

ODA Model Perfectly Predicts S&P 500 Declines of 20% or Greater

Paul R. Yarnold, Ph.D.
 Optimal Data Analysis, LLC

The Shiller price-to-earnings (P/E) ratio assesses mean inflation-adjusted earnings in the prior decade.¹ Five times over the past 150 years this index consistently exceeded 30 during a bull market run. Two times were during the Great Depression and the “dot-com” bubble, and both resulted in record declines in equity prices. Two other times occurred in the past three years and ended in declines of 20% to 34% in the S&P 500 index. The fifth time began on February 3 of 2021, with the Shiller P/E ratio for the S&P 500 approaching 35.

Historically there were four periods when the Shiller P/E ratio exceeded 30 and sustained this level (always for less than a year) during a bull market rally.¹ Each period terminated with a decline of $\geq 20\%$ in the price of equities.²

For this result the exact Type I error^{3,4} is $p < 7.2 \times 10^{-43}$. The corresponding optimal discriminant analysis (ODA⁵⁻¹²) model is:

If BULL MARKET and SHILLER > 30,
 predict S&P 500 will decline $\geq 20\%$; otherwise predict S&P 500 will decline $< 20\%$.

The “confusion” (i.e., cross-classification) table summarizing classification performance of the ODA model developed using the “training data” is presented in Table 1.

Table 1: Training Classification Performance of ODA Model

| <u>Actual Class</u> | <u>Predicted Class</u> | | <u>Sensitivity</u> |
|-------------------------|------------------------|---------------------|--------------------|
| | Decline < 20% | Decline $\geq 20\%$ | |
| Decline < 20% | 139 | 0 | 100% |
| Decline $\geq 20\%$ | 0 | 4 | 100% |
| <u>Predictive Value</u> | 100% | 100% | |

The fifth such period began in August of 2020. Since that date, the Shiller P/E ratio for the S&P 500 has exceeded 30, and as of this writing it exceeds 35.

Current elaboration of the ODA paradigm, called *novometric* theory, states the efficacy of a statistical model developed in training analysis is assessed in validity analysis.¹³⁻¹⁶ According to the ODA model elaborated herein, reduction in the value of the S&P 500 index of at least 20% is expected. Historical data suggest a correction will occur within six months from the date of publication of this note.

References

¹<https://www.multip.com/shiller-pe>

²<https://www.fool.com/investing/2021/02/06/you-should-expect-a-20-stock-market-crash-in-2021/>

³Soltysik RC, Yarnold PR (1994). Univariable optimal discriminant analysis: One-tailed hypotheses. *Educational and Psychological Measurement*, 54, 646-653.

⁴Carmony L, Yarnold PR, Naeymi-Rad F (1998). One-tailed Type I error rates for balanced two-category UniODA with a random ordered attribute. *Annals of Operations Research*, 74, 223-238.

⁵Yarnold PR, Soltysik RC (1991). Theoretical distributions of optima for univariate discrimination of random data. *Decision Sciences*, 22, 739-752.

⁶Yarnold PR, Soltysik RC (2005). *Optimal data analysis: Guidebook with software for Windows*. Washington, D.C.: APA Books.

⁷<https://github.com/njrhodes/ODA>

⁸Rhodes NJ (2020). Statistical power analysis in ODA, CTA and Novometrics (Invited). *Optimal Data Analysis*, 9, 21-25.

⁹Yarnold PR, Brofft GC (2013). ODA range test vs. one-way analysis of variance: Comparing strength of alternative line connections. *Optimal Data Analysis*, 2, 198-201.

¹⁰Yarnold PR (2018). Visualizing application and summarizing accuracy of ODA models. *Optimal Data Analysis*, 7, 85-89.

¹¹Rhodes NJ, Yarnold PR (2020). Generating novometric confidence intervals in R: Bootstrap analyses to compare model and chance ESS. *Optimal Data Analysis*, 9, 172-177.

¹²Yarnold PR (2017). What is optimal data analysis? *Optimal Data Analysis*, 6, 26-42.

¹³Yarnold PR (2020). What is novometric data analysis? *Optimal Data Analysis*, 9, 195-206.

¹⁴Yarnold PR, Linden A (2016). Novometric analysis with ordered class variables: The optimal alternative to linear regression analysis. *Optimal Data Analysis*, 5, 65-73.

¹⁵Yarnold PR (2016). Novometric vs. ODA reliability analysis vs. polychoric correlation with relaxed distributional assumptions: Interrater reliability of independent ratings of plant health. *Optimal Data Analysis*, 5, 179-183.

¹⁶Yarnold PR, Linden A. (2016). Theoretical aspects of the D statistic. *Optimal Data Analysis*, 5, 171-174.

Author Notes

No conflicts of interest were reported.