Review of G. Shafer and V. Vovk “Probability and Finance: It’s Only a Game”, Wiley, 2001, xii+414 pp.

V. N. Tutubalin

English translation by Glenn Shafer and Vladimir Vovk
December 30, 2001

This monograph reports a series of outstanding and entirely unexpected results by the authors, in part purely mathematical, and in part belonging to the branch of mathematical science that studies financial markets. It involves a new conception of probability, which the authors call “game-theoretic probability”, but I would prefer to call “market probability”. (Whenever a new scientific approach is created, the somewhat mystical essence of things is at stake, and this is why the name is important. My reasons for preferring “market probability” will become clear as we proceed.) At the base of this new approach lies not the assumption that some source of randomness like coin tossing exists, but only the assumption that one can bet money on the future course of events, as market speculators have done from time immemorial.

There are two styles in scientific writing. One style, modern, moves the scientific novelty of the authors’ discoveries to the foreground. This style is flawed, because novelty is elusive, and in general there is nothing worse than starting a dispute over priority (the classical example—Leibniz and Newton). The other style, ancient, emphasizes continuity. In the Middle Ages it was thought to be in good taste to say first that, well, Hermes Trismegistus (a.k.a. the Egyptian god Thoth) and Albertus Magnus, as it turns out, had about the same opinions as the author about the matters under discussion. This is a very good style for a monograph, and it is the style followed by the authors of this book. Admittedly, they begin not from Hermes Trismegistus himself, but only from Pascal and Fermat, going from them to Mises and

Kolmogorov and highlighting as ideologically close the work of Jean Ville (1910–1988), author of the probabilistic notion of a martingale. (It turns out that J. L. Doob wrote a review of Ville's book in 1939.) The brevity required of a review makes it impossible to follow the ancient style here. Invoking my prerogative as a reviewer, I declare that, in my opinion, the authors' approach is entirely original. It seems to me that their predecessors, for all their splendor, had nothing like it.

Demonstrating this claim requires giving an example of the authors' approach. I will adapt one from Chapter 3 of the book. Consider a financial market operating at discrete steps in time (say once a day), let S_n denote the price of some asset on day n, and suppose the price's evolution is described by the simple and natural formula $S_{n+1} = S_n(1 + x_n)$. When we take the usual probabilistic view of the dynamics of the asset's price, we say that the x_n—the relative increments of the price—are random variables. Moreover, we often assume that they are independent or at least stationary in time. But these assumptions are always a source of doubt and discomfort; after all, we cannot exhibit a random number generator. Maybe the x_n are ultimately controlled by the devil, who would hardly be obliged to respect statistical homogeneity by making their distribution stationary. Amazingly, the book's authors have discovered that the devil has no way out! He can be forced to respect many probabilistic laws, including the strong law of large numbers. He can be tamed in this way by the most modest and cautious market speculator—a speculator who has modest initial capital and is so cautious that he avoids any risk of bankruptcy. If the devil tries to violate the law of large numbers, he will have to pay this speculator an unboundedly large amount of money, something that even he, the devil, will not like. Moreover, the speculator's strategy is not complicated—say something only known to be computable. It is very simple, almost trivial. Algorithmic probability has nothing to do with the matter.

Indeed, suppose the speculator's initial capital is just one dollar, and suppose his strategy is to use, on each day n ($n = 1, 2, \ldots$), a fixed fraction ϵ ($0 < \epsilon < 1$) of his current capital K_n to buy the asset. This is a bet on the price increasing from day n to day $n+1$, because he always sells on day $n+1$ what he bought on day n, even if the price went down. The evolution of the speculator's capital is described by the equation

$$K_{n+1} = K_n(1 - \epsilon) + \epsilon K_n(1 + x_n) = K_n(1 + \epsilon x_n),$$

where x_n is the devil's move on day n. Taking into account the initial condition $K_0 = 1$ we obtain

$$K_{n+1} = \prod_{i=1}^{n} (1 + \epsilon x_i).$$

Now it can be made perfectly clear, just from the properties of logarithms, that the behavior of the speculator's capital is closely connected with the behavior of the sum $\sum_{i=1}^{n} x_i$. To make the mathematical demonstration simple, assume that the devil's moves satisfy the additional condition $|x_n| \leq 1$. (By common sense $x_n \geq -1$, and the restriction $x_n \leq 1$ is quite reasonable for daily price increments.) Suppose the devil violates the strong law of large numbers in such a way that

$$\limsup n^{-1} \sum_{i=1}^{n} x_i > \epsilon > 0. \tag{1}$$

This produces $\sup K_n = \infty$; the total reluctantly paid out by the devil, considered at selected times, will tend to infinity. Indeed, if one assumes to the contrary that $K_n < C$, then $\sum_{i=1}^{n} \ln(1+\epsilon x_i) < D$. But since for $t \geq -1/2$ the inequality $\ln(1 + t) \geq t - t^2$ holds, we obtain the following inequalities:

$$\epsilon \sum_{i=1}^{n} x_i - \epsilon^2 \sum_{i=1}^{n} x_i^2 \leq D, \quad \epsilon \sum_{i=1}^{n} x_i - \epsilon^2 n \leq D,$$

whence

$$n^{-1} \sum_{i=1}^{n} x_i \leq D/(\epsilon n) + \epsilon$$

for all n, in contradiction with the inequality (1). Since the speculator only risks the fraction ϵ of his current capital on each round, his capital never becomes negative. By mixing strategies with decreasing ϵ and decreasing initial capitals the speculator can prevent the devil from satisfying the inequality (1) for any positive ϵ. The same strategy with the speculator's bets opposite in sign (bearish gambling) prevents him from violating the law of large numbers in the negative direction as well.

Our speculator is so modest (small initial capital) and cautious (never goes bankrupt) that one does not even want to call him a speculator. The book's authors call him Skeptic. But I will again allow myself to argue about the name, because this is important for the general philosophy of the market that follows from the book under review: I suggest we call Skeptic Controller.[1] Controller would be employed and paid by the institution operating the market, just as an accountant is employed by a firm and an actuary by an insurance company. Like an accountant or an actuary, Controller would have the responsibility, in the eyes of society and the state, for making sure his employer not violate certain laws. The accountant must make sure taxes are paid. The actuary must calculate insurance premiums fairly, so the insurance company does not go bankrupt. The position of market controller does not exist at the moment, but it would be good to establish it by legislation. (Who would license a firm without an accountant or an insurance company without an actuary?! Similarly, a financial exchange should not be able to get a license without a controller.) In the authors' model, Controller ensures, using purely market means, that the market does not violate the law of large numbers.

Of course, it has long been known that the market is a very powerful machine. But the book under review reveals new capabilities of the market. It might seem that the possibility of betting on future events differs little from Mises's requirement that frequencies be stable under selection of subsequences. But the assumption that one cannot change a frequency when one is allowed to use an arbitrary algorithm to decide whether or not to include future events in the calculation of the frequency turns out to be much more complicated and much weaker than the assumption that one cannot earn a very large sum without risking bankruptcy when one is allowed to put arbitrary bets on the future events. The authors show that the latter assumption implies many probabilistic laws that we are accustomed to deduce using the notion of measure, relying ultimately on the assumption of the existence of a source of randomness. So how is this new probability theory developed further in the book?

I must admit that what I have written so far does not do justice to the book. The book considers not only speculators in financial markets, but also much more general games, in which there are three players: Skeptic, Forecaster and Reality. In standard probability theory, one cannot do without the notion of variance. In game-theoretic probability, this notion is replaced by the assumption that Skeptic can bet not only on Reality's future move x_n (devil's move in the case of the speculator in a financial market) but also on

[1] *Translators' Note*: The Russian word used by the reviewer is ординатор. The English form, *ordinator*, appears in the Oxford English Dictionary but is labeled obsolete.

x_n^2. (In the second part of the book, this assumption is reshaped in a way that is important and very useful for hedging options; see below.) The order of play is described by a protocol (i.e., scenario). Though small and simple, such protocols are extraordinarily rich in implications. In other words, the authors have discovered in such protocols quite a simple and apt way of expressing the conditions of mathematical theorems.

As an example, I will quote the protocol of Chapter 4 (p. 79 of the book):

Skeptic's initial capital is $K_0 = 1$.
For $n = 1, 2, \dots$ the players act in the following order: first
Forecaster announces $m_n \in R$ and $v_n \geq 0$, and then
Skeptic announces $M_n \in R$ and $V_n \geq 0$, and finally
Reality announces $x_n \in R$.
As a result, Skeptic's capital takes the value

$$K_n = K_{n-1} + M_n(x_n - m_n) + V_n((x_n - m_n)^2 - v_n).$$

Additional restrictions: Skeptic is obliged to respect the condition $K_n \geq 0$ for all n, and Reality is obliged not to allow unlimited growth of Skeptic's capital.

One can deduce the following mathematical theorem from this protocol. If $\sum_{n=1}^{\infty} v_n/n^2 < \infty$ is satisfied, then $\lim n^{-1} \sum_{i=1}^{n}(x_i - m_i) = 0$. This is the game-theoretic variant of Kolmogorov's law of large numbers. (In my opinion, it would be better to say "market variant of Kolmogorov's law".) Chapter 5 establishes the corresponding variant of the law of the iterated logarithm.

What about the central limit theorem: is the devil obliged to obey it? The authors develop their variant of the theory of weak limit theorems in Chapters 6 and 7. The general philosophy changes here: now it is not a matter of Skeptic (or Controller) forcing the market to obey laws of probability theory under threat of having to make an infinitely large payout. In fact, the number N of moves in the game is now assumed to be bounded, and although N formally tends to infinity, the authors always take care to estimate how large N must be for a limit theorem to hold with a given accuracy. No, now it is a matter of Skeptic hedging a future financial obligation in a way that accounts for any possible behavior by Reality. Another surprise: it turns out that calculation of the initial capital needed for such hedging reduces to integration over the classical normal law. Let us move to more precise statements.

Consider a *variable* x. This is a function defined on the set of elementary events, i.e., strings of possible moves by Reality of length N. (Since one does not assume any probability measure on the set of elementary events, the authors avoid the term "random variable".) The variable x denotes the future obligation. The *upper price* of the obligation x is the value α for Skeptic's initial capital K_0 for which there exists a strategy for Skeptic (this is the hedge) under which his capital K_N at the end of the game satisfies $K_N \geq x$ for all elementary events. (The *lower price* of the obligation x is equal, by definition, to minus the upper price of the obligation $(-x)$.) The upper price is the lowest price, if you will, at which Skeptic can buy x, and the lower price is the highest price at which x can be sold. By common sense, the upper price is no smaller than the lower price; otherwise Skeptic would make easy money by repeatedly selling x and buying it back. So one allows only *coherent* protocols of the game, protocols for which this condition is satisfied. Yet another surprise: it turns out that if $x = U(S_N)$, where U is a sufficiently smooth function and S_N is the sum of Reality's moves in the game, then the upper price is close to the lower price and both are approximately given by the integral of the function $U(z)$ with the weight $\phi(z)$, where ϕ is the standard normal density. The proof is based on Lindeberg's method, which uses the heat equation satisfied by the normal distribution. The rule for finding the hedging strategy is similar to the Black-Scholes rule of "delta-neutral" hedging (i.e., the number of shares in the hedging portfolio is equal to the derivative of the option's price with respect to the current share price). As an example of a precise statement of a theorem we will quote a sample protocol (page 161 of the book):

First one announces the parameters $A \geq 1$, $B \geq 1$, $\sigma^2 > 0$, K_0. Then for $n = 1, 2, \ldots, N$ the following happens.

Forecaster announces $v_n \in [0, B\sigma^2/N]$ respecting the condition $\sum_{n=1}^{N} v_n = \sigma^2$.

Skeptic announces $M_n \in R$ and $V_n \in R$.

Reality announces $x_n \in [-A\sqrt{v_n}, A\sqrt{v_n}]$.

Skeptic's capital takes the value $K_n = K_{n-1} + M_n x_n + V_n(x_n^2 - v_n)$.

The conclusion of the theorem whose hypothesis is given by the protocol is that for the value $x = U(S_N)$ the upper and lower prices are approximately equal to the integral of the function $U(z)$ with respect to the normal density with mean zero and standard deviation σ^2. The constants A and B influence the speed of convergence of the upper and lower price to that integral: the number N should be large as compared with the product $A^2B\sigma^2$. Notice that in this protocol it is important that the number σ^2 be announced in advance and that Reality's moves x_n be small—of order $O(1/\sqrt{N})$. The theorem by no means asserts that the sum S_N of the devil's moves will obey the normal distribution in any sense. It only says that we will not be much mistaken if we calculate the required initial capital and hedging strategy starting from the normal distribution; we will approximately meet our obligation.

Chapter 8 of the book is called "The Generality of Probability Games". This chapter shows, first of all, that a series of measure-theoretic probability theorems can be regarded as special cases of game-theoretic probability theorems. This is a beautiful result. Second, the most general formal scheme of game-theoretic probability first appears in this chapter, and this is also very good, because the reader would hardly be able to understand the most general abstractions if the book started with them. The authors show great concern for the situation of the reader, who would like to understand the essence of the matter first, from simpler examples. (Remember the beginning of this review, where I followed the authors in deriving the law of large numbers from the properties of logarithms only—nothing more complicated.) Third, this chapter gives two applications of the book's concept of probability to topics not connected with any game or market, namely tests of quantum theory and tests of Cox's model for dependence of the mortality of a member of a population on medical parameters. A third similar application, connected with testing the correctness of a probabilistic weather forecast, is given slightly earlier, in Chapter 7 (pages 162–164).

These applications seem to me less promising than applications to the study of the market. I will start a controversy by asserting that a single conception of probability suitable for all possible applications may scarcely be possible. Take, for example, the theory of errors. Here the devil's move x_n is not a change in the price of an asset but rather an error in the measurement of a physical quantity in some experiment. It would be very good to be able to study conditions under which the law of large numbers is satisfied, i.e., the error disappears as the result of averaging a great number of observations. But who is in a position to make monetary bets on the errors of experiments? The game-theoretic conception clearly has nothing to do with this situation.

In the matter of weather forecasting, the authors propose to test the soundness of a calculation of probabilities for rain on successive days using a statistic that can also be obtained from the frequentist conception of probability, normalized as if the differences $x_i - p_i$ were statistically independent, where x_i is the indicator of the event "presence of rain on day i", and p_i is the forecasted probability of rain. Since nothing guarantees the independence of such differences, I would also recommend thinning out the sequence of observations in different ways, so as to see how the statistic comes out just on Mondays, just on Tuesdays, etc., and then comparing these values with the common value of the statistic. Who will convince me that such an attempt to weaken the assumed statistical dependence is meaningless? Yet it is not, for some reason, recommended under the game-theoretic approach.

In connection with quantum mechanics, I will note that testing a physical theory is a complicated matter and goes rather deep in physics. Testing is never done directly by observing the results of an experiment: one must first put forward alternatives to the theory being tested. This problem cannot be solved by making game-theoretic bets.

As for the model of mortality, I think the exposition is too brief; I did not understand it. In any case, this kind of work needs to be accompanied by analysis of empirical data, which is absent in this case.

These are the reasons for my remarks about terminology. As it appears to me, the conception put forward in the book is most effective precisely in those cases where the nature of the phenomenon itself brings monetary bets into the picture—in cases where we are studying speculation in a market. Hence my suggestion to replace "game-theoretic probability" by "market probability". In the area of the market, the authors' approach seems to me extremely interesting, and this brings us to the second part of the book (Chapters 9–15), which is called "Finance without Probability".

Chapters 9 and 10 are devoted to hedging options in the case of discrete time. Let us first consider the general philosophy concerning the market that is implied by the mathematical results. Perhaps after trying other things, the authors arrived at the conclusion that good mathematical theorems about option hedging in the game-theoretic situation can be obtained if a new assumption is made. They assume that we can have the market price a new asset, which pays its owner dividends in the amount $(dS(t)/S(t))^2$. (Here $S(t)$ is the price of the asset and $dS(t)$ is the change in price over a short period of time, say one day.) No such asset exists at the moment in the market. It is interesting that the reviewer, reflecting independently on the philosophy of the market, also arrived at something very similar.

The train of thought runs as follows. What do we want from the market in general, and how can probabilistic financial mathematics be helpful? In its current form, the market clearly lacks stability. Speculators, while providing the economy with the necessary dynamism, can easily create all kinds of market crises in pursuit of their own profit. It is unfortunate, for example, that fluctuations in the market price of oil have to be smoothed out by changing the level of production: oil futures, being merely pieces of paper or even electronic code, can change hands very easily in the market, and the destabilization of production is a much more unpleasant matter. How can probabilistic financial mathematics help here? The point is that this science developed the notion of the volatility of market prices, which can be estimated adequately by the sum of squared increments of the logarithm of the price (which coincides for practical purposes with $(dS(t)/S(t))^2$). We may suggest that each speculator should pay some amount of money for the excessive nervousness shown by the market as a whole, and so we might impose a tax on the (squared) volatility. If we connect this with the train of thought of the authors of the book under review, then we find (quite in the spirit of a market economy) that as we introduce such a tax, we must issue in the market securities (let us call them *indulgences*) that exempt the bearers from the tax on volatility for a set period of time. This is precisely the asset needed for the mathematical theorems given by the authors of the book. So a hedging portfolio consists not of shares and riskless securities but rather of shares and indulgences. (Of course, all this is in the simplified situation where only one asset is initially traded in the market).

To explain the meaning of the hedging protocol used in the book, one has to add that hedging can end with a zero change in the hedger's capital only when the future volatility is known precisely.[2] The authors of the book give the market responsibility for providing values D_n that estimate $\sum_{k=n}^{N}(\Delta S_k/S_{k-1})^2$, but are not yet given by the market at time $n=0$, when hedging begins. Only D_0 is given at that point.

After these remarks one can give an example of a hedging protocol. This is the so-called Black-Scholes protocol (page 249).

At the beginning of the game, parameters N, $I_0 > 0$ (Investor's initial capital), $\delta > 0$, $C > 0$ are announced. (See below for the meaning of

[2] *Translators' Note*: Here the reviewer means only that conventional Black-Scholes hedging, which involves only trading in the underlying asset (and a risk-free bond, if one does not ignore interest rates), requires precise knowledge of the future volatility of the asset's price in order to succeed fully. The method he is about to explain replaces this precise knowledge with market prices for future volatility.

the last two parameters.) Market announces parameters S_0 and D_0.

For $n = 1, \ldots, N$ the following happens:

Investor announces $M_n \in R$ and $V_n \in R$.

Market announces $S_n > 0$ and $D_n \geq 0$.

Investor's capital evolves by the equation

$$I_n = I_{n-1} + M_n \Delta S_n + V_n((\Delta S_n / S_{n-1})^2 + \Delta D_n).$$

We impose only these additional restrictions: $0 < S_n < C$ for $n = 1, \ldots, N$, $0 < D_n < C$ for $n = 0, 1, \ldots, N-1$, $D_N = 0$, and

$$\inf_{\epsilon \in (0,1)} \max \left(var_S(2+\epsilon), var_D(2-\epsilon) \right) < \delta.$$

(Here $var_F(\alpha)$ for a function $F = F(n)$ means the sum over n of the increments $|\Delta F(n)|^\alpha$.)

The conclusion of the theorem (Proposition 10.3) is that for a payoff function $U(S_N)$ the price of hedging the obligation is approximately equal to the integral of the function $U(S_0 \exp z)$ with the weight $n(z; -D_0/2, D_0)$, where the latter function is the density of the normal law with mean $(-D_0/2)$ and variance D_0. (The theorem also estimates the approximation error.)

One might add that the investor's move V_n in this protocol is a bet on the difference between the actual value $(\Delta S_n / S_{n-1})^2$ and the market's estimate $(-\Delta D_n)$ for it.

Chapters 11–14 of the book are devoted to a theory of hedging in continuous time, based on non-standard analysis. Not having command of the apparatus of non-standard analysis, I may say, as Socrates about Parmenides: "what I understand in Parmenides is beautiful; from this I conclude that what I do not understand is also beautiful". I myself explain continuous-time hedging to students as the limiting case of discrete hedging, basing myself on Kolmogorov's theorem about convergence of sequences of Markov chains to a diffusion process (convergence in the sense of probability distributions). While doing this, I emphasize that neither paths of dynamical systems in mathematical physics nor paths of asset prices on a fine time scale have anything to do with paths of diffusion processes. Only the probabilities converge. But what can we do if there are no probability distributions at all? Apparently the authors of the book are absolutely correct in their appeal to non-standard analysis. I will say merely that Chapter 11 is the analogue of

Chapter 10 for continuous time, and Chapters 12 and 13 cover different generalizations: price processes with jumps, taking account of the interest rate on capital, and American options (where the holder of the option decides when to exercise it). In Chapter 14, diffusion processes are embedded in the general game-theoretic scheme.

Chapter 15, which does not use non-standard analysis, concludes the book by considering the game-theoretic variant of the efficient-market hypothesis. The authors aptly suggest that we take the total value of the market's assets (more precisely, 10^{-10} times this total) as our numéraire. It turns out, of course, that under quite reasonable conditions Skeptic can prevent all the other speculators from redistributing the total value of the market to their benefit, no matter how they choose their portfolios (this is a variant of the law of large numbers). One cannot outwit the market as a whole. I will also point out the striking bound, given in §15.4, on the return that Investor can receive on capital invested in shares of only one company. Namely, if r_n is the return on the nth day, and $\sigma_0^2 = N^{-1}\sum_{n=1}^{N} r_n^2$ (this is an estimate of the squared volatility of the asset), then the average return $\mu = N^{-1}\sum_{n=1}^{N} r_n$ must satisfy the inequality $\mu \leq \sigma_0^2 - \ln\alpha/N +$ some small number. This happens with lower probability $\geq 1-\alpha$. (The lower probability of an event is the lower price of its indicator.) The inequality means that a large average return is only possible for shares with high volatility. In the case of Microsoft, the inequality is satisfied for $\alpha = 0.01$, but not for $\alpha = 0.1$. The authors explain this last observation by saying that Microsoft is clearly in the top 10% of all companies with respect to the growth rate of share prices, but it is not quite clear how game-theoretic probability is related to such a comparison, which probably refers to frequentist probability.

Now a few words about the book as a whole. I would evaluate the book as a classic work of the highest rank, which explains entirely new ideas in a way that it is easy and pleasant to read page by page, from beginning to end. In this sense the book can be compared with Laplace's "Philosophical essay" (but not with his "Analytical theory of probability", which is quite obscure) or Mises's "Probability, statistics and truth". Of course, it is difficult to say if the book will actually become an equally popular work, since fame depends not only on the intrinsic virtues of a book, but also on how it is marketed, and this lies outside our influence. In and of itself, the book merits wide popularity, and one would like to see it translated into Russian.

A natural application domain for the book's mathematics and philosophy is the market, and, in my opinion, the authors commit, in a very slight degree, an error that is customary for pioneers (including Laplace and Mises): they overestimate somewhat the range of applicability of their discoveries. But in this case the magnitude of the sin is unusually small. So what directions can we point out (very speculatively, of course) for further study of the market?

The market is very important for for the modern economy and for life in general, yet it is very poorly understood scientifically. For the sake of science, it would be good to study the market with the methods of modern experimental psychology, e.g., to attach to every speculator a psychomotor gauge of psychic excitement, something like a lie detector. But this would involve offending all speculators, and one cannot do without them. Equip the mice and keyboards used in electronic trading with such detectors? This may be too much as well. One is left with the price data. Unfortunately, the market researcher has to be content with mere crumbs of information in comparison to what is technically possible in a computer age. One can obtain data about the final prices and volumes of deals, but who exactly made the deals and the dynamics of their electronic bidding all remain unknown. The exchanges (and other systems of electronic trading) have these data, but they themselves neither give them out nor analyze them. They have no need to do so.

If only a situation were created where exchanges would want to analyze this information. . . . Such a situation can be created by imposing a volatility tax on the exchange, which will then be redistributed among the traders. This would make the profession of controller necessary. Moreover, it would automatically create the asset ("indulgence") that the authors need for their theoretical constructions. Even so, no genuine understanding of the market dynamics of prices will emerge, because the market surely has periods of rise and fall that are not understood by any science, and in particular, by probabilistic financial mathematics of any variant: classic (with probability) or game-theoretic (without probability). After all, in any variant the starting point is the rejection of the possibility of profitable speculation, and such speculation would be quite possible for those who understood the nature of the market's rises and falls. But it is quite possible that the proposals under discussion will be sufficient to smooth these rises and falls, and this is a very worthwhile goal.