Multiagent Cooperative Search for Portfolio Selection

David C. Parkes[*]

Department of Computer and Information Science,
University of Pennsylvania,
Philadelphia, PA 19104
E-mail: dparkes@unagi.cis.upenn.edu

and

Bernardo A. Huberman

Internet Ecologies Group,
Xerox Palo Alto Research Center,
Palo Alto, CA 94304
E-mail: huberman@parc.xerox.com

We present a new multiagent model for the multiperiod portfolio selection problem. A system of cooperative agents divide initial wealth and follow individual worst-case optimal investment strategies from random portfolios, sharing their final profits and losses. The multiagent system achieves better average-case performance than a single agent with the same initial wealth in a simple stochastic market. A further increase in performance is achieved through communication of hints between agents and probabilistic strategy-switching. However, this explicit cooperation is redundant in a market that approximates the Capital Asset Pricing Model, a model of equilibrium stock price dynamics.

1. INTRODUCTION

An investment portfolio is an effective way to increase expected long-term return and decrease risk when investing in a stock market [42]. Instead of investing in a single stock, an investor can select a balanced portfolio across different stocks. The portfolio selection problem has received considerable attention in both the financial [12, 14] and statistics literature [49, 15, 3, 16]. In this paper we introduce a new *multiagent* model for multiperiod portfolio selection, the problem of choosing a sequence of portfolios over time to

maximize a measure of performance that is appropriate for the risk-return preferences of an investor.

Our model builds on a recent computationally efficient portfolio-selection rule for a single agent [25]. We assume a system of agents that share their initial wealth and make individual investment decisions, before sharing profits and losses at the end of the final period. In one variation, called *independent search*, each agent selects an initial portfolio and follows the single agent portfolio-selection rule without communication with other agents. In another variation, called *cooperative search*, agents can communicate about the recent performance of their portfolio-selection strategies. Agents choose to switch probabilistically to the strategy in the population that has been performing best in the recent past. Similar cooperative search models have enabled exponential performance improvements in problem solving domains within artificial intelligence [13, 26].

The independent multiagent search model outperforms a single agent in a simple simulated market. The simple market is characterized with price changes independently distributed across stocks. Furthermore, the prices are exogenous inputs to the system because we assume that the system of investors is small with respect to the total size of the market. Cooperative search with explicit communication boosts performance in the same market. However, cooperative search performs no better than independent search in a more realistic market, an approximation to the influential Capital Asset Pricing Market (CAPM) model [53], which introduces correlations between stock prices and constraints on volatility to model the equilibrium between agent investment decisions and stock price movements.

The difference in the performance of cooperative search across the markets can be explained by the nature of the statistics of the market, and the effect that the statistics have on the nature of the optimal portfolio-selection strategy. There is a meaningful long-term optimal portfolio in both markets, at least for a risk-averse expected-utility maximizing investor, because we assume price dynamics with stationary statistics. Independent multiagent search boosts the rate of convergence of the overall portfolio of the system towards the optimal portfolio in both markets, compared with the rate of convergence to the optimal portfolio by a single agent's portfolio-selection strategy.

Cooperative multiagent search provides a further performance improvement in the simple market. Hint exchange and portfolio switching helps to eliminate bad portfolios from the population in early investment periods. In comparison, agent communication has a negligible effect on system performance in the CAPM market, in which price dynamics already reflect implicit communication between many agents and all balanced portfolio-selection strategies perform quite well.

Although a single agent can in principle achieve the same performance as a multiagent system by simulating the investment strategy of the entire system, because all agents receive the same stock price information, multi-agent portfolio selection is useful for *bounded-rational* agents with limited computational and information processing resources. A single bounded-rational agent is unable to simulate the entire system and can benefit from an exchange of information on the performance of the strategies of other agents. Bounded-rational constraints are especially relevant for investors in a real market, in which there is a large variety of stocks and financial instruments and many different sources of information in addition to price information. Indeed, investment decisions are seldom made by a single investor in isolation, but only after extensive consultation and research.

Here is an outline of the paper. In Section 2 we define the multiperiod portfolio selection problem, and introduce two common approaches to solve the problem, *model-based* and *model-free* portfolio selection. Model-free portfolio selection has a number of advantages, and we provide each agent in our multiagent model with a model-free strategy. Section 3 defines the individual agent portfolio selection strategy, and demonstrates its performance in a simple example. In Section 4 we introduce our multiagent portfolio-selection model. We present its performance in the simple market in Section 5, and its performance in the CAPM market in Section 6. We summarize our results in Section 7, before presenting our main conclusions. The appendix contains proofs and algorithmic descriptions of each multiagent model.

1.1. Related Work

Cooperative search has been applied to hard computational problems in artificial intelligence with agents that have diverse search heuristics [13, 26]. Agents exchange useful information to avoid redundant search and accelerate problem solving. The communication of hints between agents can be more sophisticated than direct imitation because hints derived from one problem-solving heuristic can be introduced into a different problem solving heuristic. A general theory predicts superlinear speedup in the performance of individual agents when the search methods are diverse and the agents are able to utilize information found in other parts of the search space [28]. The current problem of portfolio selection is interesting because it is a stochastic *online* decision problem [30], agents must invest as they receive incremental information about stock prices.

Many other problem-solving techniques that are related to cooperative search have been proposed for solving hard computational problems, including: sequential restart strategies with diverse heuristics [52, 41, 31, 10]; and parallel independent search with stochastic search algorithms [45, 40, 34, 36, 29]. Knight [35] compares the performance of a system of many co-

operative agents with simple search heuristics to a system of a few agents with more complex search heuristics, and Aldous and Vazirani [1] describe a cooperative search technique called "Go with the winners".

Game theorists have proposed a model of *social learning*, or learning by imitation, to generate solutions in coordination games [18, 7]. This is related to our model of cooperative search. Agents share information as they learn, about their recent strategies and payoffs received, and take advantage of information about the payoffs and strategy choices of other agents with similar goals.

We model bounded-rational agents that can benefit computationally from cooperation with other agents because they are unable to compute optimal investment strategies directly. The agents in our model use a first-order approximation to a worst-case optimal portfolio-selection rule, which is similar to the approach to bounded-rationality in game-theory, placing a static constraint on the complexity on agents [47]. In comparison, economic models of metadeliberation select a level of deliberation within a decision-theoretic framework, based on the expected-value of further deliberation [22, 54, 48].

In contrast to the recent literature on bounded-rational learning in games [32, 44], we assume in the portfolio selection problem that an agent's opponent (the market) plays the same strategy for all agent strategies. Prices do not depend on investment actions. Furthermore, there is no exploration versus exploitation problem, as occurs for example in the classic Multi-armed bandit problem [46, 5], because we assume that all agents receive the same price information, irrespective of their portfolio selection.

The usual emphasis in game theory is on model-based learning, for example Fictitious Play [20], and other models of myopic best-response dynamics [57, 33], where agents play a best-response to a model that they learn of their opponents. Recent models of multiagent learning within artificial intelligence provide a hierarchy of agent models and allow *strategic learning* [21, 56, 55], where agents take advantage of models of the learning of other agents. Instead, we follow the framework of *model-free learning*: the agents in our model do not maintain an explicit model of the stock market. The current portfolio of an agent represents the cumulative learning of the agent, and the agent makes a small adjustment to its portfolio every time it observes new stock prices [25]. For an example of model-free learning in games, see Sandholm and Crites' [50] application of Q-learning to the classic Prisoners' Dilemma game.

The portfolio selection problem has different characteristics than the multiagent load-balancing problem studied by Schaerf *et al.* [51]. In that study communication between agents reduces performance because a resource which is lightly loaded when used by a single agent becomes heavily loaded when used by many agents. Communication reduces the heterogeneity of

agent decisions and leads to unbalanced loads and system instability. In our model of portfolio selection the stock prices are exogenous and independent of investment actions. Hence, all agents can achieve a good performance even if they all follow the same strategy.

A similar comparison can be made with work in the agent-based computational economics (ACE) literature. ACE studies the dynamics of prices generated *endogenously* through the actions of many simple agents [37, 19, 4, 38], and builds markets from the "bottom up" in order to understand the connection between simple agent actions and price dynamics. We consider investment in a large market with stock price dynamics that are *independent* of the investment decisions of the investment group.

2. THE MULTIPERIOD PORTFOLIO SELECTION PROBLEM

In general terms, the multiperiod portfolio selection problem is to invest over a sequence of periods to maximize some measure of performance over the final return-on-investment.

Consider a market of N stocks with $t = 1, \ldots, T$ discrete investment periods. Let x_i^t denote the *price relative* of stock i in period t, the ratio of closing price to opening price over the period. This is non-negative by definition. The vector $\mathbf{w} = (w_1, \ldots, w_N)$, where $w_i \geq 0$ and $\sum_{i=1}^{N} w_i = 1$ defines a portfolio, where w_i is the fraction of total investment in stock i. The *return-on-investment* in a single period, for an agent with portfolio $\mathbf{w}$ and price relatives $\mathbf{x} = (x_1, \ldots, x_N)$, is given by the weighted sum over all stocks, $\mathbf{w} \cdot \mathbf{x} = \sum_{i=1}^{N} w_i x_i$.

In multiperiod portfolio selection over T investment periods the return on investment[1], $\mathcal{R}$, for an agent with a sequence of portfolios, $\{\mathbf{w}^T\} = \mathbf{w}^1, \ldots, \mathbf{w}^T$, and a sequence of stock price relatives, $\{\mathbf{x}^T\} = \mathbf{x}^1, \ldots, \mathbf{x}^T$, is the product of single period returns, $\mathcal{R} = \prod_{t=1}^{T} \mathbf{w}^t \cdot \mathbf{x}^t$. The goal of multiperiod portfolio selection is to select a sequence of portfolio strategies, $\{\mathbf{w}^T\}$, to maximize a measure of performance over the final return-on-investment. The appropriate performance metric depends on an agent's risk-return preferences.

The *offline* multiperiod portfolio selection problem with knowledge of the sequence of stock prices is trivial. The optimal strategy with hindsight switches all investment at the start of each period to the stock with the greatest return in that period. In our main experimental results the performance of the optimal constant portfolio with hindsight provides a

---

[1] We use *return-on-investment* and *wealth* interchangeably in this paper because we assume throughout that the total initial wealth of all systems of agents is $1.

useful benchmark for the performance of our multiagent portfolio selection models.

The *online* multiperiod portfolio selection problem is hard because future stock prices are unknown and an agent must choose a portfolio $\mathbf{w}^t$ for investment period t without knowledge of the price relatives $\mathbf{x}^t$. This is an *online decision problem* [30] because decisions must be made as new information arrives.

There are two common approaches to online multiperiod portfolio selection: model-based and model-free portfolio selection. In *model-based portfolio selection*, agents have access to a statistical model of stock price dynamics. This allows the problem to be formulated and solved as a stochastic optimization problem. In *model-free portfolio selection*, agents have no statistical model of stock dynamics and stock prices can be arbitrary sequences. The agents in our model of portfolio selection follow a model-free portfolio selection strategy.

2.1. Model-Based Portfolio Selection

Briefly, model-based portfolio selection assumes a statistical model of stock price dynamics, based for example on the past performance of the market, and agents choose a sequence of portfolios to maximize expected utility over return-on-investment. An agent learns a model of its environment, and then plays a best-response to that model.

Non-linear stochastic dynamic programming techniques can solve the portfolio selection problem directly for a restricted class of utility functions and market models [6]. The optimal portfolio strategy depends on the risk preferences of the agent, as represented by a *utility function* over return-on-investment [11]. A good investment strategy trades off expected return and variance to maximize expected utility. For example, a concave-increasing utility function represents a risk-averse agent which is prepared to reduce expected return in favor of lower risk.

Although model-based approaches can be computationally tractable, their performance depends on the accuracy of the underlying stock market model. Parameter estimation for a stochastic economic model is a difficult problem [11, 12]. Furthermore, solution techniques assume stock price dynamics with stationary statistics, and investment strategies are not robust to shocks.

2.2. Model-Free Portfolio Selection

Having noted the limitations of model-based portfolio selection, we now introduce *model-free portfolio selection*, which makes no assumptions about the underlying stock prices and avoids the parameter estimation problem [14]. Agents learn an optimal portfolio-selection strategy directly, without forming an explicit model of stock price dynamics.

An immediate problem in model-free portfolio selection is how to measure performance. Average-case analysis, which makes claims about the expected performance of a strategy, is not meaningful without a statistical model. Furthermore, all strategies have bad worst-case performance, consider an adversary that chooses stock prices such that the stocks held in the portfolio in each investment period devalue.

A useful technique in the design and analysis of online algorithms is *competitive analysis*, in which performance is measured by comparison with an optimal offline algorithm which takes the same decisions but has information about all future inputs. A competitive algorithm must only perform well relative to the difficulty of a problem instance, as measured by the performance of the offline algorithm. A number of portfolio selection strategies exist with *competitive performance* with respect to a class of offline strategies.

In this paper we provide each agent with a *strongly-competitive* portfolio selection strategy [25]. A strongly-competitive strategy has an optimal worst-case performance guarantee. We assume in the following definition that the decision problem is a *maximization problem*.

Competitive analysis requires a *comparison set* of algorithms from which to choose the optimal offline algorithm, and a *performance metric*, $Perf_{comp}$. Let $online(\{\mathbf{x}^T\})$ denote the return-on-investment from an online portfolio selection strategy, and $offline(\{\mathbf{x}^T\})$ denote the return-on-investment from an optimal offline portfolio selection strategy (that can invest with hindsight of all future stock prices), perhaps constrained to a comparison set.

The online portfolio strategy is strongly competitive with respect to performance-measure $Perf_{comp} : \mathbb{R} \to \mathbb{R}$, such that the agent wants to maximize $Perf_{comp}(\mathcal{R})$ given return $\mathcal{R}$, when:

DEFINITION 2.1. Strongly competitive. Online algorithm $online(\{\mathbf{x}^T\})$ is strongly-competitive with respect to performance measure $Perf_{comp}$ if its worst-case performance is equal to the optimal offline algorithm $offline(\{\mathbf{x}^T\})$ in the long-term:

$$\lim_{T \to \infty} \min_{\{\mathbf{x}^T\}} \left[\frac{Perf_{comp}(online(\{\mathbf{x}^T\}))}{Perf_{comp}(offline(\{\mathbf{x}^T\}))} \right] = 1$$

where the minimization is over all feasible input sequences, $\{\mathbf{x}^T\}$ of length T.

An offline portfolio-selection strategy can invest with *hindsight*, that is with information about all future stock price changes. The optimal *unrestricted* offline investment strategy, which shifts all investment at the start

of each period to the single stock that will show the greatest return in that period, does not provide a useful benchmark.

Cover [14] proposed a model-free portfolio-selection algorithm, UNIVERSAL, and demonstrated that it is strongly-competitive with the set of *constant rebalanced portfolios* (CRP) in terms of *per-period return-on-investment*; performance is defined as $Perf_{\text{comp}}(\mathcal{R}) = \mathcal{R}^{1/T}$, for return $\mathcal{R} = \Pi_{t=1}^{T} \mathbf{w}^t \cdot \mathbf{x}^t$ after T investment periods.

A strongly-competitive portfolio-selection strategy, such as UNIVERSAL, achieves the same long-term per-period return as the best offline CRP for any sequence of stock prices. In Section 3 we introduce a simple model-free portfolio-selection rule from Helmbold *et al.* [25] with the same property. The rule is followed by individual agents in our multiagent system.

The offline portfolio strategies are constrained to the set of constant rebalanced portfolios (CRP), which are multiperiod portfolio strategies that maintain the same portfolio across all periods. An agent with a constant rebalanced portfolio trades in each period to rebalance its investment, selling stocks that outperform the portfolio and buying stocks that underperform the portfolio.

DEFINITION 2.2.   Best offline constant rebalanced portfolio. The best offline CRP, $\mathbf{w}^*_{\text{CRP}}$, computed with complete information on the sequence of stock prices, $\{\mathbf{x}^T\} = \mathbf{x}^1, \ldots, \mathbf{x}^T$, maximizes final return-on-investment:

$$\mathbf{w}^*_{\text{CRP}}\left(\{\mathbf{x}^T\}\right) = \arg\max_{\mathbf{w}} \prod_{t=1}^{T} \mathbf{w} \cdot \mathbf{x}^t \tag{1}$$

where the maximization is over all constant rebalanced portfolios.

The offline problem is *deterministic*, and the objective can be accurately stated in terms of return alone, irrespective of an agent's risk-preferences.

2.3.   An Economic Interpretation of Competitive Portfolio Selection

The performance of the *long-term* optimal offline CRP, denoted $\mathbf{w}^*$, which solves (1) as $T \to \infty$, provides a good benchmark because: its return is at least as large as the return-on-investment from the best single stock, since buy-and-hold of a single stock is a special case of a CRP; and its return is at least as large as the return-on-investment from the best online strategy when price changes are independent and identically distributed from period to period [2]. This is quite surprising, given that we only allow offline strategies that are constant rebalanced portfolios.

In a market with stationary, independent and identically distributed price relatives $\mathbf{x}^t$ from period-to-period, the long-term optimal offline CRP,

$\mathbf{w}^*$, is well-defined. It maximizes the single-period expected log return-on-investment (see Appendix A).

When we also make a common assumption in the literature on financial optimization that agents have a *logarithmic* utility function for return-on-investment, $u_i : \mathbb{R} \rightarrow \mathbb{R}$, such that $u_i(\mathcal{R}) = \log(\mathcal{R})$, then there is an *economic interpretation* of the performance of the long-term optimal CRP over a finite number of investment periods.

A logarithmic utility function represents the preferences of a risk-averse investor, and is useful because it allows tractable analysis. With a utility function the performance of a portfolio selection strategy can be measured in terms of expected utility.[2]

DEFINITION 2.3. Performance measure. The performance, *Perf*, of a multiperiod portfolio strategy $\{\mathbf{w}^T\}$ for an agent with a logarithmic utility for return-on-investment is:

$$Perf = E_{\{\mathbf{x}^T\}} \left[\log \left(\prod_{t=1}^{T} \mathbf{w}^t \cdot \mathbf{x}^t \right) \right]$$

where the expectation is taken over sequences of price relatives, $\{\mathbf{x}^T\}$, distributed according to market price dynamics, and T is the number of investment periods.

The long-term optimal offline CRP maximizes the expected utility of an agent for a finite number of investment periods:

THEOREM 2.1. *The long-term optimal offline constant rebalanced portfolio, $\mathbf{w}^*$, in a market with non-negative, independent and identically distributed price relatives, maximizes expected utility after any finite number of investment periods for an agent with a logarithmic utility for return-on-investment.*

The long-term optimal portfolio, $\mathbf{w}^*$, also lies on the *efficient frontier* [42]. Markowitz [42] introduced a single period mean-variance approximation to simplify portfolio selection; with "risk" quantified as the standard deviation of return from period to period, and "return" quantified as the expected single-period return. With this approximation, portfolio selection reduces to the selection of a portfolio on the efficient frontier for a particular return:

---

[2] The reader should be careful not to confuse this term with $Perf_{comp}$, the metric to measure the strong-competitiveness of a strategy.

DEFINITION 2.4.   The efficient frontier. The efficient frontier is the set of all portfolios that minimize risk for some level of return.

It is not optimal to merely invest in the single stock with the highest return. The variance in return from period to period is also important because wealth is reinvested at the start of each period, and also because agents tend to be risk-averse.

THEOREM 2.2. *The long-term optimal offline constant rebalanced portfolio, $\mathbf{w}^*$, in a market with non-negative, independent and identically distributed price relatives lies on the efficient frontier.*

Given this analysis, a model-free portfolio-selection strategy that is strongly competitive with the long-term optimal CRP, $\mathbf{w}^*$, should have useful economic properties, so long as it converges quickly enough to $\mathbf{w}^*$.

3. A COMPETITIVE PORTFOLIO-SELECTION STRATEGY

Each agent in our multiagent model for portfolio selection follows an approximation, χ^2, to a model-free portfolio-selection rule EG which is strongly competitive with the best offline constant rebalanced portfolio [25].

The update-rule, χ^2, adjusts the portfolio on the basis of its recent performance and recent price dynamics. It has worst-case time and space complexity that is *linear* in the number of stocks.[3] It was developed within a framework of multiplicative updates for online prediction in machine learning theory [39].

DEFINITION 3.1.   The χ^2 Portfolio-Selection Rule. Choose an initial portfolio at random. The portfolio in period $t + 1$, portfolio $\mathbf{w}^{t+1}$, is computed from the current portfolio, $\mathbf{w}^t$, and the price relatives in the most recent investment period, $\mathbf{x}^t$:

$$w_i^{t+1} = w_i^t \left(\eta(\frac{x_i^t}{\mathbf{w}^t \cdot \mathbf{x}^t} - 1) + 1 \right)$$

where $\eta > 0$ is the *learning rate*.

The rule increases the fraction of wealth invested in stocks that outperform the portfolio and decreases investment in stocks that underperform

---

[3] In comparison UNIVERSAL has exponential worst-case time and space complexity in the number of stocks.

the portfolio, notice that $w_i^{t+1} > w_i^t \iff x_i^t > \mathbf{w}^t \cdot \mathbf{x}^t$. It is model-free because the strategy is updated directly from the price changes, without forming an explicit model of the stock market dynamics. The current portfolio strategy, $\mathbf{w}^t$, implicitly represents the cumulative information that an agent has learned about the stock price dynamics up to period t.

Portfolio update with χ^2 is a compromise between long-term learning, retaining information about previous stock-price dynamics, and responsiveness, moving in a direction that will give a better performance if price relatives in the current period characterize future periods. The learning rate, η, determines how this tradeoff is made (see Section 5).

Recall that a constant rebalanced portfolio (CRP) maintains the same proportion of wealth invested across all stocks in every investment period, trading to sell stocks that outperform the portfolio and buy stocks that underperform the portfolio. The χ^2 portfolio-selection rule converges towards the optimal offline CRP over time in a market with stationary statistics.

However, strongly-competitive performance with metric $Perf_{\text{comp}}(\mathcal{R}) = \mathcal{R}^{1/T}$ is not sufficient for optimal expected end-period utility. Although, Theorem 2.1, the optimal offline portfolio $\mathbf{w}^*$, maximizes end-period expected utility, the performance of χ^2 depends on the speed-of-convergence to $\mathbf{w}^*$.

We show that our multiagent model, in which agents each follow local χ^2-update rules from a random initial portfolio, achieves better expected utility because it boosts the rate of convergence of the overall portfolio in the system towards the optimal portfolio. An agent that quickly adjusts its portfolio to the optimal CRP achieves a greater return-on-investment than an agent that adjusts its portfolio more slowly.

3.1. Example: Single-Agent Portfolio Selection

This example shows the effect of investment by a single agent with the χ^2 portfolio-selection rule in a simulated market. The best offline CRP exponentially outperforms the best single stock buy-and-hold policy. Furthermore, χ^2 tracks the wealth from the best CRP to within a constant logarithmic difference and exponentially outperforms the best single stock buy-and-hold policy.

Consider 2 stocks with Normally distributed price relatives, $x_i \sim N(\mu_i, \sigma_i^2)$, with mean σ_i and variance σ_i^2. This is the standard *geometric Brownian motion* model of stock price dynamics (see Section 5 for more details). Recall that the price relative is the ratio of price in period $t + 1$ to price in period t. In this example stock 1 is generated with price relatives $x_1 \sim N(1.005, 0.1)$, and stock 2 with price relatives $x_2 \sim N(1.0005, 0.05)$. Stock 1 has a high expected single period return, and a high volatility across periods, while stock 2 has a low expected single period return and a
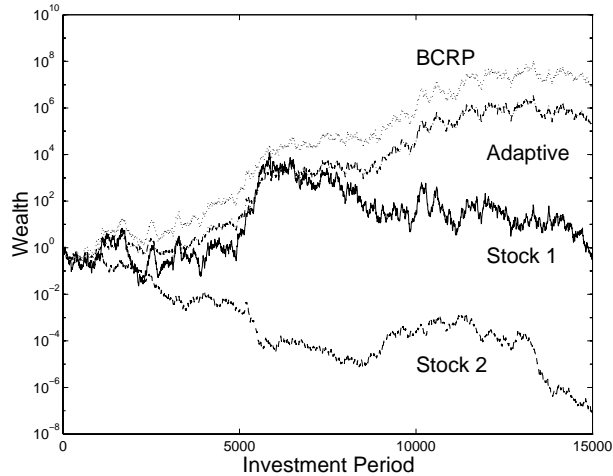
FIG. 1. Wealth (log-scale) versus investment period for the best offline constant rebalanced portfolio (Bcrp), the portfolio selection rule χ^2 (adaptive), and buy-and-hold in each stock. Stock 1 closes at \$0.69, stock 2 at \$0.00, while the final wealth from χ^2 is \$218,000 and from the best Crp is \$9,860,000.

low volatility across periods. This partial correlation between return and risk across stocks is typical of real markets.

Figure 1 plots final wealth from the χ^2 rule, buy-and-hold in each stock, and the best offline Crp, for a particular sequence of simulated stock prices. The χ^2 rule exponentially outperforms both stocks (note that wealth is plotted on a log-scale). Stock 1 closes at \$0.69, and stock 2 at \$0.00, both from an initial price of \$1.00, while the final wealth of the agent is \$218,000. Remember that the final wealth from the best offline Crp, \$9,860,000, is unattainable. The adaptive agent is able to maintain a constant logarithmic difference between its wealth and the wealth of the best Crp; this indicates its strongly competitive performance.

Figure 2 shows that the portfolio selected by χ^2 (solid) tracks the best Crp (dashed), and converges to the best long-term offline Crp, which is $\mathbf{w}^* = (0.547, 0.453)$ for stocks with these statistics. The best offline Crp is plotted incrementally, for prices up to period t, to provide a comparison with the portfolio selected by the online portfolio selection rule in each period.

In Figure 3 we plot the single-period risk-return characteristics for all portfolios. The best long-term Crp lies on the efficient frontier, the set of portfolios that minimize variance in period-to-period return for some expected period-to-period return. It is interesting that a model-free portfolio-selection rule, such as χ^2, can select a portfolio that lies on the efficient frontier without learning an explicit model of market price dynamics.
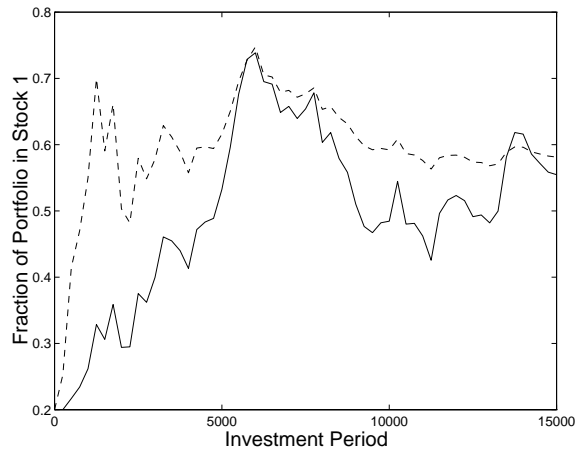
FIG. 2. The χ^2 portfolio (solid) and the best offline constant rebalanced portfolio (dashed). The best offline CRP is computed incrementally for each investment period t, on the basis of the stock price information up to period t.
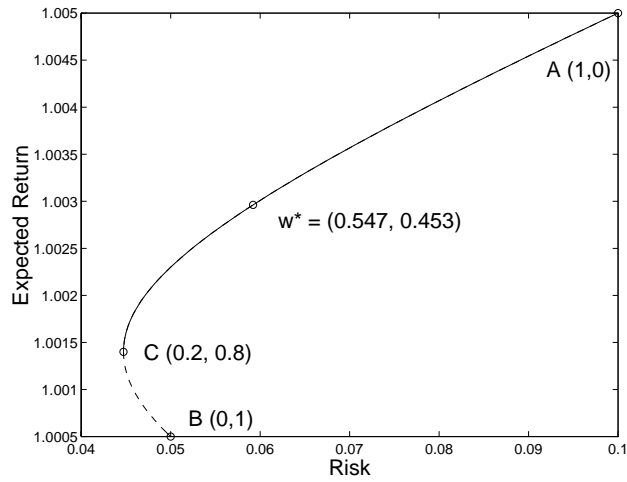

FIG. 3. Expected single period return-on-investment versus variance in return (risk). The efficient frontier, the set of portfolios that minimize risk for some level of return, is illustrated with a solid line between C and A. The long-term optimal constant rebalanced portfolio, $\mathbf{w}^* = (0.547, 0.453)$, selected by χ^2, lies on the efficient frontier.

4. COOPERATIVE MULTIAGENT SEARCH

We propose a new multiagent model for portfolio selection which combines the investment decisions of a system of agents that follow local χ^2 portfolio-selection strategies. We model an "investment group", in which agents combine their initial wealth and divide it between the agents. Individual agents make autonomous investment decisions in each investment period, before sharing profits and losses at the end of the investment.

It is useful to define the *overall portfolio* $\mathbf{w}_e^t$ of a system of agents in round t. This is the single portfolio with the same return-on-investment as the joint return from each agent's portfolio. It is computed as the weighted average of each agent's portfolio, with weight proportional to an agent's wealth:

DEFINITION 4.1.   Overall Portfolio. Given portfolio, $\mathbf{w}_i^t$, for agent i in period t, the overall portfolio, $\mathbf{w}_e^t$ is

$$\mathbf{w}_e^t = \sum_{i=1}^{M} \left(\frac{wealth_i^t}{\sum_{j=1}^{M} wealth_j^t} \right) \mathbf{w}_i^t$$

where $wealth_i^t$ is the wealth of agent i at the start of period t, and there are M agents.

Here are brief descriptions of the three models of multiagent portfolio selection. See Appendix B for algorithmic descriptions of each model.

Non-adaptive Independent Search

First, we consider a very simple multiagent system that performs *non-adaptive independent search*, in which agents choose initial portfolios at random and invest in the same portfolio for all investment periods, trading to rebalance the portfolio as necessary. The non-adaptive independent system provides a performance baseline. It separates the effect of agent heterogeneity (from selecting random initial portfolios) and the effect of single-agent learning. In fact, the system is a multiagent approximation to Cover's UNIVERSAL portfolio-selection algorithm. The approximation is exact in the limit, as the number of agents gets large [9].

Given N stocks, in each model we select initial portfolios for agents at random from the Dirichlet$(1/N, \ldots, 1/N)$ distribution, which is a generalization of the uniform distribution to the space of feasible portfolios which generates N-dimensional vectors with non-negative components that sum to one and mean $(1/N, \ldots, 1/N)$. Each agent trades to rebalance its portfolio and maintain its initial portfolio across all investment periods. An

agent sells stocks that outperform the portfolio and buys stocks that underperform the portfolio.

Independent Search

Then, we consider a system of *independent search*, in which each agent follows the χ^2 portfolio-selection rule from a random initial portfolio. The independent search model combines agent heterogeneity with individual agent learning.

Each agent adjusts its portfolio across investment periods with the χ^2 portfolio-update rule, which is initialized with a learning-rate from a uniform distribution, $\eta_i \sim U(\eta_l, \eta_h)$, where η_l and η_h are lower and upper bounds. Parameters η_l and η_h are selected offline to provide reasonable performance across all experiments.

The overall portfolio of the independent search model remains strongly competitive (Definition 2.1) if individual agents have strongly competitive portfolio-selection strategies.

THEOREM 4.1. *The overall portfolio-selection strategy of the independent multiagent search model is strongly competitive if individual agents have strongly competitive portfolio-selection strategies.*

Cooperative Search

Finally, we consider a model of *cooperative search*, which introduces *explicit communication* between agents to the model of independent search. Agents can exchange information about the recent performance of their portfolios, and switch probabilistically to the portfolio with the best performance. The cooperative search model is designed to speed up multiagent search for a good portfolio, through discontinuous updates in the portfolios of individual agents towards portfolios which are performing well in the population of agents.

The current portfolio and learning rate of an agent with the χ^2 rule defines its future portfolio selection for any sequence of stock price movements. Therefore, when an agent in the cooperative search model switches to the portfolio of another agent, both agents follow the same future portfolio investments to the extent that the agents have the same learning-rates, at least until either agent switches to another portfolio.

Each agent adjusts its portfolio across investment periods with the χ^2 portfolio-update rule, and also announces the recent performance of its portfolio strategy and switches probabilistically to the best system-wide portfolio.

In particular, each agent maintains the average return of its recent investment strategy over a finite number of recent periods, τ, its *performance*

window size, and posts its current portfolio and recent performance to a central *blackboard* at the end of every period. The blackboard maintains the portfolio that is performing best over all the agents. If an agent's own portfolio is performing worse than the best system-wide portfolio it switches to that portfolio with probability p, its *switching probability*. Agents only post to the blackboard and test the blackboard for hints if they have not switched portfolio for at least τ periods. This prevents thrashing of agent strategies and avoids early lock-in to a single strategy.

In our simulations we provide all agents with the same window size and switching probability. The parameters are optimized offline for each set of problems.

4.4. Market Models and Experimental Tests

Experimentally, the models are tested in two different markets. The first is a simple stochastic market with independent stock price dynamics across stocks, the second is a more realistic market that approximates the Capital Asset Pricing Model (CAPM) market and models equilibrium price dynamics [53]. In both markets, the investment groups are assumed to be small with respect to the total market, and we treat prices as exogenous variables. Independent search performs better than a single agent in both markets, and cooperative search outperforms independent search in the Simple market.

In Sections 5 and 6 we show that the structure of the search problem depends on the market statistics, and explain why cooperative search outperforms independent search in the Simple market but is less useful in CAPM. A system which is selects an overall portfolio that converges quickly towards the optimal offline portfolio has a good overall performance, and the location of the optimal portfolio in the overall search space depends on the market statistics.

5. PERFORMANCE IN A SIMPLE MARKET MODEL

In this section we consider a simple non-equilibrium stock market, called the *Simple market*. In this market each stock is a geometric Brownian motion stochastic process with price relatives (ratio of prices in successive periods) independent and identically distributed according to a Normal distribution, i.e. $x_i \sim N(\mu_i, \sigma_i^2)$ with mean μ_i and standard deviation σ_i. This model is often used in theoretical studies of investment strategies [17].

The mean and standard deviation for each stock are selected independently in each trial from uniform distributions $\mu_i \sim U(\mu_l, \mu_h)$ and $\sigma_i \sim U(\sigma_l, \sigma_h)$, for lower and upper bounds μ_l, σ_l and μ_h, σ_h. The dynamics of price changes are independent across stocks and there is no correlation between return and risk across stocks.

5.1. Experimental Details

The performance of each investment model is tested in a market with 10 stocks and an investment of duration 2000 periods, with means and standard deviations for each stock drawn from distributions $\mu_i \sim U(0.9995, 1.01)$ and $\sigma_i \sim U(0.0, 0.2)$. These statistics are appropriate for the monthly returns on real stocks. For example, the mean monthly return on stock in IBM between 1962 and 1994 was 1.0081, and the standard deviation in monthly return was 0.062 [12, Page 21]. In each trial we first selected the market parameters, $(\mu_1, \dots, \mu_N)$ and $(\sigma_1, \dots, \sigma_N)$, and then generated a sequence of stock prices. The performance of all multiagent models is compared for the same stock prices.

The systems are tested with between 1 and 800 agents, to study the relationship between the number of agents in an investment group and its performance. Performance is measured as the average end-period log return-on-investment across 2000 independent trials. We assume expected-utility maximizing agents with logarithmic utility functions (Definition 2.3).

A random initial portfolio is generated for each agent in each trial, $\mathbf{w}^1 \sim$ Dirichlet$(1/N, \dots, 1/N)$, and a random learning rate, $\eta \sim U(0.1, 0.15)$, is assigned. This distribution of learning rates gives a good performance across all sizes of models. The switching probability and performance window size are the same for every agent, and optimized for the number of agents, with switching probability $p = 0.004$ and performance window $\tau = 200$ typical.

5.2. Results in the Simple Market

The performance of each multiagent portfolio-selection model is compared in Figure 4. The best offline CRP with hindsight of stock price movements, computed in each trial with an algorithm due to Helmbold *et al.* [24], achieved *Perf*(BCRP) = 16.0 in this market.

The experimental results show that:

[a] A single adaptive agent outperforms a single non-adaptive agent.

[b] A system of non-adaptive agents (independent non-adaptive search) outperforms a single non-adaptive agent.

[c] A system of adaptive agents (independent search) outperforms a single adaptive agent and a system of non-adaptive agents.

[d] Hint exchange and strategy switching (cooperative search) provides a further increase in performance, and the value of communication increases as the number of agents increases. Cooperative search outperforms independent search when there are more than 50 agents.[4]

---

[4] The null hypothesis that the mean end-period log wealth for a system of communicating agents and a system of non-communicating agents is equal is rejected with a significance level of less than 0.01 for systems with more than 50 agents.
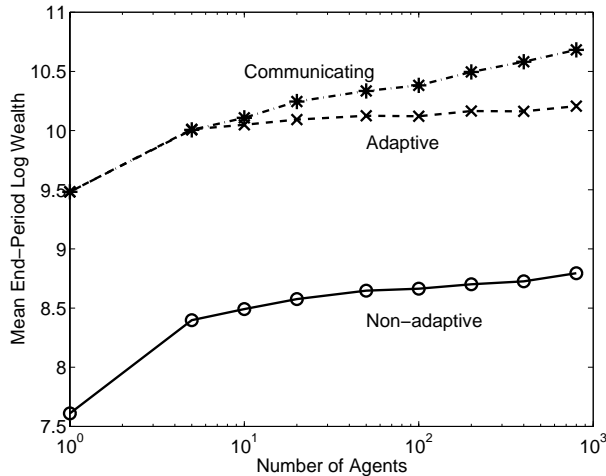
FIG. 4. Simple market. The performance of the non-adaptive independent search (non-adaptive), independent search (adaptive) and cooperative search (communicating) models.

Figure 5 (a) shows an example of the effect of introducing adaptive agents. It plots the performance of a system with 100 non-adaptive independent agents (line) and a system with 100 adaptive independent agents (dots). Each data point corresponds to the overall performance of a system in a single trial, and the trials are sorted by the final wealth of the non-adaptive system of agents for clarity. The system of adaptive agents outperforms the system of non-adaptive agents, achieving a better return-on-investment in almost every trial.

Figure 6 (a) shows an example of the effect of introducing cooperative search. It plots the ratio of final wealth of the cooperative search model to the independent search model, for 400 agents with $\tau = 200$ and $p = 0.004$. The cooperative search system achieves a better return in 75% of the trials, with a final return-on-investment 1.47 times greater on average.

5.3.   Analysis for the Simple Market

Result [a] can be explained with existing theory; we expect a single adaptive agent to outperform a single non-adaptive agent from the analysis and empirical results for χ^2 in Helmbold *et al.* [25]. Similarly, we expect a system of independent non-adaptive agents to outperform a single non-adaptive agent, result [b], because the system implements a randomized approximation to UNIVERSAL, a strongly-competitive portfolio selection strategy [9].

Results [c] and [d] demonstrate new and interesting effects. The independent search model, which combines individual-agent learning with di-
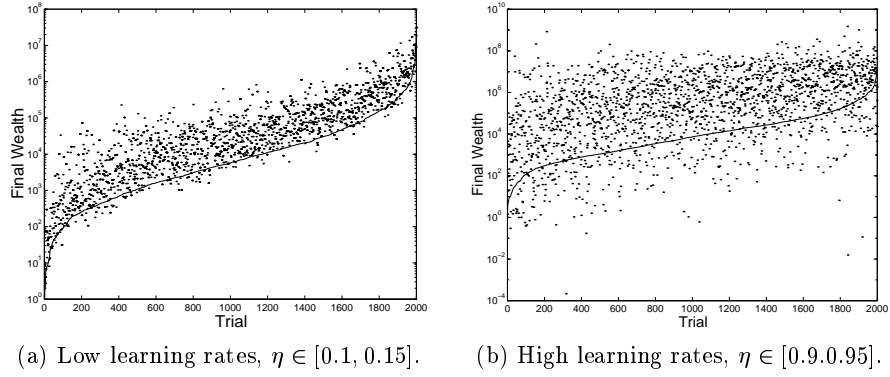
(a) Low learning rates, $\eta \in [0.1, 0.15]$.

(b) High learning rates, $\eta \in [0.9.0.95]$.

FIG. 5. Simple market. Final wealth in 2000 trials of a system of independent search (dots) and a system of non-adaptive independent search (line, for investment groups of 100 agents. The trials are sorted by the final wealth of the non-adaptive agents.
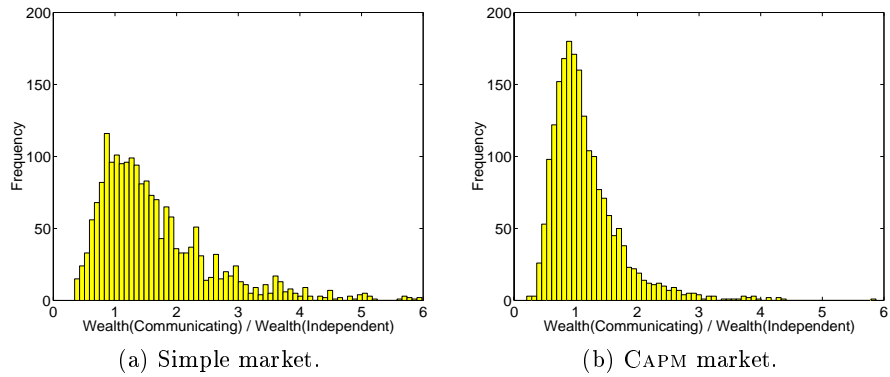

(a) Simple market.

(b) CAPM market.

FIG. 6. Distribution over 200 trials of the ratio of the final wealth of a system of cooperative multiagent search (communicating) to final wealth of a system of independent multiagent search (independent), with 400 agents. In the Simple market communication improves final wealth in 75% of the trials, with an average wealth 1.47 times greater. In the CAPM market communication improves final wealth in 53% of the trials, with an average wealth 1.05 times greater.

versification from the random initial portfolios of each agent, outperforms both a system of non-adaptive agents and a single adaptive agent [c]. The effect on performance from adaptive individual-agent portfolio selection appears to be independent and additive of the effect from agent diversification with initial random portfolios. The difference in performance between non-adaptive and adaptive independent search is approximately constant across all numbers of agents (Figure 4). Furthermore, introducing communication and strategy switching improves performance [d].

We can interpret the results from a search perspective. Let us consider the rate of convergence of the overall portfolio selected by each multiagent model towards the optimal offline CRP. Recall that each individual agent uses the χ^2 rule, which converges to the optimal offline CRP over time. In fact, it is the speed of convergence towards the optimal offline CRP that determines the performance of a system of agents.

In Figure 7 we plot the average distance of the overall portfolio of a system of 200 agents to the best offline CRP in each investment period, for each multiagent investment model. The L_2 norm

$$\mathcal{D}(\mathbf{u}, \mathbf{v}) = \left(\sum_{i=1}^{N} (u_i - v_i)^2 \right)^{1/2}$$

is used to compute the distance between two portfolio vectors, $\mathbf{v}$ and $\mathbf{w}$, which is denoted $\mathcal{D}(\mathbf{u}, \mathbf{v})$.

The ability of each system to select a portfolio that is close to the best offline CRP appears to be a good indicator of its performance: there is a strong relationship between the final average distance of the overall portfolio to the best offline CRP and the performance of each multiagent system. Although the portfolios selected by all systems converge towards the best CRP (even for the non-adaptive agents), explicit cooperative search with communication between the agents boosts the rate of convergence, especially during the early periods when agents with bad strategies still have a large proportion of wealth.[5]

A comparison with the performance of the *market portfolio*, a simple buy-and-hold strategy across all stocks, provides further insight.

DEFINITION 5.1.   Market portfolio.   The market portfolio is a simple buy-and-hold strategy across all stocks, with initial investment equally distributed across all stocks.

---

[5] No agents switch strategy during the first 200 periods because the performance window is $\tau = 200$, and agents must wait τ periods to switch.
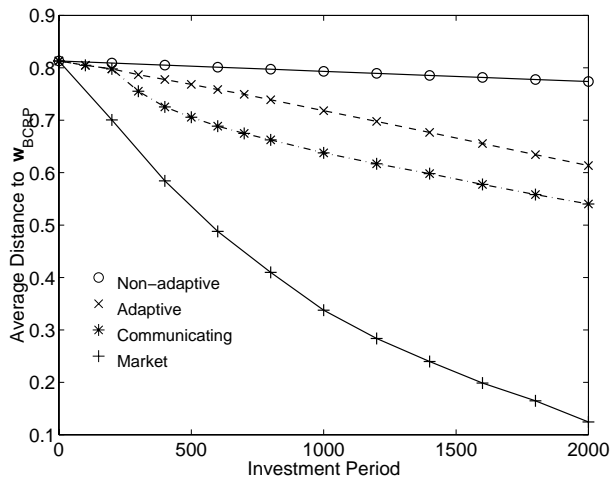
FIG. 7. Simple market. The average distance in each period between the portfolio selected in each system and the best offline constant rebalanced portfolio, in systems with 200 agents. Communication boosts convergence of the overall portfolio of the multiagent portfolio selection model towards the optimal portfolio in the early investment periods. The Market portfolio converges very quickly to the best CRP.

Surprisingly, we found that the performance of the market portfolio dominates the performance of the other strategies, with $Perf(Market) = 14.3$. Indeed, Figure 7 shows that the market portfolio is the most effective at selecting the optimal CRP in the Simple market. Simple statistical analysis explains this result.

Figure 8 shows that the value of the maximum component of the best offline CRP (the greatest weight of investment in any single stock) is often very close to one. All components are non-negative and sum to one, hence the best CRP is typically very close to a single stock buy-and-hold strategy. It is very likely that there is a single stock with a high return and a low volatility in the simple market because we select the mean and standard deviation parameters for the price distributions independently.[6]

The market portfolio performs well because the best CRP is often approximately a single stock, and the market portfolio is provably competitive with the best single stock [9]. The market portfolio shifts towards the stocks with the best performance over time, as the investment in stocks

---

[6] Given 10 stocks with $\mu_i \sim U(0.9995, 1.01)$ and $\sigma_i \sim U(0, 0.2)$, there is a probability of 0.0643 that a single stock has $\mu > 1.0073$ and $\sigma < 0.05$, and one stock from ten will have these statistics with probability 0.354. A stock with these statistics has an expected single period log return of 0.0062, and $Perf = 12.3$ for 2000 investment periods. Any stock with a mean greater than 1.0073, or a standard deviation less than 0.05, will have a better performance than this.
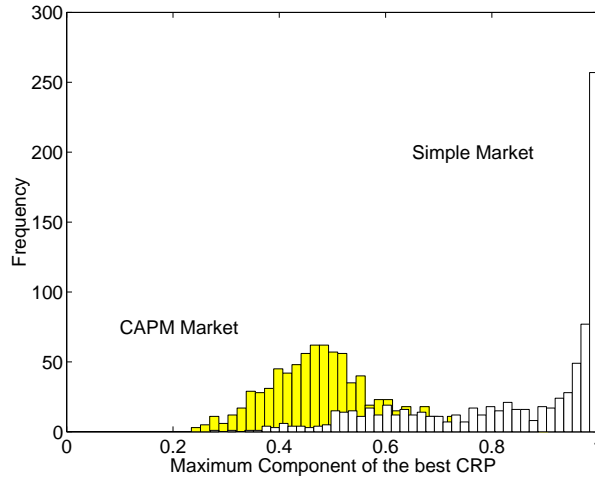
FIG. 8. The distribution over 835 trials of the value of the maximum component of the best offline constant rebalanced portfolio, plotted in the Simple market and the CAPM market.

that perform badly decreases and the investment in stocks that perform well increases. In this sense the portfolio selection problem in the Simple market is easy.

This analysis also explains why communication boosts the performance of the multiagent portfolio selection model. The agents' individual investment strategies have poor performance because the strongly-competitive χ^2 portfolio-selection rule is "too sophisticated" for the statistical realities of the market. The agents are slow to learn that the best CRP is extremal and at a corner of the the simplex of portfolio strategies with non-zero components that sum to one. Communication and strategy switching helps, especially in the early periods, because while the optimal portfolio (approximately a single stock) is a long way from the initial overall portfolio, it can be close to the random initial portfolio of one of the agents. The system of cooperative search takes advantage of this by adjusting the overall portfolio towards the portfolio of the single agent that is performing best.

In the next section we study the performance of our multiagent model for portfolio selection in a more realistic market that models correlations between stock prices, an approximation to the Capital Asset Pricing Model (CAPM) market. Although the model of independent search continues to perform well, the model of explicit cooperation with hint exchange and strategy switching has a negligible effect on performance in CAPM. The optimal portfolios tend to be more balanced in CAPM because of equilibrium price dynamics, and communication about different parts of the search space is less important.
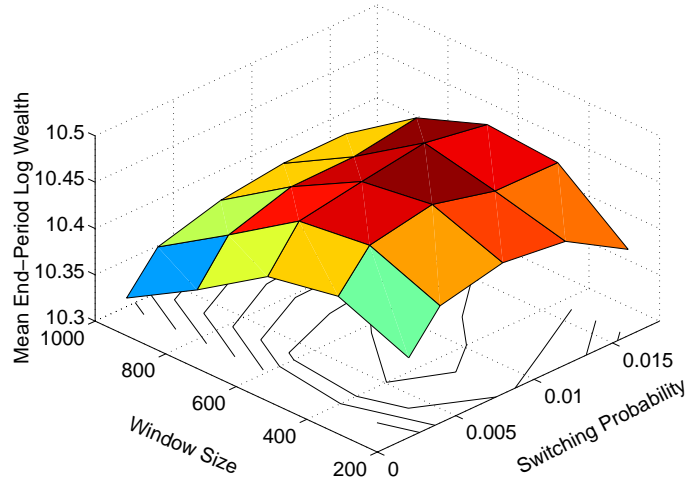
FIG. 9. Simple market. Performance of cooperative search with 100 agents as the switching probability, p, and the performance window size, τ, are varied, with learning-rate $\eta \sim U(0.1, 0.15)$. There is a peak at around $(p, \tau) = (0.008, 500)$.

5.4. The Choice of Model Parameters

Briefly, let us address the choice of model parameters and the robustness of our results to parameter selection.

Learning rate

The choice of learning rate, η, in the portfolio-selection rule, χ^2, is a tradeoff between accuracy and speed of convergence. With a small learning rate the portfolio moves slowly towards the optimal strategy with little sensitivity to period-to-period fluctuations in prices. With a large learning rate the portfolio is more sensitive to period-to-period fluctuations in stock prices and can fail to converge.

Figures 5 (a) and (b) illustrate the effect of learning rate within the independent search model, for a system of 100 agents. In (a) the agents have a small learning rate, $\eta_i \sim U(0.1, 0.15)$, and the adaptive agents achieve a final wealth that is an average of 4.3 times greater than the non-adaptive agents (line), while performing more than 80% worse than the non-adaptive agents in only 3.7% of the trials. In (b) the agents have a large learning rate, $\eta_i \sim U(0.9, 0.95)$, and the adaptive agents achieve a final wealth that is an average of 26 times greater than the non-adaptive agents, but perform more than 80% worse than the non-adaptive agents in 16.5% of the trials. The adaptive agents are either lucky and perform well or unlucky and perform badly.

The best learning rate depends on the number and volatility of stocks in the market, and also on the number of investment periods. The learning rate in our main experiments is optimized offline to maximize expected utility for agents with logarithmic utility functions over return-on-investment, and selected from $\eta_i \sim U(0.1, 0.15)$ for each agent.

Switching probability and Performance-window size

The optimal choice of the switching probability, p, and the performance-window size, τ, in the cooperative search model depends on factors such as the volatility of the market, the number of agents, and the number of investment periods T.

At one extreme, the performance of the cooperative multiagent system reduces to that of a single adaptive agent when $\tau/p \ll T$, because there is too much strategy switching and all agents converge to the portfolio-selection strategy of a single agent. At the other extreme, when $\tau/p \gg T$, the cooperative multiagent system reduces to a system of independent agents because the interval between switching is very large and the probability of switching is very low, and there is too little strategy switching.

Figure 9 plots the performance of cooperative search with 100 agents and different communication parameters. The optimal combination of parameters in this case is approximately $p = 0.008$ and $\tau = 500$. Notice that the performance drops off when τ is too large or p is too small because of too little strategy switching, and also when τ is too small or p is too large because of too much strategy switching.

In our main experiments we select the parameters offline for each problem size to maximize the benefit from cooperation. The optimal parameters are similar across all problems, with $p = 0.004$ and $\tau = 200$ typical, and a trend to larger switching probabilities as the number of agents increases.

6. PERFORMANCE IN AN EQUILIBRIUM MARKET MODEL

Finally, we tested the performance of our multiagent portfolio-selection models in a more realistic market, an approximation to the influential Capital Asset Pricing Model (CAPM) [53], an equilibrium model of stock price dynamics.

For our purposes, CAPM augments the geometric Brownian motion model of the Simple market with correlations between stock price dynamics. The system of agents remains small with respect to the total market, and we

treat prices as exogenous inputs, although the prices are now part of a larger equilibrium system.

Although independent search continues to perform well, explicit cooperation through hint exchange and strategy-switching is redundant in CAPM. Cooperative search performs no better than independent search. The difference is explained by the statistical properties of price dynamics in the Simple and CAPM markets.

6.1. The Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) explains a significant fraction of the price dynamics observed in real stock markets [8, 43]. Price changes across stocks in real markets are in fact highly (but not perfectly) correlated [11]. For example, many stocks in the same industrial sector will move in the same direction in a given day, and few will move against the overall trend of the market. It is this partial correlation which allows diversification through portfolio investment to reduce but not eliminate risk [53, 12].

CAPM predicts the equilibrium statistics of price dynamics in a market with mean-variance optimal investors with *homogeneous beliefs* about the future dynamics of stock prices [53].[7] A central result is that in equilibrium all stocks have risk-return statistics that fall on the *Security Market Line*, such that the expected excess return of a stock (return above the risk-free rate of return in the market) is proportional to the covariance of return with the market portfolio.

DEFINITION 6.1.   Security Market Line. In CAPM there is a linear relation between expected excess return and volatility with respect to the market across all stocks.

$$E[X_i] - \mathrm{R}_f = \beta_i \left(E[X_m] - \mathrm{R}_f \right)$$

where $E[X_i]$ is the expected return on stock i, $E[X_m]$ is the expected return on the market portfolio, and R_f is the available rate of risk-free return (zero in our model). The constant β_i, the "beta" of stock i, is a measure of the volatility of the stock's rate of return in relation to changes in the market rate of return, defined as:

$$\beta_i = \frac{\mathrm{Cov}[X_i, X_m]}{\mathrm{Var}[X_m]}$$

---

[7] The main additional assumptions are lending and borrowing at a risk-free rate of interest, and a competitive and costless market (no market friction).
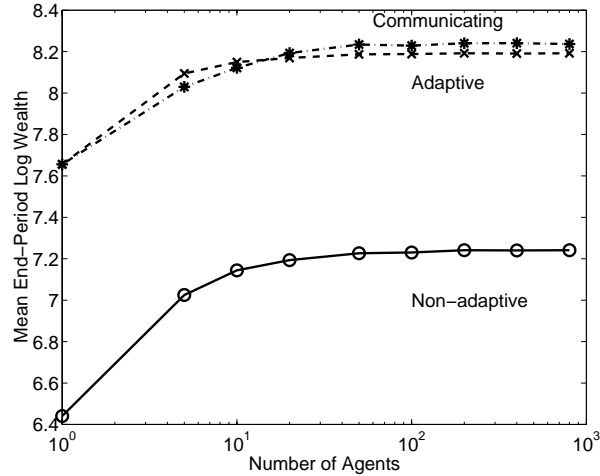
FIG. 10. CAPM market. The performance of the non-adaptive independent search (non-adaptive), independent search (adaptive), and cooperative search (communicating) models.

A stock with $\beta_i > 1$ has a high expected return and a high risk, while a stock with $\beta_i < 1$ has a low expected return and a low risk.

To generate stock prices offline with statistics that fall approximately onto the Security Market Line we follow Huang and Litzenberger [27] and augment the geometric Brownian motion model of stock prices with quantified correlations between the price relatives of stocks. The price relatives $\mathbf{x} = (x_1, \dots, x_N)$ are distributed according to a multivariate Normal distribution, $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with mean $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)$ and $N \times N$ covariance matrix, $\boldsymbol{\Sigma}$, which specifies the correlations between price relatives across stocks.

The mean μ_i for each stock is distributed according to the same uniform distribution as in the Simple market, $\mu_i \sim U(0.9995, 1.01)$. We set this and then perform random "generate-and-search" for a covariance matrix that generates a sequence of stock prices with approximate CAPM statistics. The standard deviation of each price relative is constrained to $\sigma_i \in [0.0, 0.2]$, the same as in the Simple market.

6.2. Results in the CAPM Market

As before we measured the performance of our multiagent portfolio selection models over 2000 investment periods, for systems with between 1 and 800 agents. Each agent was initialized from the same distribution of initial portfolios and learning rates as in the Simple market model. The

switching probability, p, and performance window size, τ, were selected offline to maximize expected utility for each number of agents, with switching probability $p = 0.004$ and window size $\tau = 200$ typical.

Figure 10 shows the performance of each multiagent model, averaged over 2000 trials. The best constant rebalanced portfolio strategy in this market yields an expected end-period log wealth, $Perf(\text{BCRP}) = 12.3$, while the market achieves a performance of $Perf(Market) = 5.31$.

Adaptive multiagent search continues to outperform both the non-adaptive multiagent search and a single adaptive agent: it remains useful to invest in diverse initial portfolios with individual agents that follow the χ^2 portfolio-update rule.

However, in this market cooperative search performs no better than independent search.[8] For example, consider Figure 6 (b), which compares the performance of cooperative and independent search with 400 agents. The system with communication and strategy switching often performs worse than without communication, and agents achieve approximately the same return on average.

The ability to select a portfolio close to the best offline CRP remains a good indicator of the performance of the agents, and explains the difference in performance of cooperative search in the Simple and CAPM markets. Figure 11 plots the average distance from the best offline CRP of the overall portfolio selected in each model. First, the initial portfolios are closer to the optimal offline portfolio than in the Simple market (0.5 compared to 0.81, see Figure 7). Furthermore, the market and non-adaptive agent models are ineffective at selecting the optimal offline portfolio in this market. Indeed, the multiagent portfolio-selection models outperform the market portfolio. Finally, although cooperative search initially boosts convergence towards the optimal portfolio, the independent search model closes the gap by the end of the investment.

Recall that in the Simple market the best CRP is often a single stock buy-and-hold strategy, which is a long way from the average initial overall portfolio of the system. In comparison, Figure 8 shows that the best CRP in CAPM typically invests around 50% in one stock, and it is rare for a single stock to receive more than 80% of the investment. This explains why the market portfolio, which can only achieve a performance that is competitive with the best single stock, does not perform well.

---

[8] There is only weak support for rejecting the null hypothesis that the non-communicating and communicating systems of agents have the same performance, with a minimum significance level of around 0.3 for systems with 50 or more agents.
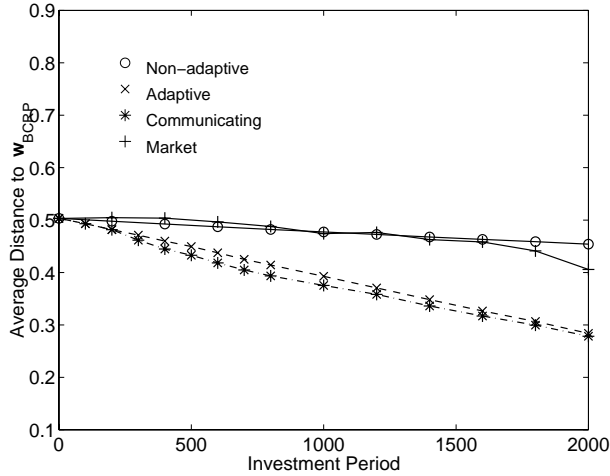
FIG. 11. CAPM market. The average distance in each period between the portfolio selected by each agent and the best offline constant rebalanced portfolio, in systems with 200 agents. The cooperative (communicating) and independent search (adaptive) models are most effective at selecting the optimal CRP.

Paradoxically, although the CAPM market has more statistical structure than the Simple market, the investment problem for an agent with a portfolio-selection strategy that is strongly competitive with the best CRP is easier. The correlation between return and volatility across stocks in CAPM makes the optimal portfolio closer to the initial portfolio of the agents, and there is "less to learn". All balanced portfolios perform quite well, and communication and strategy switching tends to have little effect on performance.

6.3.   Mean-Variance Analysis

Finally, it is interesting to compare the mean-variance efficiency [42, 53] of the portfolio-selection strategies in CAPM. A well-developed theory states that all adequately diversified portfolios, including the market portfolio, should have approximately the **same** *Sharpe Ratio* in a market which satisfies CAPM.

DEFINITION 6.2.   Sharpe Ratio. The Sharpe Ratio, r_p, for a portfolio is the ratio of excess expected return over the risk-free rate of return, to standard deviation in period-to-period return:

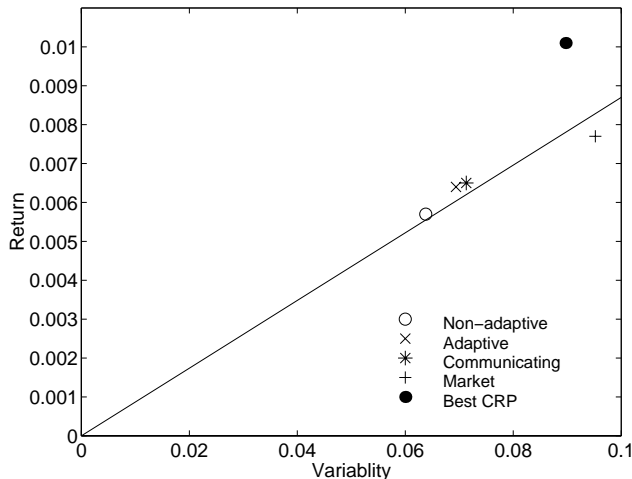$$r_p = \frac{E_p - \mathrm{R}_f}{\sigma_p}$$

FIG. 12. CAPM market. Expected per-period return versus standard deviation in per-period return (variability), averaged over 2000 trials. Each multiagent model has almost identical statistics for group sizes greater than one, so we plot one representative point for each model. The online portfolio-selection strategies all have statistics with approximately the same Sharpe ratio, falling on the Capital Market line.

where E_p is the expected per-period return from portfolio p, R_f is the risk-free rate of return (zero in our model), and σ_p is the standard deviation in per-period return from portfolio p.

Figure 12 plots the expected per-period return versus standard deviation in per-period return (variability) for each multiagent system. We plot only one point for each model because the return and variability varies only slightly for systems with more than one agent.

The analysis confirms that all online portfolio selection models, including the Market portfolio, have statistics with approximately the same Sharpe ratio. The best offline CRP, which uses complete information about all future stock prices when selecting a portfolio strategy, is the only exception. Of course, its performance is unattainable and not explained with standard portfolio theory.

However, there is no contradiction between Sharpe-ratio equivalence and our results. We have showed that adaptive multiagent search outperforms both non-adaptive multiagent search and the market portfolio, for agents with logarithmic utility functions over return-on-investment. Although the *ratio* of return to variability is not a sufficient statistic to predict the per-

TABLE 1.

Table I: Summary of experimental results. All multiagent models include 200 agents.

		Market Portfolio	Non-adaptive independent search	Adaptive independent search	Cooperative search	Optimal Offline Portfolio
Simple	$Perf^a$	14.3	8.6	10.1	10.5	16.0
	$Dist^b$	0.124	0.774	0.614	0.540	0
	$Corr^c$	0.953	0.208	0.347	0.428	1
Capm	$Perf$	5.3	7.25	8.2	8.25	12.3
	$Dist$	0.406	0.454	0.284	0.278	0
	$Corr$	0.389	0.587	0.941	0.965	1

[a] Average logarithmic return-on-investment.
[b] Average final distance from the optimal offline constant rebalanced portfolio.
[c] Correlation coefficient between return-on-investment and optimal offline return-on-investment.

formance of a portfolio strategy, the *difference* in return and variability across the models explains much of the variation in performance.

For example, the market portfolio has a per-period return distributed with $\mu = 0.0077$ and $\sigma = 0.0952$, while a system of 400 adaptive agents has a per-period return distributed with $\mu = 0.0064$ and $\sigma = 0.0694$. Assuming Normally distributed single-period returns with these statistics, this implies a distribution over end-period return-on-investment with expected utility of 6.4 and 8.0 for the market and adaptive multiagent search respectively. This is close to the empirically measured performances of 5.3 and 8.2 in each model.

7. DISCUSSION

In Table I we summarize our empirical results in the two market models. All results are presented for systems of 200 agents. The performance of each investment model is measured as the expected utility for an agent with a logarithmic utility function over return-on-investment. We include the performance of the optimal offline portfolio, reflecting the performance of an investor that knows all future stock price changes.

The distance measure compares the average final distance of the portfolio selected with each model with the optimal offline constant rebalanced portfolio. The correlation coefficient is computed between the return-on-investment of each model in each trial and the return of the optimal offline portfolio.

First, notice that performance is consistently higher in the Simple market than the CAPM market. The statistical analysis in Section 5 showed that there is often a single stock with high return and low risk in the Simple market, while all stocks in the CAPM market must lie on the Security Market Line with the same β (see Definition 6.1).

There is a strong relationship between *Dist*, a model's ability to select a portfolio close to the optimal portfolio, and its performance. This is where adaptive independent search helps in both markets, and cooperative search helps in the Simple market.[9]

The *correlation*, *Corr*, with the final return-on-investment from the optimal offline portfolio provides another indicator of performance. While the distance measurement compares the ability of a model to select a portfolio close to the optimal offline portfolio, correlation compares the ability of a model to achieve a high return-on-investment in the trials in which the offline portfolio achieves a high return-on-investment.

The independent search model achieves a remarkable correlation of 0.941 with the return from the optimal portfolio in the CAPM market, compared to 0.347 in the Simple market. In contrast, the market portfolio achieves a correlation of 0.389 in CAPM, compared to 0.953 in the Simple market. This is the sense in which the portfolio selection problem is easier for sophisticated portfolio-selection strategies in CAPM than in the Simple market, but harder for simple strategies. This also explains why communication boosts performance in the Simple market, but has a negligible effect in CAPM.

8. CONCLUSION

Multiagent portfolio selection, with agents that manage an equal share of initial investment and divide profits and losses at the end of investment, can improve the performance of a single agent portfolio-selection algorithm in simple market environments.

We compared a model of independent multiagent search, where the investment decisions of agents are completely autonomous, with a model of cooperative multiagent search, where the agents take account of the success of portfolio strategies followed by other agents. Experimental tests were performed in two markets: a Simple market with independent stock price

---

[9] The only exception is the performance of the market portfolio in CAPM, it performs worse than the non-adaptive agents despite finding portfolios closer to the optimal portfolio. However, the correlation between the return of the non-adaptive agents and the return of the optimal offline portfolio is higher than for the market portfolio, perhaps explaining this anomaly.

dynamics across stocks; and an approximation to the Capital Asset Pricing Model (CAPM), which is an equilibrium model of stock price dynamics that introduces correlations in price dynamics across stocks.

The independent multiagent search model performed better than a single agent in both markets. Furthermore, explicit cooperation, with communication and strategy switching between agents, improved performance in the Simple market. We showed that individual agents can follow strategies with very different performance (even though each agent's performance is worst-case optimal in the long-term). Communication and strategy-switching helps because it eliminates poor strategies.

However, communication provided only a small boost in performance in the CAPM market. The optimal portfolio in CAPM tends to be more balanced and close in search space to agents' initial strategies because stock dynamics are correlated. We showed that the statistical profile of a market influences that performance of cooperative search techniques. That said, explicit cooperation, perhaps with richer types of information exchange, remains a useful technique to make a multiagent portfolio selection models more *robust* to a variety of market environment.

In future work we would like to test our multiagent models on real stock market data, where stock price dynamics are a true reflection of an efficient marketplace. An interesting extension is to consider a system with agents that receive partial information about the prices in the market. This will serve to illustrate the role of collaborative computational agents and hint exchange when agents have limited or costly computation and are not able to process all information individually.

Efficient portfolio theory has also been applied to a technique for combining multiple heuristics to solve hard computational problems [29]. Future work may show that multiagent models provide a useful collaborative technique for the dynamic selection of an effective portfolio of heuristics, for example as a meta-rule for online problem solving in an environment with a sequence of stochastic problem instances [23].

APPENDIX A
Proofs

Proof (Theorem 2.1). Let $\{\mathbf{x}^T\} = \mathbf{x}^1, \ldots, \mathbf{x}^T$, denote a sequence of price relatives, independent and identically distributed across investment periods. The long-term optimal offline constant rebalanced portfolio, $\mathbf{w}^*$,

solves:

$$
\begin{aligned}
\mathbf{w}^* &= \arg\max_{\mathbf{w}} \lim_{T\to\infty} \left(\prod_{t=1}^{T} \mathbf{w} \cdot \mathbf{x}^t \right) \\
&= \arg\max_{\mathbf{w}} \lim_{T\to\infty} \left(\sum_{t=1}^{T} \log \mathbf{w} \cdot \mathbf{x}^t \right) \\
&= \arg\max_{\mathbf{w}} E_{\mathbf{X}} \log \mathbf{w} \cdot \mathbf{x} \qquad\qquad (*)\\
&= \arg\max_{\mathbf{w}} \sum_{t=1}^{T} E_{\mathbf{X}} \log \mathbf{w} \cdot \mathbf{x} \\
&= \arg\max_{\mathbf{w}} E_{\{\mathbf{X}^T\}} \sum_{t=1}^{T} \log \left(\mathbf{w} \cdot \mathbf{x}^t \right)
\end{aligned}
$$

∎

Proof (Theorem 2.2). We first prove Lemma A.1

LEMMA A.1. *Given two portfolios with the same expected single period return but different period-to-period variance, the portfolio with the smaller variance has larger expected log return in a market with non-negative, symmetric, independent and identically distributed price relatives.*

Proof. This follows from Jensen's inequality, which states that a concave function $f : \mathbb{R} \to \mathbb{R}$ is characterized by the condition that

$$
\int f(x)\mathrm{d}F \le f\left(\int x\mathrm{d}F \right)
$$

for any probability measure over x, $F : \mathbb{R} \to [0,1]$. Let us assume that the distribution F, that represents the distribution over price relatives, is non-negative. Substituting $f(x) = \log(x)$,

$$
E[\log(\mathbf{X})] \le \log\left(E[\mathbf{X}] \right)
$$

Now, since $E[\mathbf{X}] = \mu$, we have

$$
E[\log(\mathbf{X})] \le \log(\mu)
$$

for all x drawn from F. When we also assume a *symmetric* distribution F, then $E[\log(\mathbf{X})]$ is strictly monotonic in the standard deviation, σ, of the distribution (proof omitted). Given that: (1) $\lim_{\sigma\to 0} E[\log(\mathbf{X})] = \log(\mu)$;

(2) $E[\log(\mathbf{X})] \leq \log(\mu)$; (3) $E[\log(\mathbf{X})]$ is strictly monotonic in standard deviation, then $E[\log(\mathbf{X})]$ must be a strictly *decreasing* function in σ. ∎

Therefore any portfolio that maximizes the expected single period log return in a market with non-negative, independent and identically distributed price relatives, lies on the efficient frontier. Recall that the efficient frontier is the set of all portfolios that minimize standard deviation for some expected return (Definition 2.4). Hence, because the long-term best CRP, $\mathbf{w}^*$, also maximizes the expected single period log return (*), then $\mathbf{w}^*$ is on the efficient frontier. ∎

Proof (Theorem 4.1). Consider a group of M agents, each with an equal share of total initial wealth, assumed (without loss of generality) to be \$1. The final total wealth of the group of agents after T periods is given by

$$\frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \mathbf{w}_m^t \cdot \mathbf{x}^t$$

where $\mathbf{w}_m^t$ is the portfolio of agent m in period t, and $\mathbf{x}^t$ is price relatives in period t. To prove that the overall portfolio of the system of agents is strongly competitive (Definition 2.1), with $Perf_{comp}(\mathcal{R}) = \mathcal{R}^{1/T}$, we need to prove that:

$$\lim_{T \to \infty} \min_{\{\mathbf{x}^T\}} \left[\frac{\left(\frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \mathbf{w}_m^t \cdot \mathbf{x}^t \right)^{1/T}}{\left(\prod_{t=1}^{T} \mathbf{w}^* \cdot \mathbf{x}^t \right)^{1/T}} \right] = 1$$

where $\mathbf{w}^*$ is the best long-term offline portfolio, solving (1) as the number of investment periods $T \to \infty$. We require for all stock price sequences, $\{\mathbf{x}^T\} = \mathbf{x}^1, \ldots, \mathbf{x}^T$, that

$$\lim_{T \to \infty} \left[\frac{\left(\frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \mathbf{w}_m^t \cdot \mathbf{x}^t \right)^{1/T}}{\left(\prod_{t=1}^{T} \mathbf{w}^* \cdot \mathbf{x}^t \right)^{1/T}} \right] \geq 1 \qquad \text{(A.1)}$$

With out loss of generality, assume a particular sequence of stock prices, $\{\mathbf{x}^T\}$, and T investment periods. Let l denote the agent that achieves the least return-on-investment of all the agents:

$$\prod_{t=1}^{T} \mathbf{w}_l^t \cdot \mathbf{x}^t \leq \prod_{t=1}^{T} \mathbf{w}_m^t \cdot \mathbf{x}^t \qquad \forall m \neq l$$

Substituting the return-on-investment of agent l for every agent in (A.1) gives a strictly smaller performance ratio, and leads to the following inequality:

$$\lim_{T \to \infty} \left[\frac{\left(\frac{1}{M} \sum_{m=1}^{M} \prod_{t=1}^{T} \mathbf{w}_l^t \cdot \mathbf{x}^t \right)^{1/T}}{\left(\prod_{t=1}^{T} \mathbf{w}^* \cdot \mathbf{x}^t \right)^{1/T}} \right] \geq 1$$

Simplifying, since $\frac{1}{M} \sum_{m=1}^{M} \left(\prod_{t=1}^{T} \mathbf{w}_l^t \cdot \mathbf{x}^t \right) = \prod_{t=1}^{T} \mathbf{w}_l^t \cdot \mathbf{x}^t$, and because the portfolio strategy of agent l is strongly competitive, the inequality holds by construction as $T \to \infty$. ∎

APPENDIX B
Algorithms for Multiagent Portfolio Selection

Consider M agents, and let the variable m denote a single agent. We present simple object-oriented pseudo-code, with comments prefixed with //.

Independent Non-adaptive Multiagent Search

NON-ADAPTIVE() is the algorithm implemented by a non-adaptive system of agents. The agents are each assigned a share of initial wealth V_0, $m.$wealth $= V_0/M$, and a random portfolio, $m.\mathbf{w} = \text{rand}(N)$, where $\text{rand}(N)$ generates a random vector from the Dirichlet$(1/N, \dots, 1/N)$ distribution. At the start of a new period every agent rebalances its portfolio according to the current stock prices, $m.$trade($price$), to maintain a constant strategy. Then the price relatives for the period are generated, price_relatives(t), and the stock prices are updated. At the end of an investment period the agents observe the new prices. The algorithm terminates after T periods, when NON-ADAPTIVE() returns the total wealth over all agents.

ALGORITHM 1 (NON-ADAPTIVE()).

```
Global: T ;                          // number of investment periods
        N ;                                  // number of stocks
        M ;                                  // number of agents
        V_0 ;                                // initial wealth
   for each m ∈ M {
      m.wealth = V_0/M;                       // share initial wealth
```

```
        m.w = rand(N); }                              // assign random CRP
    price = (1, . . . , 1);                           // set initial prices
    t = 1;
    while (t ≤ T) {
        for each m ∈ M
            m.trade(price);
        (price, x) = price_relatives(t);
        for each m ∈ M
            m.wealth = m.wealth * m.w · x;
        t = t + 1; }
    return (∑_{m∈M} m.wealth); }
```

Independent Multiagent Search

ADAPTIVE(η_l, η_h) is the algorithm that is implemented by an adaptive, non-communicating system of agents. Agents have heterogeneous learning rates, uniformly distributed between η_l and η_h. The algorithm proceeds as with NON-ADAPTIVE() except that the agents update their portfolio strategies at the end of each investment period. The function m.update($\mathbf{x}$) returns the new portfolio of agent m using the χ^2 update rule, computed on the basis of the current portfolio $m.\mathbf{w}$, the current stock price relatives $\mathbf{x}$, and the agent's learning rate. The algorithm terminates after T periods, when ADAPTIVE(η_l, η_h) returns the total wealth over all agents.

ALGORITHM 2 (ADAPTIVE(η_l, η_h)).

```
Global: T ;
        N ;
        M ;
        V_0 ;
    for each m ∈ M {
        m.wealth = V_0/M;
        m.w = rand(N);
        m.η ∼ U(η_l, η_h); }                          // assign random learning rates
    price = (1, . . . , 1);
    t = 1;
    while (t ≤ T) {
        for each m ∈ M
            m.trade(price);
        (price, x) = price_relatives(t);
```

```
    for each m ∈ M {
        m.wealth = m.wealth * m.w · x;
        m.w = m.update(x); }                    // update portfolio
      t = t + 1; }
  return (∑_{m∈M} m.wealth); }
```

Cooperative Multiagent Search

COMMUNICATING$(\eta_l, \eta_h, p, \tau)$ is the algorithm that is implemented by a group of adaptive and communicating agents. The additional parameters, p and τ, define the switching probability and window size, common to every agent in the system.

The function $m.\text{average}(\mathcal{R}, \tau)$ returns the average per-period return for the past τ investment periods for agent m given return $\mathcal{R}$ in the current period, where $m.\text{count}$ is the number of periods since the agent last switched strategy; $m.\text{post}()$ posts the current portfolio of agent m and its recent performance to the blackboard; and $m.\text{switch}(p)$ either returns the current portfolio of agent m and increments $m.\text{count}$ or returns the best portfolio of the other agents (from the blackboard) and resets $m.\text{count}$ to zero.

The algorithm proceeds as for ADAPTIVE(η_l, η_h), except that at the end of every period each agent updates $m.\text{perf}$, its per-period return over the past τ investment periods. Each agent that has not switched strategy for at least τ investment periods posts its current strategy and recent performance to the blackboard, and switches to the best strategy in with probability p. The algorithm terminates after T periods, when COMMUNICATING$(\eta_l, \eta_h, p, \tau)$ returns the total wealth over all agents.

ALGORITHM 3 (COMMUNICATING$(\eta_l, \eta_h, p, \tau)$).

```
Global: T ;
        N ;
        M ;
        V_0 ;
  for each m ∈ M {
     m.wealth = V_0/M;
     m.w = rand(N);
     m.η ∼ U(η_l, η_h);
     m.count = 0; }                            // initialize count
  price = (1, ... , 1);
  t = 1;
```

```
while (t ≤ T) {
    for each m ∈ M
        m.trade(price);
    (price, x) = price_relatives(t);
    for each m ∈ M {
        m.wealth = m.wealth * m.w · x;
        m.perf = m.average(m.w · x, τ);          // update performance
        m.count = m.count + 1;                    // increment count
        if (m.count ≥ τ)
            m.post();                             // post performance
    }
    for each m ∈ M {
        if (m.count ≥ τ)
            (m.count, m.w) = m.switch(p);     // probabilistic switching
                                              // reset count if switch
        m.w = m.update(x);
    }
    t = t + 1; }
    return (∑_{m∈M} m.wealth); }
```

REFERENCES

1. David Aldous and Umesh Vazirani. "Go with the winners" algorithms. In *Proc. of the 35th Symp. on Found. of Comp. Sci.*, pages 492–501, 1994.

2. Paul H Algoet. Universal schemes for prediction, gambling and portfolio selection. *The Annals of Probability*, 20(2):901–941, 1992.

3. Paul H Algoet and Thomas M Cover. Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *The Annals of Probability*, 16(2):876–898, 1988.

4. W Brian Arthur, Steven N Durlauf, and David A Lane, editors. *The Economy as an Evolving Complex System II*. Addison-Wesley, 1997.

5. Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *36th Annual Symp. on Found. of Comput. Sci.*, pages 322–331, 1995.

6. Dimitri P Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall, 1987.

7. Sushil Bikchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100:992–1026, 1992.

8. Fischer Black, Michael C Jensen, and Myron S Scholes. The Capital Asset Pricing Model: Some empirical tests. In M C Jensen, editor, *Studies in the Theory of Capital Markets*. Praeger, NY, 1972.

9. Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35:193–205, 1999. Earlier in Proc. 10th Annual Conf. on Comput. Learning Theory (COLT'97).

10. Kenneth D Boese, Andrew B Kahng, and Sudhakar Muddu. A new adaptive multi-start technique for combinatorial global optimizations. *Operations Research Letters*, 16:101–113, 1994.

11. Karl Henrik Borch. *The Economics of Uncertainty*. Princeton University Press, 1968.

12. John Y Campbell, Andrew W Lo, and A Craig MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1997.

13. Scott H Clearwater, Bernardo A Huberman, and Tad Hogg. Cooperative solution of constraint satisfaction problems. *Science*, 254:1181–1183, 1991.

14. Thomas M Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

15. Thomas M Cover and David H Gluss. Empirical Bayes stock market portfolios. *Advances in Applied Mathematics*, 7:170–181, 1986.

16. Thomas M Cover and Erik Ordentlich. Universal portfolios with side information. *IEEE Transactions on Information Theory*, 42(2):348–363, mar 1996.

17. Avinash K Dixit and Robert S Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994.

18. Glenn Ellison and Drew Fudenberg. Rules of thumb for social learning. *Journal of Political Economy*, 101(4):612–643, 1993.

19. Joshua M Epstein and Robert Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press/Brookings, 1996.

20. Drew Fudenberg and David M Kreps. Learning mixed equilibria. *Games and Economic Behavior*, 5:320–367, 1993.

21. Piotr J Gmytrasiewicz and Edmund H Durfee. A rigorous, operational formalization of recursive modeling. In *Proc. 1st Int. Conf. on Multi-Agent Systems (ICMAS-95)*, pages 125–132. AAAI Press, CA., 1995.

22. I J Good. Twenty-seven principles of rationality. In V P Godambe and D A Sprott, editors, *Foundations of Statistical Inference*. Toronto: Holt, Rinehart and Winston, 1971.

23. J Gratch and S Chien. Adaptive problem-solving for large-scale scheduling problems: A case study. *Journal of Artificial Intelligence Research*, 4:365–396, 1996.

24. David P Helmbold, Robert E Schapire, Yoram Singer, and Manfred K Warmuth. A comparison of new and old algorithms for a mixture estimation problem. *Machine Learning*, 27(1):97–119, 1997. Earlier in Proc. 8th Annual Conf. on Comput. Learning Theory (COLT'95).

25. David P Helmbold, Robert E Schapire, Yoram Singer, and Manfred K Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8(4):155–177, 1998.

26. Tad Hogg and Colin P Williams. Solving the really hard problems with cooperative search. In *Proc. 11th Nat. Conf. on Artificial Intelligence (AAAI-93)*, pages 231–236. AAAI Press, CA, July 1993.

27. Chi-fu Huang and Robert H Litzenberger. *Foundations for Financial Economics*. North-Holland, 1988.

28. Bernardo A Huberman. The performance of cooperative processes. *Physica D*, 42:38–47, 1990.

29. Bernardo A Huberman, Rajan M Lukose, and Tad Hogg. An economics approach to hard computational problems. *Science*, 275:51–54, 1997.

30. Sandy Irani and Anna R Karlin. Online computation. In Dorit S Hochbaum, editor, *Approximation Algorithms for NP-Hard Problems*, chapter 13, pages 521–564. PWS Publishing, 1997.

31. D S Johnson, C R Aragon, L A McGeoch, and C Schevon. Optimization by simulated annealing: An experimental evaluation. *Operations Research*, 37:865–892, 1989.

32. Ehud Kalai and Ehud Lehrer. Rational learning leads to Nash equilibrium. *Econometrica*, 61(5):1019–1045, 1993.

33. Michihiro Kandori, George J Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993.

34. S Kauffman and S Levin. Toward a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128:11–45, 1987.

35. Kevin Knight. Are many reactive agents better than a few deliberative ones? In *Proc. 13th Int. Joint Conf. on Artificial Intelligence (IJCAI-93)*, pages 432–437. AAAI Press, CA, 1993.

36. W A Kornfeld. The use of parallelism to implement a heuristic search. In *Proc. 7th Int. Joint Conf. on Artificial Intelligence (IJCAI-81)*, pages 575–580. AAAI Press, CA, 1981.

37. Blake LeBaron, W Brian Arthur, John H Holland, Richard Palmer, and Philip Taylor. Asset pricing under endogenous expectations in an artificial stock market. In Arthur et al. [4].

38. Haim Levy. Risk and return: An experimental analysis. *International Economic Review*, 38:119–149, 1997.

39. Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994.

40. Michael Luby and Wolfgang Ertel. Optimal parallelization of Las Vegas algorithms. In *Proc. 11th Annual Symp. on Theoretical Aspects of Computer Science*, pages 463–474. Springer-Verlag, 1994.

41. Michael Luby, Alistair Sinclair, and David Zuckerman. Optimal speedup of Las Vegas algorithms. *Information Processing Letters*, 47(4):173–180, 1993.

42. Harry M Markowitz. *Portfolio Selection*. Wiley, New York, 1959.

43. Robert C Merton. *Continuous-Time Finance*. Blackwell, MA, 1997.

44. Paul Milgrom and John Roberts. Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3:82–100, 1991.

45. V Rao and V Kumar. On the efficiency of parallel backtracking. *IEEE Trans. on Parallel and Dist. Systems*, 4:427–437, 1993.

46. Michael Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9:185–202, 1974.

47. Ariel Rubinstein. *Modeling Bounded Rationality*. MIT Press, Cambridge, Mass., 1998.

48. Stuart Russell and Eric Wefald. Principles of metareasoning. *Artificial Intelligence*, 49:361–395, 1991.

49. Paul A Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *Review Econom. Statist.*, 51:239–246, 1969.

50. Tuomas W Sandholm and R H Crites. Multiagent reinforcement learning and the iterated Prisoner's Dilemma. *Biosystems Journal*, 37:147–166, 1995.

51. Andrea Schaerf, Yoav Shoham, and Moshe Tennenholtz. Adaptive load balancing: A study in multi-agent learning. *Journal of Artificial Intelligence Research*, 2:475–500, 1995.

52. Bart Selman, Hector Levesque, and David Mitchell. A new method for solving hard satisfiability problems. In *Proc. 10th Nat. Conf. on Artificial Intelligence (AAAI-92)*, pages 440–446. AAAI Press, CA, July 1992.

53. William F Sharpe. *Portfolio Theory and Capital Markets*. McGraw-Hill, 1970.

54. Herbert A Simon. From substantive to procedural rationality. In S J Latsis, editor, *Method and Appraisal in Economics*, pages 129–148. Cambridge University Press, 1976.

55. José M Vidal and Edmund H Durfee. Learning nested agent models in an information economy. *Journal of Experimental and Theoretical Artificial Intelligence*, 10:291–308, 1998.

56. Michael P Wellman and Junling Hu. Conjectural equilibrium in multiagent learning. *Machine Learning*, 33:179–200, 1998.

57. H Peyton Young. The evolution of conventions. *Econometrica*, 61(1):57–84, 1993.