ABSTRACT. Simulation evidence obtained within a Bayesian model of price-setting in a betting market, where anonymous gamblers queue to bet against a risk-neutral bookmaker, suggests that a gambler who wants to maximize future profits should trade on the advice of the analyst cum probability forecaster who records the best probability score, rather than the highest trading profits, during the preceding observation period. In general, probability scoring rules, specifically the log score and better known “Brier” (quadratic) score, are found to have higher probability of ranking rival analysts in predetermined “correct” order than either (i) the more usual method of counting categorical forecast errors (misclassifications), or (ii) an economic measure of forecasting success, described here as the “Kelly score” and defined as the trading profits accumulated by making log optimal bets (i.e. Kelly betting) against the market maker based on the probability forecasts of the analyst being assessed. This runs counter to the conventional wisdom that financial forecasts are more aptly evaluated in terms of their financial consequences than by an abstract non-monetary measure of statistical accuracy such as the number of misclassifications or a probability score.

KEY WORDS: bid-ask spread, economic forecast evaluation, Kelly criterion, probability forecasting, probability scoring rules, Kelly score

JEL CLASSIFICATION: C11, C44, D40, D81, C52, G11

1. INTRODUCTION

Economic “survival of the fittest” has been supposed by Alchian (1950) and Friedman (1953) to let rational investors prosper, and reduce others to bankruptcy. Or more accurately, wealth is predicted to gravitate over economic history (repeated trials) to agents who act as (or as if) expected utility maximizers trading on the basis of rational expectations.
Others inevitably decline into penury, and leave no impression on the long run evolution of market prices. This is called the market selection hypothesis, and is sufficient of itself to imply a theory of efficient markets.

Rather than accept the market selection hypothesis as axiomatic, and partly in response to studies in finance on the limits to arbitrage and the sometimes reality that noise traders prevail rather than disappear (De Long et al., 1990, 1991), Blume and Easley (1992) demonstrated a rich array of market equilibrium conditions under which certain investor types are bound to prosper, survive or ultimately fail. The most severe circumstances of economic Darwinism, formalized by Blume and Easley (1992) and extended by Sandroni (2000) and Blume and Easley (2001, 2002), occur in a market populated by heterogeneously informed and possibly irrational traders (all with the same discount rate or savings-versus-consumption rate) where expected utility maximizers with (i) log utility \( U(x) = \log(x) \), and (ii) objectively “correct” beliefs (i.e. “true” probabilities), are mathematically bound to prosper and eventually ruin all others.

Expected log-utility maximization, “log optimal” investment or “Kelly betting” as it is known to professional gamblers, implicitly maximizes the investor’s expected (long run average) rate of capital growth (Breiman, 1961; Kelly, 1956; MacLean et al., 1992). When based on physically or objectively “true” probabilities, no other decision rule produces the same wealth over the long run.

More generally, decision makers cannot know “true” probabilities, but among all Kelly bettors those with the “most accurate” probabilities can expect to accrue the greatest wealth:

In the log utility case, with equal discount factors [savings rules], the market selects for those with the most nearly correct beliefs. (Blume and Easley, 1992, p. 23)

The question then is: who has the “most nearly correct beliefs”—and in what sense? Suppose, for example, that there are two competing analysts who each report personal