8 Anomalies	2.3	2.3	3.0	–	–
9 Anomalies	2.5	2.5	3.9	–	–
10 Anomalies	2.8	2.8	5.8	–	–
11 Anomalies	3.5	3.5	11.5	–	–
Combination	2.0	2.0	2.0	2.2	4.4
11 Plus the combination	3.5	3.5	16.3	4.4	70.8

71 supports the three SYH hypotheses as strongly as y_t . In other words, the Table 2 value of 71 is a benchmark for interpreting the Table 1 value of 200 million: it is what one expects the Table 1 value to be if S_t is truly spurious. Dividing the Table 1 value by the Table 2 value gives what might be characterized as the "effective" value of the former. For example, dividing 200 million by 71 gives an effective value of about 2.8 million—still very large. Similar comparisons to Table 2 can be made for other values in Table 1. For example, recall from Table 1 that only one spurious regressor out of 468,000 supports the first two SYH hypotheses as strongly as S_t . The corresponding benchmark value in Table 2 is 4.4, and dividing 468,000 by 4.4 still gives over 106,000. In general we see that, while the joint-comparison issue is important, interpreting the Table 1 values in light of the Table 2 benchmarks still yields the overall conclusion that the SYH results are extremely unlikely if S_t is a spurious regressor.

2.3. Additional comparisons

To judge whether a spurious regressor supports the SYH hypotheses as strongly as the actual investor-sentiment series, one must define "supports as strongly." While the definition employed above in Tables 1 and 2 seems a

reasonable way to capture the consistency of results across anomalies, there are of course alternative definitions. For example, we could instead examine the k least favorable t -statistics for a given hypothesis, comparing those produced by x_t to those obtained using S_t . To illustrate, let $k=1$ and consider the first hypothesis, which predicts $b > 0$ when $R_{i,t}$ is the long-short return difference. The lowest t -statistic produced by S_t among the 11 anomalies is equal to 0.22, and less than one x_t series out of every 50 produces a minimum t -statistic greater than that value. For the second hypothesis, which predicts $b < 0$ when $R_{i,t}$ is the short-leg return, the weakest t -statistic using S_t is -1.11 , and only one x_t in every 2,300 produces a weakest t -statistic less than -1.11 . Now let $k=2$, and note that the second-lowest t -statistic produced by S_t for the first hypothesis equals 0.76. Only one x_t series out of every 163 produces a lowest t -statistic greater than 0.22 as well as a second-lowest t -statistic greater than 0.76. With hypothesis 2, for only one x_t out of 10,000 are the two weakest t -statistics more favorable to the hypothesis than the two weakest t -statistics using S_t . Proceeding through additional k values and the remaining third hypothesis would produce a table in the same format as Table 1, with entries in the final three rows identical to those in Table 1 and larger entries in the first ten rows, corresponding to

Table 3

Number of randomly generated predictors required to obtain one predictor that enters with the correct sign.

The table reports the reciprocal of the frequency with which a randomly generated predictor x_t produces an estimate of b with the predicted sign when x_t replaces S_t in the regression,

$$R_{i,t} = a + bS_{t-1} + cMKT_t + dSMB_t + eHML_t + u_t,$$

where $R_{i,t}$ is the excess return in month t on an anomaly's long leg, short leg, or the difference, S_t is the level of the investor-sentiment index of Baker and Wurgler (2006), and MKT_t , SMB_t , and HML_t are the three stock-market factors defined in Fama and French (1993). The predictor x_t is generated as a first-order autoregression with autocorrelation equal to 0.988, the bias-corrected estimate of the autocorrelation of S_t .

The row for j anomalies reflects the frequency with which a simulated predictor produces an estimate of b for at least j anomalies with the predicted sign (positive in the long-short column and negative in the short-leg column). The "combination" row reflects the frequency with which a simulated predictor produces an estimate of b with the predicted sign when $R_{i,t}$ is an equally weighted combination of the 11 anomaly strategies. The last column reflects the frequencies with which the predicted signs are obtained jointly across the previous columns.

Comparisons	(1)	(2)	(1) and (2)
	Long-short	Short leg	
1 Anomaly	1.0	1.1	–
2 Anomalies	1.1	1.1	–
3 Anomalies	1.3	1.3	–
4 Anomalies	1.4	1.4	–
5 Anomalies	1.7	1.7	–
6 Anomalies	2.0	2.0	–
7 Anomalies	2.5	2.5	–
8 Anomalies	3.3	3.3	–
9 Anomalies	4.9	4.9	–
10 Anomalies	8.8	8.5	–
11 Anomalies	25	21	–
Combination	2.0	2.0	2.2
11 Plus the combination	25	21	43

longer odds.⁷ Thus, comparing the weakest results across the individual anomalies would deliver a similar message as Table 1, if anything, even more strongly.

Of course, conducting joint comparisons of weakest results raises the same benchmarking issue discussed in the previous subsection. That is, an alternative version of Table 1 based on comparing weakest results could be accompanied by the corresponding weakest-result version of Table 2. For example, when $k=1$, the alternative Table 1 values of 50 and 2,300 reported above for the first and second hypotheses have corresponding "effective" values of 25 and 1,150 when divided by the values that would appear in the alternative version of Table 2. Similarly, when $k=2$, the alternative Table 1 values of 163 and 10,000 reported above have corresponding effective values of 70 and 4,367. As before, the low frequencies still seem low when interpreted in the context of joint comparisons.

Another approach that to some degree captures consistency across anomalies is simply comparing median t -statistics. For example, across the 11 individual anomalies

as well as the combination strategy, the median t -statistic for the first hypothesis equals 2.41 using S_t , and one x_t out of every 1,650 produces a median t -statistic as large. For the second hypothesis, the median t -statistic using S_t equals -2.57 , and one x_t out of every 1,186 produces a median t -statistic greater in negative magnitude. Only one x_t out of every 7,103 produces median t -statistics that are simultaneously as favorable to both hypotheses. For the third hypothesis, the median absolute t -statistic using S_t is 0.46. One x_t out of every 15 produces a median absolute t -statistic that low, but only one x_t out of 562,000 does so while simultaneously producing median t -statistics as favorable to the first two hypotheses as those obtained using S_t . The effective frequency of such an outcome is still less than one out of 123,000 if one adjusts for the joint-comparison issue in the same manner as discussed earlier.

The average t -statistic across anomalies says little about consistency across anomalies. Nevertheless, it appears rather unlikely that a spurious regressor can produce even comparably favorable average t -statistics. For example, the averages of the SYR-reported t -statistics across the 11 anomalies and the combination strategy are 2.14 and -2.38 for the first and second hypotheses, respectively. The average absolute value of the SYR-reported t -statistics is 0.69 for the third hypothesis. An average t -statistic supporting the first hypothesis as strongly (i.e., greater than 2.14) is produced by one x_t out of every 554. An average t -statistic supporting the second hypothesis as strongly (i.e., less than -2.38) occurs for one x_t out of every 1,393. Average t -statistics simultaneously supporting both hypotheses as strongly occur once every 2,412. An x_t producing that simultaneous support for the first two hypotheses while also being as favorable to the third hypothesis—delivering an average absolute t -statistic less than 0.69—occurs only once in every 237,000. Adjusting for the joint-comparison issue still leaves that effective frequency at less than one in every 53,000.

Finally, fairly unlikely is just the possibility that a spurious regressor would give \hat{b} 's with the predicted signs consistently across all anomalies. Table 3 reports the frequencies with which a spurious regressor gives the predicted sign across anomalies for the long-short difference (first hypothesis) and the short-leg return (second hypothesis). For the first hypothesis, one in every 25 spurious regressors gives the predicted positive sign for all 11 anomalies. For the second hypothesis, the frequency of getting the predicted negative sign for all 11 anomalies is one in every 21. A spurious predictor that produces all 22 coefficients with the predicted signs, as does investor sentiment, occurs only once in every 43 randomly generated regressors.

3. Conclusions

It appears to be extremely unlikely that the observed role of investor sentiment in stock-return anomalies can be filled by a spurious regressor. Out of 200 million simulated regressors, we find none. These very long odds—seemingly no better than those attached to winning the Powerball Jackpot with a single play—reflect the consistency with which investor sentiment produces results across multiple

⁷ To see this, note that the k -th row of Table 1 reports the frequency with which any k of the ordered t -statistics using x_t is as favorable to the given hypothesis as are the corresponding ordered t -statistics using S_t . The k -th row of the alternative table would consider instead the least favorable k t -statistics, constituting only a subset of the outcomes included in the frequency in Table 1.

anomalies for the three SYH hypotheses.⁸ Simultaneous support of the SYH hypotheses is important, by itself, in that the odds of a spurious regressor supporting them as strongly as investor sentiment are only one in 6,580 even when all of the anomalies are combined into a single long-short strategy. It is the consistency across the individual anomalies, however, that raises the highest hurdle for a spurious regressor to clear in order to play the role of investor sentiment.

References

- Antonioni, C., Doukas, J., Subrahmanyam, A., 2013. Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis* 48, 245–275.
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61, 1645–1680.
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129–152.
- Baker, M., Wurgler, J., 2012. Comovement and predictability relationships between bonds and the cross-section of stocks. *Review of Asset Pricing Studies* 2, 57–87.
- Baker, M., Wurgler, J., Yuan, Y., 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics* 104, 272–287.
- Brown, G., Cliff, M., 2004. Investor sentiment and the near-term stock market. *Journal of Empirical Finance* 11, 1–27.
- Brown, G., Cliff, M., 2005. Investor sentiment and asset valuation. *Journal of Business* 78, 405–440.
- Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. *Journal of Finance* 63, 2899–2939.
- Chen, L., Novy-Marx, R., Zhang, L., 2010. An alternative three-factor model. Unpublished working paper. CKGSB, University of Rochester, and Ohio State University.
- Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. *Journal of Finance* 63, 1609–1652.
- Daniel, K.D., Titman, S., 2006. Market reactions to tangible and intangible information. *Journal of Finance* 61, 1605–1643.
- Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.
- Fama, E., French, K., 2006. Profitability, investment, and average returns. *Journal of Financial Economics* 82, 491–518.
- Ferson, W., Sarkissian, S., Simin, T.T., 2003. Spurious regressions in financial economics? *Journal of Finance* 58, 1393–1413.
- Granger, C.W.J., Newbold, P., 1974. Spurious regressions in economics. *Journal of Econometrics* 4, 111–120.
- Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38, 297–331.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for market efficiency. *Journal of Finance* 48, 65–91.
- Kendall, M.G., 1954. Note on bias in the estimation of autocorrelation. *Biometrika* 41, 403–404.
- Keynes, J.M., 1936. *The General Theory of Employment, Interest, and Money*. Macmillan, London.
- Lemmon, M., Portniaguina, E., 2006. Consumer confidence and asset prices: some empirical evidence. *Review of Financial Studies* 19, 1499–1529.
- Livnat, J., Petrovits, C., 2009. Investor sentiment, post-earnings announcement drift, and accruals. Unpublished working paper. New York University and College of William and Mary.
- Loughran, T., Ritter, J.R., 1995. The new issues puzzle. *Journal of Finance* 50, 23–51.
- Marriott, F.H.C., Pope, J.A., 1954. Bias in the estimation of autocorrelations. *Biometrika* 41, 390–402.
- Miller, E.M., 1977. Risk, uncertainty and divergence of opinion. *Journal of Finance* 32, 1151–1168.
- Novy-Marx, R., 2013. The other side of value: the gross profitability premium. *Journal of Financial Economics* 108, 1–28.
- Novy-Marx, R., 2014. Predicting anomaly performance with politics, the weather, global warming, sunspots, and the stars. *Journal of Financial Economics* 112, 137–146.
- Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18, 109–131.
- Ritter, J.R., 1991. The long-run performance of initial public offerings. *Journal of Finance* 46, 3–27.
- Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289–315.
- Stambaugh, R.F., 1999. Predictive regressions. *Journal of Financial Economics* 54, 375–421.
- Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: investor sentiment and anomalies. *Journal of Financial Economics* 104, 288–302.
- Stambaugh, R.F., Yu, J., Yuan, Y., 2013. Arbitrage asymmetry and the idiosyncratic volatility puzzle. Unpublished working paper. University of Pennsylvania, University of Minnesota, and Shanghai Jiao Tong University.
- Titman, S., Wei, K., Xie, F., 2004. Capital investments and stock returns. *Journal of Financial and Quantitative Analysis* 39, 677–700.
- Wang, H., Yu, J., 2010. Dissecting the profitability premium. Unpublished working paper. University of Delaware and University of Minnesota.
- White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48, 817–838.
- Xing, Y., 2008. Interpreting the value effect through the Q-theory: an empirical investigation. *Review of Financial Studies* 21, 1767–1795.
- Yu, J., 2013. A sentiment-based explanation of the forward premium puzzle. *Journal of Monetary Economics* 60, 474–491.
- Yule, G.U., 1926. Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series. *Journal of the Royal Statistical Society* 89, 1–64.

⁸ Powerball is a multi-state lottery in which the odds of a single combination of numbers claiming a share of the top “Jackpot” prize are roughly one in 175 million.