Behavioral Finance and Market Anomalies: An Agent-Based Computational Economics Model

Doctoral Thesis

Dr. rer. nat.

Faculty of Human and Social Sciences

University of Vienna

Advisors:

Univ.-Prof. Dr. Oliver Vitouch

Univ.-Prof. Dr. Georg Dorffner

Ing. Mag. Michael Schönhart

Vienna, January 2005

“Traditionally, much of economic research has relied on the assumption of “homo oeconomicus" motivated by self-interest and capable of rational decision-making. Economics has also been widely considered a non-experimental science, relying on observation of real-world economies rather than controlled laboratory experiments. Nowadays, however, a growing body of research is devoted to modifying and testing basic economic assumptions; moreover, economic research relies increasingly on data collected in the lab rather than in the field. This research has its roots in two distinct, but currently converging, areas: the analysis of human judgment and decision-making by cognitive psychologists, and the empirical testing of predictions from economic theory by experimental economists.”

Press Release: The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel, October 9th 2002

“for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty”

The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel 2002, awarded to Daniel Kahneman, Princeton University, USA

ABSTRACT

Financial markets exhibit dynamics and behavior which are not completely explainable in the traditional neoclassical economic framework based on the assumption of rationally acting agents ("Homo Oeconomicus"). Even though there is strong empirical evidence that financial markets are highly efficient, the existence of these market "anomalies" is well accepted. In the last decades academic studies have revealed dozens of examples of repeated patterns of irrationality, inconsistency, and errors in judgment when human beings are required to reach decisions while faced with the condition of uncertainty. Behavioral finance incorporates this body of knowledge and argues that market anomalies can plausibly be understood using novel models in which agents are not fully rational.

In the first part of the thesis, seminal theoretical and experimental work on behavioral finance and market anomalies will be reviewed. Furthermore the underlying psychological mechanisms and empirical evidence of robust and systematic effects observed in experiments and over a wide area of financial markets data will be emphasized.

The main objective is the simulation of selected empirical effects based on the novel methodology of agent-based computational economics, which provides a framework to study an economic system in a controlled computational environment. Therefore an integrated markets model consisting of a financial market with trading agents and a consumer market with cognitively and socially bounded consumer agents will be introduced. The markets are coupled via learning production firm agents offering their products and shares. The consumer agents are embedded in a social structure based on "small-world network" principles. The integrated markets model will serve as a testbed, which allows the investigation of market dynamics under conditions which are too complex to be addressed analytically. The underlying behavioral, cognitive and social mechanisms will be explored.

CONTENTS

1 Introduction

Financial markets exhibit dynamics and behavior which are not completely explainable by traditional economic concepts. Despite strong evidence that financial markets are highly efficient, the existence of these “anomalies” is well accepted. In the last decades academic studies have revealed dozens of examples of repeated patterns of irrationality, inconsistency, and errors in judgment when human beings are required to reach decisions while faced with the condition of uncertainty (see for example Simon, 1955 and 1982; Kahneman and Tversky, 1974 and 1979; Statman, 1997; Dörner, 1997; Gigerenzer et al., 1999; Barberis and Thaler, 2003).

Behavioral finance argues that these financial phenomena can plausibly be understood using models in which agents are not fully rational. On the contrary neoclassical economic theory is based on the assumption of rationally acting agents (lovingly named “Homo Oeconomicus”). Used in this context rationality usually means two things. First, agents are able to update their beliefs correctly following the rules described by Bayes’ law (see section 2.2.1.2). Second, agents make choices which are consistent with Savage’s notion of Subjective Expected Utility (Savage, 1954). Savage’s work has once been described by Fishburn (1970) as “the most brilliant axiomatic theory of utility ever developed”, and by Kreps (1988) as “the crowning achievement of single-person decision theory". Since in reality probabilities are rarely objectively known, Savage (1954) developed a counterpart to expected utility theory (von Neumann and Morgenstern, 1944) known as Subjective Expected Utility. Under certain axioms of Subjective Expected Utility, preferences can be represented by the expectation of a utility function weighted by an individual’s subjective probability assessment. Nevertheless experimental work in the last few decades has been as unkind to Subjective Expected Utility as it was to expected utility (see section 2.2.3.4).

Moreover in traditional economics most models of asset pricing use the rational expectations equilibrium framework, which assumes consistent beliefs in addition to individual rationality (Sargent, 1993). This means that the subjective distribution an agent uses to forecast future realizations of unknown variables equals the distribution that those realizations are drawn from. Hence agents’ beliefs are correct if they are able to process new information correctly and if they are able to consider enough information in their decision-making process to find out the correct distribution for the unknown variables they are interested in. These traditional economic assumptions are appealingly simple, but after decades of research, it has become clear that basic facts about the aggregate stock market, the cross-section of average returns and individual trading behavior are not easily understood within this framework (Barberis and Thaler, 2003).

As an early critic on economic agents with unlimited information processing capabilities Herbert Simon (1955 and 1982) suggested the term “bounded rationality” to describe a more realistic approach to cover human problem solving competence. It has long been recognized that a source of judgment and decision biases is that cognitive resources such as time, memory, and attention are limited. Since human information processing capacity is not infinite, there is a need for imperfect decision making procedures, or heuristics that arrive at reasonably good decisions cheaply (see for example Simon, 1955; Tversky and Kahneman, 1974). The necessary abbreviation of decision processes can be called heuristic simplification (Daniel, Hirshleifer, and Teoh, 2002; see section 2.2.1). Indeed, the complexity of human behavior suggests that a choice model should explicitly capture uncertainty factors. Real economic agents are restricted at least in their cognitive (for example knowledge) and computational abilities (Mullainathan and Thaler, 2000).

Behavioral Finance is a “new” approach to financial markets. To overcome the difficulties faced by the traditional paradigm, behavioral finance argues that some financial phenomena can be better understood using models in which (some) agents are not fully rational. More specifically, it analyzes what happens when the assumptions that underlie individual rationality are relaxed. For example, if agents fail to update their beliefs correctly or agents apply Bayes’ law properly but make choices that are normatively questionable since they are incompatible with Subjective Expected Utility (Barberis and Thaler, 2003).

In the first part of the thesis, seminal theoretical and experimental work on behavioral finance and market anomalies will be reviewed. Furthermore the underlying psychological mechanisms and empirical evidence of robust and systematic effects observed in experiments and over a wide area of financial markets data will be emphasized.

The main objective is the simulation of selected empirical effects based on the novel methodology of agent-based computational economics, which provides a framework to study an economic system in a controlled computational environment. Therefore an integrated markets model consisting of a financial market with trading agents and a consumer market with cognitively and socially bounded consumer agents will be introduced. The markets are coupled via learning production firm agents offering their products and shares. The consumer agents are embedded in a social structure based on “small-world network” principles. The integrated markets model will serve as a testbed, which allows the investigation of market dynamics under conditions which are too complex to be addressed analytically. The underlying behavioral, cognitive and social mechanisms will be explored.

2 Behavioral Finance

Behavioral finance is a “new” approach to finance, where financial markets are investigated using models that are less restricted than those based on Von Neumann-Morgenstern (1944) expected utility theory and (no-)arbitrage assumptions[1]. Specifically, behavioral finance has two main building blocks:

- limits to arbitrage and
- cognitive psychology or psychology of decision making.

Limits to arbitrage refers to the effectiveness of arbitrage forces under different conditions. Cognitive biases refers to the huge psychological evidence documenting that people make systematic errors in the way they come to decisions under the condition of uncertainty. For example, they can be overconfident, they may put too much weight on recent experience, etc. Behavioral finance incorporates this body of knowledge rather than taking the approach that it should be ignored (Ritter, 2003).

Over the last decades, prominent researchers in both economics and psychology have criticized the view of neoclassical economics as psychologically unrealistic and proposed alternative assumptions. The underlying idea of this research is compelling in its simplicity: the more realistic the assumptions about economic actors, the better the economics. Thus economists should aim at making assumptions of humans as psychologically realistic as possible. The idea that economists should incorporate behavioral evidence from psychology that indicate systematic and important departures from the discipline’s habitual assumptions is so fundamental and manifest for economics that it will have long-term influence. Maybe a good analogy is the introduction of game theory in economics. While game theory is now a required core topic of every major U.S. Economics Department, it was said in 1985 by more than one respected and thoughtful economist that it will be a passing fad. Like game theory, psychological economics clearly expands the range of phenomena which economists can successfully study, and does so in

[1] Strictly defined, an arbitrage is an investment strategy that offers riskless profits at no cost (Barberis and Thaler, 2003).

what clearly is the spirit of economics. Like game theory, it is based not on a proposed paradigm shift in the basic approach of economics, but rather represents a natural broadening of the field. And finally, like game theory, behavioral finance is intended to be absorbed by economics, not to exist as an alternative approach (Rabin, 2002a).

Up to now one of the biggest successes of behavioral finance is a series of theoretical publications showing that in an economy, which includes interacting rational and irrational traders, irrationality can have a substantial and long-living impact on prices. One reason is that there are some psychological biases which virtually no one can escape. A second reason is that when traders are risk averse, prices reflect a weighted average of beliefs. Just as rational investors trade to arbitrage away mispricing, irrational investors trade to arbitrage away rational pricing. The presumption that rational beliefs will be victorious is based on the premise that wealth must flow from foolish to wise investors. But if investors are foolishly aggressive in their trading, they may earn higher rewards for bearing more risk (see for example DeLong et al., 1990b and 1991) or for exploiting information signals more aggressively (Hirshleifer and Luo, 2001). Thus irrational traders may gain from intimidating competing informed traders (Kyle and Wang, 1997). Indeed, one would expect wealth to flow from smart to dumb traders exactly when mispricing becomes more severe (Shleifer and Vishny, 1997; Xiong, 2000), which could contribute to self-feeding bubbles.

This stands in contrast to the neoclassical view that even if some agents in the economy are less than fully rational, rational agents will prevent them from influencing security prices for very long, through a process called arbitrage. The literature on "limits to arbitrage" (see for example Shleifer and Vishny, 1997), which overcomes this classic notion, forms one of the two buildings blocks of behavioral finance and will be reviewed in the next section. To specify the form of agents' irrationality behavioral economists typically turn to the extensive experimental evidence compiled by cognitive psychologists. Psychological research has discovered countless biases that arise when people form beliefs, and on people's preferences, or on how they make decisions, given their beliefs. Thus psychology represents the second building block of behavioral finance, which is reviewed in chapter 2.2 (Shleifer and Summers, 1990; Barberis and Thaler, 2003).

2.1 Limits to Arbitrage

Two economists are walking down the street and one spots a $100 bill lying on the ground. He turns to the other economist and says, "Look, a $100 bill!" The other economist looks at him in disbelief and answers, "If it were real, someone would have already picked it up."

(Bodie, Kane, and Marcus, 2002)

Limits to arbitrage refers to the conditions when arbitrage forces will be effective and when they will be not effective. Practitioners are familiar with the fact that misvaluations of financial assets are common, but nonetheless it is not easy to reliably make abnormal profits from these misvaluations. This is because misvaluations are mainly of two types: those that are recurrent and arbitrageable, and those that are non-repeating with a rather long-term time horizon. Trading strategies can reliably make money for the recurrent misvaluations. For example some hedge funds trade on these mispricings, thus keeping them from getting too big. Hence the market is pretty efficient for these assets. For the long-term, non-repeating mispricings, it is hard to identify the peaks and floors until they have passed. Even worse, if limited partners or other investors are supplying funds, withdrawals of capital after a losing streak may actually result in trading pressure that exacerbates the inefficiency. As Shleifer and Vishny (1997) state, the efforts of arbitrageurs to make money will make some markets more efficient, but they will not have any effect on other markets (Shleifer and Vishny, 1997; Ritter, 2003).

2.1.1 Market Efficiency

> *"If rational speculation makes markets efficient, then because the market is efficient, no profits can be made and all the rational speculators should leave, thus causing the market to revert to an inefficient state."*
>
> Milton Friedman

The traditional framework states that agents are rational and there are no market frictions. Thus a security's price equals its "fundamental value" which equals the discounted sum of expected future cash flows. This holds under the assumptions that investors correctly process all available information when forming expectations and that the discount rate is consistent with a normatively acceptable preference specification. The hypothesis that actual prices reflect fundamental values is known as the Efficient Markets Hypothesis (Barberis and Thaler, 2003).

The Efficient Markets Hypothesis represents a main building block of modern finance. It states that competition between investors seeking abnormal profits drives prices to their "correct" value. Under this hypothesis "prices are right" since they are set by agents who understand Bayes' law and have rational preferences. In an efficient market, there is "no free lunch". There does not exist an investment strategy that can earn excess risk-adjusted average returns. The Efficient Markets Hypothesis does not assume that all investors are rational, but it does assume that markets are rational. Furthermore it does not assume that markets can anticipate the future, but the hypothesis does assume that markets make unbiased forecasts of the future. In contrast to this traditional framework, behavioral finance comes to the conclusion that, under some circumstances, financial markets are informationally inefficient (Shleifer, 2000; Barberis and Thaler, 2003; Ritter, 2003).

2.1.2 Free Lunch or Wrong Prices?

One hypothesis of behavioral finance is that asset price deviations from their fundamental values are induced by the presence of traders who are not fully rational. The neoclassical economic approach assumes that rational traders are able and willing to quickly undo any dislocations caused by irrational traders (Friedman, 1953). For example let us assume that the market value of a share of VW equals its fundamental value. If a group of irrational traders becomes extremely pessimistic about VW's future prospects, they are able to push the price by selling. In such a case the Efficient Markets Hypothesis states that rational traders, anticipating an attractive investment opportunity, will buy the security at its undervalued price and at the same time, hedge their bet by shorting a "substitute" security, for example BMW, which exhibits similar cash flows in future states of the world. The resulting buying pressure on VW shares will finally bring their price back to the fundamental value.

Although Friedman's argument seems to be compelling at first sight, it has not survived careful theoretical examination. It is based on two assumptions. First, if there is a deviation from the fundamental value an attractive investment opportunity is created. Second, rational traders will immediately take advantage of the opportunity, thereby correcting this mispricing. Behavioral finance goes along with the second step in this argument: when attractive investment opportunities are discovered, it seems hard to believe that they are not quickly exploited. Nonetheless behavioral finance challenges the first step. Even when an asset is extremely mispriced, strategies designed to correct the mispricing can be both risky and costly, making them unattractive. As a result, the mispricing can remain. Strictly defined, an arbitrage is an investment strategy that offers riskless profits at no cost. Behavioral finance questions the belief that a mispriced asset immediately creates an opportunity for riskless profits. The strategies that Friedman's rational traders would adopt are not necessarily arbitrages, instead they often are very risky.

A consequence of this argumentation is that "prices are right" and "there is no free lunch" are not necessarily equivalent statements. In an efficient market, as described by the Efficient Markets Hypothesis, both statements are true. In an inefficient market, although there may be wrong prices there must not necessarily be a "free lunch" for any market participants. Thus if prices do differ from fundamental value, that does not mean that there are any excess risk-adjusted average returns even for the smartest investors (Barberis and Thaler, 2003).

Nevertheless many researchers still interpret the inability of professional money managers to beat the market as strong evidence of market efficiency (see for example

Rubinstein, 2001; Ross, 2001). Underlying this argument is the assumption that “no free lunch” implies “prices are right.” But if this condition is violated, the performance of money managers tells us little about whether prices reflect fundamental value or not.

2.1.3 The Risky Lunch

Trading strategies that are designed to correct existing mispricings can be both risky and costly. Thus the mispricings most likely will survive in the market. Some of the risks and costs that have been identified will be discussed in this section (according to Barberis and Thaler, 2003).

2.1.3.1 Fundamental Risk

For example let us again assume that the market value of a share of VW equals its fundamental value. If a group of irrational traders becomes extremely pessimistic about VW’s future prospects, they are able to push the price down by selling. A rational trader, anticipating an attractive investment opportunity, will buy the security at its undervalued price.

One obvious risk the arbitrageur faces if he buys VW’s stock below its fundamental value is that new announced bad news about VW can cause the stock to fall further, leading to losses. Arbitrageurs are well aware of this risk, since they usually short a substitute security such as BMW at the same time that they buy VW. The problem with substitute securities is that they are often highly imperfect, which makes it impossible to remove all the fundamental risk. Shorting BMW protects the arbitrageur from adverse news about the car industry as a whole, but still leaves him exposed to news that are specific to VW. Another problem that can occur even if a substitute security exists, is that the substitute itself may be mispriced. This can happen for example in situations where a whole industry is mispriced.

2.1.3.2 Noise Trader Risk

Noise trader risk is the risk that the mispricing being exploited by the arbitrageur worsens in the short run (De Long, Shleifer, Summers and Waldmann, 1990a; Shleifer and Vishny, 1997). For example under the assumption that BMW is a perfect substitute security for VW, the arbitrageur still faces the risk that pessimistic investors, who caused VW to be undervalued, become even more pessimistic, pushing VW's price even further.

Noise trader risk is important because it can drive arbitrageurs to liquidate their positions too early, bringing them potentially exorbitant losses. Shleifer and Vishny (1997) point out that there is "a separation of brains and capital" since most real-world arbitrageurs (for example professional portfolio managers) are not dealing with their own money, rather managing money for other people. This agency problem can have important implications. Investors usually evaluate portfolio managers based on their returns. If a mispricing that the manager is trying to exploit worsens in the short run, he is generating negative returns and investors may decide to withdraw their funds. Thus the arbitrageur will be forced to liquidate his position too early. Furthermore the fear of such premature liquidation can make him less aggressive in trying to exploit the mispricing.

2.1.3.3 Implementation Costs

Another barrier which may keep arbitrageurs from exploiting mispricings are transaction costs such as commissions, bid-ask spreads, and price impacts (see for example Chen, Stanzl, and Watanabe, 2001).

Short-sale constraints like the fee charged for borrowing a stock are rather small in general. For most stocks, they range between 10 and 15 basis points (D'Avolio, 2002) but they can be much larger and in some cases arbitrageurs are not able to find shares to borrow at any price. Furthermore there may be legal constraints. A large fraction of money managers, for example many pension fund and mutual fund managers, are simply not allowed to short-sell.

Another kind of implementation costs, horizon risk, is the risk that a mispricing takes so long to close that any profits are vanished due to accumulated transaction costs of the holding periods (for example lending fees). This applies even when the money manager is confident that no outside party can force him to liquidate too early. "Synchronization risk" is a specific type of horizon risk introduced by Abreu and Brunnermeier (2002). For example assume that the exploitation of a mispricing requires the participation of a sufficiently large number of different arbitrageurs. In that case and in the presence of per-period transaction costs, arbitrageurs may be unwilling to take advantage of the mispricing because they don't know how many other arbitrageurs have heard about the opportunity. Thus they are unsure about how long they will have to wait before prices revert to the correct fundamental values.

Finding and learning about a mispricing may count as an additional type of implementation costs, as well as the costs of the resources needed to exploit it (Merton, 1987). In economics it was thought for a long time that if noise traders are able to influence stock prices to any substantial degree, their actions would quickly be reflected in the form of predictability in returns. Shiller (1984) and Summers (1986) demonstrate that this argument is completely misleading. Moreover Shiller (1984) calls it "one of the most remarkable errors in the history of economic thought". They are able to show that even if noise traders' demand becomes so strong as to cause a large and persistent mispricing, it may create so little predictability in returns, that the anomaly still is almost undetectable.

2.1.3.4 Conditions of Risky Arbitrage

So far strong arguments were presented that real world arbitrage involves both costs and risks, which under some conditions will limit arbitrage and allow deviations from fundamental value to persist. To look at these conditions in more detail it is useful to consider two cases:

- the mispriced security does not have a close substitute
- the mispriced security does have a close substitute.

In the first case when the mispriced security does not have a close substitute the arbitrageur is exposed to fundamental risk. Thus arbitrage is limited due to the facts that

- arbitrageurs are risk averse, which ensures that the mispricing will not be wiped out by a single arbitrageur taking a large position in the mispriced security, and that

- the fundamental risk is systematic since it cannot be diversified by taking many such positions. This ensures that the mispricing will not be wiped out by a large number of investors each adding a small position in the mispriced security to their current assets.

The additional presence of noise trader risk or implementation costs will limit arbitrage further.

In the second case, although a perfect substitute does exist, arbitrage can still be limited. The existence of a close substitute relieves the arbitrageur from fundamental risk. Under the assumption that there are no implementation costs then only noise trader risk remains. De Long et al. (1990a) show that noise trader risk can be influential enough to limit arbitrage, when arbitrageurs are risk averse, have short horizons, and noise trader risk is systematic. Shleifer and Vishny (1997) argue that the possibility of an early, forced liquidation implies that many real world arbitrageurs effectively have short term horizons.

In the presence of implementation costs it may not even be necessary that noise trader risk is systematic. It is sufficient that discovering a mispricing is costly or the resources required to exploit it are expensive. Thus a large number of arbitrageurs or investors will not get involved in an attempt to correct the mispricing.

2.1.3.5 The Trend is Your Friend

Real-world arbitrageurs may prefer to trade in the same direction as noise traders rather than against them. De Long et al. (1990b) introduce an economy model with positive feedback traders, who only buy more of an asset at the current period if it performed well the last period. Thus this type of noise traders will push an asset's price above the correct fundamental value. Smart arbitrageurs will not sell or short-sell the asset. They rather buy it, since they know that the earlier price rise will attract more feedback traders in the next period. This leads to even higher prices and the smart arbitrageurs can exit their trading strategy with a profit.

So far it seems not to be easy for arbitrageurs like hedge funds to exploit market inefficiencies, since this may involve risks and costs. In the next section supporting empirical evidence for the theoretical arguments that real world arbitrage can be limited will be described.

2.1.4 Empirical Evidence

As lined out in the theoretical arguments, it is reasonable to state that arbitrage may be limited or ineffective since it involves additional costs and risks. Any empirical example of persistent mispricings also represents undoubtful evidence of limited arbitrage. If arbitrage mechanisms would be effective the mispricing would rapidly disappear. Furthermore it seems not to be easy to interpret many pricing phenomena simply as deviation from fundamental value since the latter requires a proper discounting model of future cash flows. This is what Fama (1970) named the “joint hypothesis problem”. Fama states that any test of a mispricing is therefore inevitably a joint test of the mispricing and of the model of discount rates, which makes it difficult to provide definitive evidence of the market inefficiency. Despite these problems researchers have discovered certain market anomalies, which represent persistent mispricings. These examples demonstrate the limitation of arbitrage due to the involved risks and costs (according to Shleifer, 2000; Barberis and Thaler, 2003; Ritter, 2003).

2.1.4.1 Twin Shares

In 1907, Royal Dutch of the Netherlands and Shell of the UK, at the time completely independent companies, agreed to merge their interests on a 60:40 basis while remaining separate entities. They also decided to pay dividends on the same basis. Thus shares of Royal Dutch (primarily traded in the USA and in the Netherlands) are a claim to 60 percent of the total cash flow of the two companies, while Shell (primarily traded in the UK) is a claim to the remaining 40 percent. If stock prices are equal to their fundamental values, the market value of Royal Dutch equity should thus always be 1.5 times the market capitalization of Shell equity. Figure 1 presents evidence that this is barely the case. Froot and Dabora (1999) analyzed the case and calculated a ratio of Royal Dutch equity value to Shell equity value relative to the efficient markets benchmark of 1.5 as shown in figure 1.

Thus they can provide strong evidence of a persistent market inefficiency. For example, The Financial Times Germany from the 25th of June 2004 states, besides reporting a recent balance scandal of Shell, possible reason for the persistent mispricing of the two companies; namely the different tax and stock corporation law in the Netherlands and the UK. Furthermore Froot and Dabora’s (1999) analysis revealed that the deviations are rather large. Royal Dutch is sometimes underpriced by 35 percent relative to parity, while sometimes overpriced up to 15 percent.

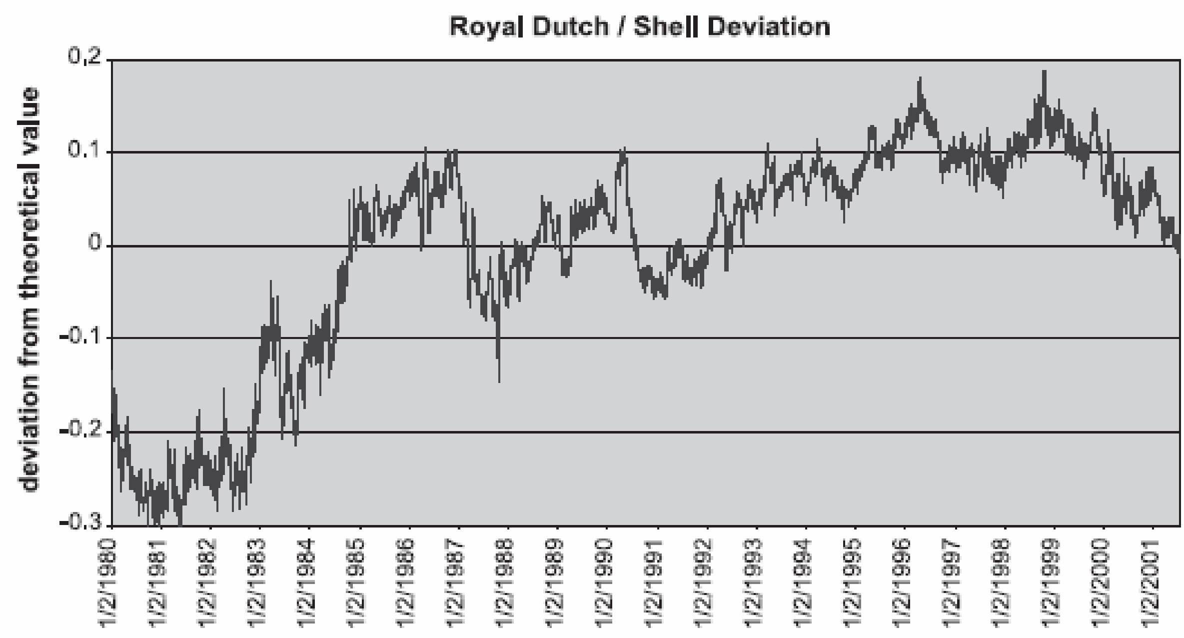


Figure 1: Deviations from Royal Dutch/Shell parity from January 1980 to December 2001, as computed by Froot and Dabora (1999) and updated by Ken Froot (from Ritter, 2003). The y-axis shows the price ratio of Royal Dutch equity value to Shell equity value relative to the theoretical efficient markets benchmark of 1.5.

The time series data in figure 1 ends in December 2001, with a price ratio close to the fundamental value of the underlying stocks. Despite that nice picture in July 2002 Standard and Poor’s announced that Royal Dutch would be dropped from the S&P 500 index because they were deleting non-American companies. Standard and Poor’s 500 is one of the most commonly used benchmarks of the overall stock market. It is a market-value weighted index, with each stock's weight in the index proportionate to its market value.

Furthermore it is an index consisting of 500 stocks chosen for market size, liquidity, and industry group representation. Royal Dutch dropped by 17 percent in the week of the S&P 500 index deletion announcement, although there was no evidence that the present value of dividends changed.

This mispricing is at the same time evidence of limited arbitrage. If an arbitrageur wanted to take advantage of this phenomenon he would buy the relatively undervalued share and short the other. Which risks is the arbitrageur implicitly bearing when applying this rather simple strategy? Given that a Royal Dutch share is a good substitute for the shares of Shell, fundamental risk is satisfactorily hedged. New information about fundamentals is supposed to affect the two shares equally. Since the shares are listed and traded in liquid and deep markets there are no major implementation costs. Thus shorting the shares of either company implies no additional risks.

The main risk that resides is noise trader risk. The opinion of market participants which cause Royal Dutch shares to be undervalued (overvalued) relative to Shell shares could also have the effect that the shares of Royal Dutch become even more undervalued (overvalued) in the short term. Figure 1 shows that this risk is very real. An arbitrageur who bought a ten percent undervalued Royal Dutch share in March 1983 would have observed that the stock price declined still further over the next six months. When a mispriced asset has a (nearly) perfect substitute, arbitrage can still be limited if

- arbitrageurs are risk averse and have short time horizons
- the noise trader risk is systematic or the arbitrage requires special skills or there are costs to learn about such opportunities.

It seems to be very plausible that for the Royal Dutch/Shell twin shares both arguments were true. Furthermore this provides a possible explanation why the mispricing persisted for so long. While prices in this case are obviously not right there are no easy profits for the taking or in other words there is no “free lunch”.

How well does this prediction work in practice? Several hedge funds, like Long-Term Capital Management, did try to exploit this anomaly with the mentioned arbitrage strategy. Finance theory has the clear prediction that whenever the Royal Dutch/Shell stock prices are not in a 60:40 ratio, there is an arbitrage profit opportunity. For the last 22 years, from 1980 to 2001, figure 1 demonstrates that there have been huge deviations from the theoretical relation. Furthermore both are large and trackable companies. Until July 2002

Royal Dutch was listed in the S&P 500 and Shell is listed in the FTSE 100 index. FTSE is a company that specializes in index calculation. Although not part of a stock exchange, co-owners include the London Stock Exchange and the Financial Times. The FTSE is similar to Standard and Poor's in the United States. They are best known for the FTSE 100, an index of blue chip stocks on the London Stock Exchange.

The before mentioned giant hedge fund, Long Term Capital Management (LTCM), was founded in 1993 by John Meriwether, former head of fixed income trading at Salomon Brothers. Even when forced to leave Salomon in 1991, in the wake of the firm's treasury auction rigging scandal, Meriwether continued to command huge loyalty from his former team of highly analytical relative-value fixed income traders. Teamed up with a handful of these traders, two Nobel laureates, Robert Merton and Myron Scholes, and former regulator David Mullins, Meriwether and LTCM had more credibility than the average broker or dealer on Wall Street (Shirreff, 1999). The fund was amazingly successful in the first few years and traded also on the Royal Dutch/Shell mispricing. In 1998, LTCM shorted the expensive stock and bought the cheaper one. But they lost money when prices diverged further from their theoretical values during the third quarter of 1998. To meet liquidity needs, LTCM and other hedge funds were forced to sell out their positions, and this selling pressure made markets more inefficient, rather than more efficient. So the forces of arbitrage failed. Moreover they had one bad quarter in which they lost four billion US dollars, wiping out their whole equity capital which forced the firm to file for chapter 11. Nevertheless they were right in the long run. LTCM mainly traded in fixed income and derivative markets. But one of the ways that they lost money was on the Royal Dutch/Shell equity arbitrage trade.

2.1.4.2 Index Changes

The Standard and Poor's 500 index is a portfolio of five hundred stocks representative for the leading industries of the U.S. economy. It is known from historical data that the S&P 500 is a good proxy for the U.S. market development and therefore commonly used as a benchmark for money managers. Furthermore it is considered as an investable index since individuals or institutions can easily invest their capital in the stocks of the index. From the current total market value of twelve trillion US$, approximately one trillion is indexed directly or indirectly to the S&P 500. Thus changes to this index are widely recognized. Not all of the companies in the S&P 500 are large. In 2002 only 340 firms in the index were also in the top 500 measured by market capitalization. The median of the market value of S&P 500 companies is approximately US$ 8 billion (Singal, 2004).

The S&P 500 index modifications are always initiated by deletions of companies due to major restructurings, like a merger, spin-off or bankruptcy. Since the S&P 500 contains no non-US companies a deletion can also happen if a company which is already member of the S&P 500 is acquired by a foreign corporation. Then it is replaced by another firm. Furthermore companies may be deleted by Standard and Poor's if they no longer represent the economy. This can happen either if their industry is no longer representative of the U.S. economy or because the firm is no longer representative of the industry. Usually the number of companies in the S&P 500 index is maintained at five hundred. Thus additions to the index are typically announced at the same time as deletions. There are four general criteria developed by Standard and Poor's serving as prerequisite for a company to be selected as an index inclusion candidate (according to Singal, 2004):

- The firm must have sufficient liquidity.
- The form of ownership must not be concentrated in a single or few entities.
- The company must be profitable.
- The firm must be a leader in an important U.S. industry.

There is no explicit market capitalization mentioned as threshold for a firm to be added. In 2002 the market capitalization of added firms was at least US$ four billion. Since many firms meet the criteria mentioned Standard and Poor's can make changes to the index subjectively. Nevertheless big investment banks and other market observers regularly try to predict the expected changes to the index more or less successfully. For example at the end of February 2002 Lehman Brothers identified nineteen companies as candidates for index addition and ten for deletion. Six months later exactly four of the nineteen firms had been added and two of the ten firms had been deleted. Since the prediction of the index changes is difficult the focus here is on market effects after the announcement of changes by Standard and Poor's. Usually changes to the S&P 500 index are announced after the market closes, while they take place at the close of the announced date. The time lag between the announcement and the effective date varies from one day to one month with infrequent lags of up to three months. Typically the effective date is declared at the initial announcement, sometimes a few days after.

There are substantial reasons for a company to be affected and why a change to the S&P 500 index is a significant event (according to Singal, 2004):

- An addition to the index indicates an additional demand for the stock of approximately 1/12 or ~8 percent of the outstanding shares since the value of assets that track the S&P 500 index is about US$ 1 trillion, while the total market value of all U.S. stocks is about US$ 12 trillion.
- Publicity stimulated by the addition will cause more investors to learn more about the company and eventually inspire them to trade.
- More analysts will follow the newly added firm because of the increased interest of investors.
- There will be increased trading in the firm, making its stock more liquid.

The effect of index changes on the stock price can be best described by viewing different historical periods. Until August 1976 stock indexing was not considered important or popular and no *public* announcements were made. Therefore nothing abnormal happened to the stock price either on the first trading day after an announcement or after that. In September 1976 Standard and Poor's began to officially announce index changes to interested investors, the media, and particularly mutual fund managers. In this period changes were announced after market close on Wednesdays, while the change to the index became effective the next morning after opening. Because index fund managers try to minimize the tracking error, which equals the difference between the return of the fund and the return of the index, of their funds they must buy the stock at the time of its addition to the index. During this second period (1976 to 1989) the stock price went up abnormally immediately after the announcement. Harris and Gurel (1986) and Shleifer (1986) document this remarkable market anomaly in their early studies. When a stock is added to the S&P 500 index its price increases dramatically by an average of 3.5 percent and much of this jump is permanent. Moreover the stock's price behavior tracks the S&P 500 index's return for the next three calendar months. In general for the first two periods there is no empirical evidence that deletions of stocks from the S&P 500 index have an impact on price.

Since the market for index funds grew impressively the buy orders from these funds at the opening of the market increased order imbalances and volatility. In October 1989 Standard and Poor's started to preannounce the intended S&P 500 index changes to help prevent order imbalances. For this period, from 1989 to 2000, index additions showed big abnormal effects. In general the excess return following the announcement date was 5.3 percent on average. Furthermore the return increased to 8.4 percent on the effective date, falling back to the announcement date return about twenty trading days later. A remarkable example was the addition of Yahoo to the index since its shares jumped by 24 percent in a single day. This again was clear evidence of a mispricing. The share price changed even though its fundamental value did not. Standard and Poor's pointed out that they basically tried to make their index representative of the U.S. economy and not to suggest any information about the state or riskiness of a company's future cash flows (Barberis and Thaler, 2003).

In the period from 1989 to 2000 deleted companies decreased by 5.4 percent in their stock price on the day of the announcement. Then they fell another 4.9 percent by the effective date resulting in a total loss of 10.3 percent. Followed by a rebound back to a net loss of 3.3 percent in the next twenty days they finally gained another 2.6 percent sixty days after the effective date of the index deletion. Thus the net effect on stock price of deletions is almost zero on average. Figure 2 shows an example.

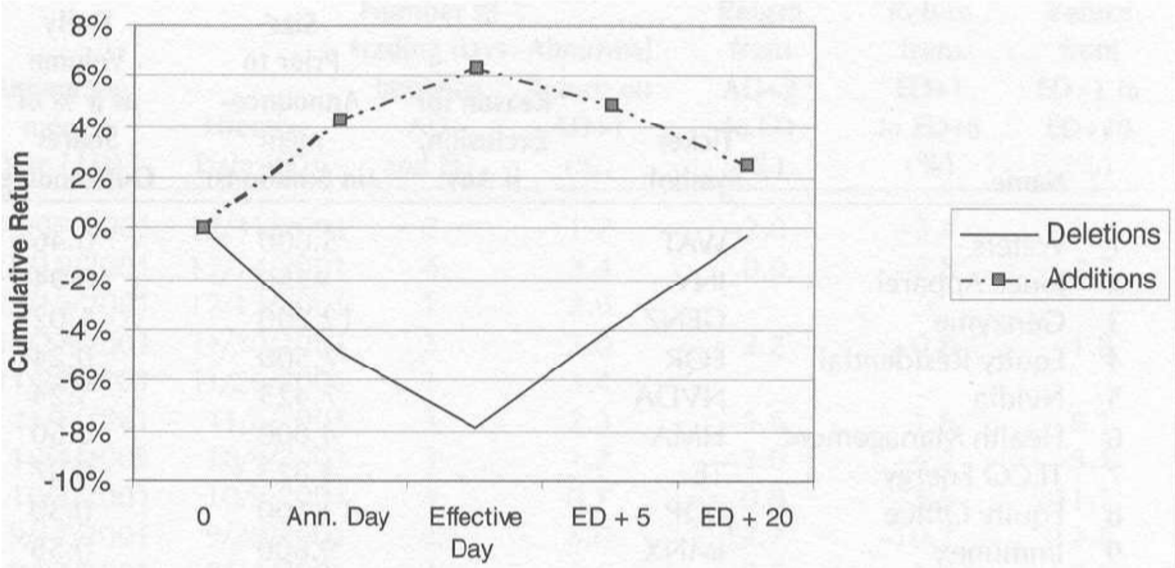


Figure 2: Changes to the S&P 500 Index: the cumulative return from the day of the index change announcement to the effective date and until twenty days after the effective date is shown for additions and deletions from the S&P 500 index during 2001 (from Singal, 2004).

Some researchers argue that the price increase might be rationally explained through information or liquidity effects. Nevertheless recent research like Kaul, Mehrotra and Morck (2000) substantiate the notion of a market anomaly due to irrational mispricings. They document the case of the TS300 Canadian equities index which in 1996 adjusted the weights of some of its stocks to meet a mild regulatory requirement. Despite that harmless adjustment the reweighing was accompanied by significant price effects. Information and liquidity explanations for the price jumps are extremely implausible in this case since the affected stocks were already in the index at the time of the event.

The effect of index inclusions as deviation from fundamental value represents evidence of limited arbitrage. Thus there are risks involved when an arbitrageur tries to exploit the anomaly and so the mispricing tends to persist. Arbitrageurs would short the included security and take a long position on a good substitute security. There is substantial fundamental risk involved because individual stocks rarely have good substitutes. Furthermore an investor who trades with this strategy bears significant noise trader risk. Thus the price may increase even further in the short run. For example Yahoo’s share price took off from US$ 115 prior to its S&P 500 index inclusion announcement to US$ 210 just one month later. Additional support for the limited arbitrage notion of S&P

500 index inclusions is provided by Wurgler and Zhuravskaya (2002). They state that the jump after a stock’s index inclusion should be particularly large for stocks with the worst substitute securities since for those stocks the arbitrage would be riskiest. To prove their hypothesis they construct the best possible substitute portfolio for each included stock. Their results reveal strong support for their hypothesis. Furthermore their analysis shows how hard it is to discover a good substitute security for an individual stock. Moreover Wurgler et al. (2000) demonstrate that for most regressions of included stock returns on the returns of their best substitute securities the regression parameter goodness of fit (R^2) is below 25 percent.

But there is also evidence that does not support the imperfect-substitutes explanation. First there is no relation between the level of indexing and the price impact for the post 1976 periods. If the hypothesis of imperfect substitutes holds then the greater the demand shock the greater the price impact should be. There is no empirical evidence for that to happen. Second the imperfect-substitutes explanation should affect both index additions and deletions. But firms deleted from the S&P 500 index do not have a permanent price impact. Thus this explanation does not seem to be the only answer (Singal, 2004).

Investor awareness or investor recognition of a company through S&P 500 index addition can improve its access to capital markets and its operating performance due to increased monitoring by investors. Thus the additional capital and the improved efficiency will allow the company to grow at a higher rate than that prior to the inclusion in the index. For index deletions investors are not able to become “unaware” of the stock. But they may reduce their holdings. Thus the investor recognition hypothesis supports an asymmetric effect since additions have a positive price impact and deletions have a muted impact on prices. Empirical evidence is consistent with this prediction. Furthermore investor recognition can explain that there was no price impact on stocks added or deleted to the S&P 500 index before 1976. Since there were no public announcements of index changes investors could not become aware of the stock and therefore there was no price impact. The same effect happens on prominently featured stocks in the media. They tend to become more widely held which results in an increase of the stock price. Thus investor recognition is an important factor for security pricing. However, it is popularly believed that the explanation of imperfect substitutes is the more appropriate one. But the investor recognition hypothesis is the one most consistent with the empirical evidence (Singal, 2004).

2.1.4.3 Initial Public Offerings

> *"Santa Clara, Calif. — March 2, 2000 — 3Com Corporation (Nasdaq: COMS). Following today's successful initial public offering of stock in Palm, Inc., 3Com reconfirmed its intention to complete a spin-off of the remaining shares of Palm owned by 3Com in approximately six months. 3Com owns approximately 532 million shares, or approximately 95%, of Palm. There are currently approximately 349 million shares of 3Com common stock outstanding. Final approval of such distribution will be made by 3Com's board of directors based on the number of shares outstanding of each company at the time of the distribution. However, if such ratio were to be calculated based on today's outstanding shares, 3Com shareholders holding 3Com stock as of the distribution date would be eligible to receive approximately 1.5 shares of Palm for each share of 3Com. The Palm initial public offering will close on Tuesday, March 7."*
>
> 3Com press release, March 2nd, 2000

On March the 2nd, 2000, 3Com Corporation successfully sold five percent of its entirely owned subsidiary Palm Inc. in an initial public offering (IPO). A shareholder of 3Com Corporation indirectly owned 1.5 shares of Palm Inc. after the IPO. Furthermore 3Com Corporation also announced its intention to spin off the remaining 95 percent equity of Palm Inc. within the next six months. Then each 3Com shareholder would receive 1.5 shares of Palm.

On the first day after the IPO Palm shares had a value of US$ 95 at the close of trading. Considering the ratio of 1.5 this would result in a lower bound on the value of 3Com of approximately US$ 143 per share. In fact 3Com's price was US$ 81 at that time. Thus this implies a market valuation of 3Com Corporation's substantial businesses outside of Palm of about US$ -60 per share. Those conditions definitely represent a severe mispricing of 3Com Corporation, which persisted for several weeks. An arbitrageur would buy one share of 3Com, while shorting 1.5 shares of Palm. Then he would wait for the spin-off, when he will earn certain profits at no cost. This strategy involves no fundamental risk and no noise trader risk (Barberis and Thaler, 2003).

The results of Lamont and Thaler (2003), who analyzed this case in detail, suggest that implementation costs were the main factor for the limitations of arbitrage and therefore the persistence of the 3Com mispricing. Investors who tried to borrow Palm Inc.'s shares to take a short position were either told by their broker that no shares were available or were quoted a very high borrowing price. These shorting barriers were not legal but they arose endogenously in the market. The demand for shorting Palm was so huge that the suppliers of Palm shorts were unable to meet it. Arbitrage was therefore limited and the mispricing persisted.

Stephen A. Ross, a serious critic of behavioral finance, confirmed the illiquidity of Palm shorts and other market anomalies from his own experience as a fund manager at his public lecture "A Neoclassical View of Behavioral Finance and the Closed-End Fund Puzzle-Implications for Asset Management" at Bank Gutmann, Vienna, Austria at January the 21st, 2004. Ross is also a principal of Roll and Ross Asset Management Corporation, which employs technology that Ross helped develop to manage over \$3 billion in investments worldwide.

Mitchell, Pulvino and Stafford (2002) examine the barriers to arbitrage in 82 situations between 1985 and 2000, where the market value of a company is less than its ownership stake in a publicly traded subsidiary. These situations suggest clear arbitrage opportunities and provide an ideal setting in which Mitchell et al. are able to analyze the risks and market frictions that prevent arbitrageurs from immediately forcing prices to fundamental values. As a result they find that there are costs that limit arbitrage in equity markets, which keep market forces from maintaining prices at their fundamental values. Mitchell et al. report that for 30 percent of the sample, the link between the parent and its subsidiary is severed before the relative value discrepancy is corrected. Furthermore, because of forced liquidation to satisfy capital requirements, they estimate that the returns to a particular arbitrageur would be 50 percent larger if the path to convergence was smooth rather than observed. Uncertainty about the distribution of returns and characteristics of the risks appear to be an important limit to arbitrage.

Ofek and Richardson (2003) explore a model based on agents with heterogenous beliefs facing short sales restrictions to explain the rise, persistence, and fall of internet stock prices. If a group of investors enters the market or becomes very optimistic then stock prices can rise quite dramatically. Pessimistic investors are required to short these „overvalued“ stocks but are prevented from doing so since they were overrun by the size and volume of optimistic trading. While this can explain any type of inflated stock price level in the context of limited arbitrage, it seems especially suited to stocks that are subject to short sale constraints and heterogeneous investors. Ofek and Richardson investigate this theory by looking at the behavior of internet stock prices during the extraordinary asset pricing period from January 1998 to February 2000. They provide three important findings:

- Using evidence on short sales, rebate rates, and option pairs, they document substantial short sale restrictions for internet stocks.
- With data on internet holdings and block trades around IPO-related events with shifting investor clientele, they are able to show a link between heterogeneity and price effects for internet stocks.
- They provide a detailed look at the impact lockup expirations have on internet stock prices. By arguing that lockup expirations are equivalent to loosening the short sale constraint, Ofek and Richardson report average, long-run excess returns as low as -34 percent for internet stocks post-lockup. Moreover, the long-run impact of the lockup expiration is related to gradual insider selling throughout the period.

These examples are not isolated cases with little relevance. The Royal Dutch/Shell mispricing demonstrates that in situations where arbitrageurs face only one type of risk (noise trader risk) securities can become mispriced by almost up to 35 percent. This suggests that if a typical stock trading on the New York Stock Exchange or NASDAQ becomes subject to investor sentiment, the mispricing could even be a magnitude larger. Arbitrageurs would then face not only noise trader risk in trying to correct the mispricing, but fundamental risk as well, not to mention implementation costs.

2.1.5 Discussion

The Efficient Markets Hypothesis has been the central proposition of finance in the last three decades (Shleifer, 2000). According to the Efficient Markets Hypothesis competition between investors seeking abnormal profits drives asset prices to their correct fundamental value. While the hypothesis does not assume that all investors are rational, it assumes that markets are rational. Furthermore the Efficient Markets Hypothesis does not assume that markets can foresee the future, but it does assume that markets make unbiased forecasts of the future (Ritter, 2003).

In his now classic paper Fama (1970) defined an efficient financial market as one in which asset prices always perfectly reflect the available information. Furthermore the Efficient Markets Hypothesis states that real-world financial markets are actually efficient according to this definition. The most radical implication of the Efficient Markets Hypothesis is that it rejects the possibility of trading strategies based solely on currently available information and has expected profits or returns in excess of equilibrium expected profits or returns (Fama, 1970). Thus an investor, whether an individual, a pension fund, or a mutal fund, is not able to consistently beat the market. Moreover all resources that such investors dedicate to analyzing, picking, and trading securities are wasted. The most excellent strategy the Efficient Markets Hypothesis supports, is a passively held market portfolio (no active money management at all), since the market truly knows best. In the first decade after its development in the 1960s, the Efficient Markets Hypothesis turned into an enormous theoretical and empirical success. In 1978, Michael Jensen, one of the creators of the Efficient Markets Hypothesis, stated that “there is no other proposition in economics which has more solid empirical evidence supporting it than the Efficient Markets Hypotheses” (Jensen, 1978, p. 95).

Shortly after this declaration by Jensen, the Efficient Markets Hypothesis was challenged on both its theoretical and empirical foundations. First it is difficult to claim that people in general, and investors in particular, are fully rational. Or in the words of Fischer Black (1986) they rather trade on noise than on information. But this seems only to be the tip of the iceberg since investors’ deviations from the predictions of economic rationality are highly pervasive and systematic. Kahneman and Riepe (1998) summarized that people deviate from the standard decision making model in a number of fundamental areas: attitude toward risk, non-Bayesian expectation formation, and sensitivity of decision making to the framing of problems. In the next sections theory and empirical evidence of the human inherent psychological biases will be reviewed.

Behavioral finance assumes that financial markets can be informationally inefficient under certain conditions. But not all mispricings are caused by psychological biases since some are due to temporary supply and demand imbalances. For example the “tyranny of indexing” can trigger demand shifts that are unrelated to the future cash flows of the particular company. In December 1999 Yahoo was added to the S&P 500 index. Therefore index fund managers had to buy the stock even though it had a limited public float. This extra demand led to a price increase by over 50 percent in one week and to over 100 percent in the following month. After eighteen months the stock price was down by over 90 percent from where it was shortly after being added to the S&P 500 index (Ritter, 2003).

Arbitrageurs will usually take positions (shorting overvalued stocks or buying undervalued stocks) and eliminate misvaluations before they become too large. But if it is difficult to take these positions, for example, due to short sales constraints, or if there is no guarantee that the mispricing will be corrected within a acceptable time horizon, then arbitrage is limited and arbitrageurs will not succeed to correct the mispricing. Technically, an arbitrage opportunity exists when one can guarantee a profit by running a certain trading strategy. For example arbitrageurs will go long in an undervalued asset and short in an overvalued asset. Unfortunately in practice, almost all arbitrage activity is risky arbitrage. Thus arbitrageurs tend to trade on opportunities where the expected profit is high relative to the involved risks or if the risks are too large arbitrageurs may even choose to avoid the markets where the mispricing is most significant (Ritter, 2003).

For example, this happened especially in large markets, such as the Japanese stock market in the late 1980s or the US market for technology stocks in the late 1990s. Arbitrageurs who shorted Japanese stocks in mid-1987 and hedged by going long in US stocks were right in the long run. Nevertheless they lost huge amounts of money in October 1987 when the US stock market crashed by more than the Japanese stock market (because of Japanese government intervention). Moreover if the arbitrageurs trade on limited funds, which is a realistic assumption, they would be forced to even up their positions just when the relative mispricings were the greatest. Conversely this resulted in additional buying pressure for Japanese stocks just when they were most overvalued (Ritter, 2003).

To analyze the implications of market efficiency for real markets it is useful to divide market related events into two different categories. First, high-frequency market events, which often take place and are support for the evidence of efficient markets. Therefore it is hard to find a trading strategy that is reliably profitable and mutual fund managers regularly struggle to beat their benchmarks. The second group are low-frequency events, which happen only infrequently and it may take a long time for markets to recover from them. The evidence from low-frequency events does not support market efficiency. According to Ritter (2003) examples of these enormous misvaluations include

- The undervaluation of world-wide stock markets from 1974 to 1982.
- The Japanese stock price and land price bubble of the 1980s.
- The Taiwanese stock price bubble that peaked in February 1990.
- The October 1987 stock market crash.
- The technology, media, and telecom (TMT) bubble of 1999-2000.

2.2 Psychology of Decision Making

In contrast to the neoclassical economic approach behavioral finance relies on models in which some agents are not fully rational, either because of their boundedly rational preferences or because of their mistaken beliefs (misguided beliefs occur because people are bad Bayesians). Thus the behavioral finance approach suggests that stock market participants suffer from certain psychological biases. For example, they under- or overreact to news; they can be overconfident about their private information; they are risk and loss averse; they put to much weight on recent information in their decision, or on the other hand, they are too conservative. Extensive empirical evidence from financial markets and innumerable psychological laboratory experiments build the foundations of behavioral finance. For example, if people are loss averse, a \$2 gain might make people feel better by as much as a \$1 loss makes them feel worse (Ritter, 2003).

The theory of limited arbitrage illustrated that if irrational traders cause deviations from fundamental value, even rational traders may often be ineffective in correcting the emerged mispricings. To analyze the structure of these deviations it is useful to view the systematic biases that arise when people form beliefs or to view people's preferences. It should be emphasized here that behavioral models do not need to make extensive psychological assumptions in order to generate testable predictions (Barberis and Thaler, 2003). For example, the theory of closed-end fund pricing (Lee, Shleifer and Thaler, 1991) makes numerous predictions using only the assumptions that there are noise traders with correlated sentiments in the economy, and that arbitrage is limited.

According to Rabin (2002a) there are many assumptions that economists often make of human nature that are not supported by behavioral and psychological research. These include the assumptions that people

- are Bayesian information processors (see section 2.2.1.2)
- have well-defined and stable preferences
- maximize their expected utility
- exponentially discount future well-being
- are self-interested
- have preferences over final outcomes, not changes
- have only instrumental/functional taste for beliefs and information

Some of the above assumptions have always been subject to doubt, while others are treated as core axioms. And some assumptions are not treated as core axioms in principle, but are pervasively maintained in all actual economic analyses. The goal of behavioral finance is to investigate behaviorally grounded departures from these assumptions that seem economically relevant. For a more concrete frame of reference, it is useful to consider the following formulation of the classical economic model of individual choice, where uncertainty is integrated as probabilistic states of the world, and the assumption that the person maximizes expected value is incorporated (equation 1):

$$\operatorname*{Max}_{x \in X} \sum_{s \in S} \pi(s) U(x|s) \tag{1}$$

While X denotes the choice set, S is state space, π(s) are the person's subjective beliefs updated using Bayes' rule, and U are stable, well-defined preferences. Furthermore, equation (1) assumes even more basic assumptions, for example, that people formulate beliefs even when no "objective" probabilities are available, and that these beliefs are correctly updated according to the laws of probability. Economic models almost always include additional strong assumptions such as "rational expectations" and common priors. From this characterization of the neoclassical model, it is useful to categorize psychological phenomena into three categories (Rabin, 2002a):

- Heuristics and biases in judgment regarding to how people really form their beliefs, implying that their beliefs $p(s) \neq \pi(s)$. For further description see section 2.2.2.
- New assumptions about preferences regarding the shape of the function $U(x|s)$ in equation (1) (evidence described in section 2.2.3).
- Lack of stable utility maximization, questioning that people do really maximize according to equation (2) (described in section 2.2.4).

$$\operatorname*{Max}_{x \in X} \sum_{s \in S} p(s) U(x|s) \tag{2}$$

In the next sections the extensive experimental findings compiled by cognitive psychologists will be reviewed. These are also crucial to explain the emergence of market "anomalies" (chapter 2.3).[2]

2.2.1 Heuristics

Heuristic is the art and science of discovery and invention. The word comes from the same Greek root (`ευρισκω) as "eureka", meaning "to find." A heuristic for a given problem is a way of effectively directing one's attention to a solution. It is different from an algorithm in that it merely serves as a rule of thumb or guideline, as opposed to an invariant procedure. Heuristics may not always achieve the desired outcome, but can be extremely valuable to problem-solving processes. Good heuristics can dramatically reduce the time required to solve a problem by eliminating the need to consider unlikely possibilities or

---

[2] Interestingly, psychology of decision making seems to be an issue not only important to psychologists or economists. Even the CIA reports on cognitive biases in human decision making (see Heuer, 1999).

irrelevant states. The mathematician George Polya popularized heuristics in the mid twentieth century in his book "How to Solve It: A New Aspect of Mathematical Method" (Polya, 1945). He was motivated by his experiences in mathematics education where students are taught mathematical proofs, without learning techniques to formulate proofs themselves. "How to Solve It" is a collection of ideas about heuristics that he taught to his students to teach them efficient ways of looking at problems and methods to find solutions very quickly.

In psychology, heuristics are simple and efficient rules of thumb that people use to make decisions, typically when facing complex problems or incomplete information. These rules work well under most circumstances, but in certain cases lead to systematic cognitive biases. For example, people may tend to perceive more expensive beers as tasting better than inexpensive ones. This finding holds even when prices and brands are switched; putting the high price on the normally relatively inexpensive brand is enough to lead experimental subjects to perceive that beer as tasting better than the beer that is normally relatively expensive. One might call this bias the "price implies quality" bias.[3] Much of the work of discovering heuristics in human decision making was ignited by Amos Tversky and Daniel Kahneman, who shared a significant influence on behavioral finance (see, for example, Kahneman, Slovic and Tversky, 1982).

The Gestalt psychologists Karl Duncker and Wolfgang Koehler preserved the original Greek definition of "serving to find out or discover" when they used the term to describe strategies such as "looking around" and "inspecting the problem" (see, for example, Duncker, 1945). For Duncker, Koehler, and some later theorists, including Herbert Simon (1955), heuristics are strategies that guide information search and modify problem representations to facilitate solutions. From its introduction in the early 1800s up until about 1970, the term "heuristics" has been used to refer to useful and crucial cognitive processes for solving problems that cannot be handled by logic and probability theory (see for example Polya, 1954; Groner, Groner, and Bischof, 1983).

In the past 30 years the definition of heuristics has changed almost to the point of inversion. In research on reasoning, judgment, and decision making, heuristics have come to denote strategies that prevent one from finding out or discovering correct answers to

---

[3] Heuristic (2004, December 15). *Wikipedia: The Free Encyclopedia*. Retrieved December 21, 2004, from URL: http://en.wikipedia.org/wiki/Heuristic

problems that are assumed to be in the domain of probability theory. From this point of view heuristics are poor substitutes for computations that are too demanding for ordinary minds to be carried out. Heuristics have even become associated with inevitable cognitive illusions and irrationality (Piattelli-Palmerini, 1994). This new view of heuristics as poor surrogates for optimal procedures emerged in the 1960s when statistical procedures such as analysis of variance and Bayesian methods became established as the psychologist's methods. These and other statistical tools were transformed into models of cognition and soon cognitive processes became viewed as mere approximations of statistical procedures (Gigerenzer, 1991 and 2000). For example, when Edwards (1968) and his colleagues concluded that human reasoning do not comply with Bayes's rule (see section 2.2.1.2) they proposed that actual reasoning is like a defective Bayesian computer with wrongly combined values (misaggregation hypothesis) or misperceived probabilities (misperception hypothesis). The view that mental processes are "poor replicas" of scientific tools became widespread (Kelley, 1973, p. 109):

> *The assumption is that the man in the street, the naive psychologist, uses a naive version of the method used in science. Undoubtedly, his naive version is a poor replica of the scientific one - incomplete, subject to bias, ready to proceed on incomplete evidence, and so on.*

ANOVA, multiple regression, first-order logic, and Bayes's rule, among others, have been proposed as optimal or rational strategies (see Birnbaum, 1983; Hammond, 1996; Mellers, Schwartz, and Cooke, 1998), and the term "heuristics" was adopted to account for inconsistencies between these rational strategies and actual human cognitive processes. For example, the representativeness heuristic (see section 2.2.1.2), introduced by Kahneman and Tversky (1974 and 1996), was proposed to explain why human inference is like Bayes's rule with the base rates left out (see Gigerenzer and Murray, 1987). The common procedure underlying these attempts to model cognitive processes is to start with a method that is considered optimal, then to eliminate some aspects or calculations, and propose that the mind carries out this naive version (Goldstein and Gigerenzer, 2002).

2.2.1.1 Naive Diversification

Heuristics are simple rules of thumb that make human decision making easier. Nevertheless they can sometimes lead to cognitive biases, especially when people are facing complex problems. These psychological biases can lead to suboptimal investment decisions. For example, when faced with *n* choices for how to invest money, many people allocate using the “1/n rule”. If there are three funds, one-third goes into each. If two are stock funds, two-thirds goes into equities. If one of the three is a stock fund, one-third goes into equities (Ritter, 2003).

Recently, Benartzi and Thaler (2001) have documented that when people diversify, they do so in a naive fashion. Thus many follow the “1/n rule”. Benartzi and Thaler investigated people’s diversification decisions regarding defined contribution saving plans and privatized social security plans. There is a worldwide trend towards defined contribution saving plans in which investment decisions are made by the plan participants themselves (Employee Benefit Research Institute, 1997). Nevertheless there has been expressed concern by economists and financial advisors about the quality of the decisions being made by the participants (see, for example, Olivia, Mitchell and Stephen, 1996). One of the reasons for concern is the lack of financial sophistication in the general public (Bernheim, 1996). For example, a survey by John Hancock Financial Services (1995) found that the greater part of respondents thought that money market funds were riskier than government bonds, and furthermore they felt that their own company stock was safer than a well diversified portfolio. Nevertheless, it is possible that poorly informed employees still make good investment decisions. Benartzi and Thaler (2001) found evidence that participants made decisions that seem to be based on naive notions of diversification. One extreme example they discuss is what they named the “1/n heuristic”. Someone using this rule simply divides her contributions evenly among the *n* options offered in her retirement savings plan. Interestingly, the use of the “1/n rule” has a long history in asset allocation. Indeed, it was recommended in the Talmud. Writing in about the fourth century, Rabbi Issac bar Aha gave the following asset allocation advice: “A man should always place his money, a third into land, a third into merchandise, and keep a third at hand.”[4] There is anecdotal evidence that the rule is still in use. For example, for many

---

[4] The reference to the original Aramaic is „Talmud Bavli, Baba Metzia 42a“.

years TIAA-CREF, one of the largest contribution saving plans in the world, offered two investments: TIAA (bonds) and CREF (stocks). The most common allocation of contributions was 50:50, which was chosen by about half of the participants (Samuelson and Zeckhauser, 1988; TIAA-CREF, 1997). In fact, Harry Markowitz, a pioneer in the development of modern portfolio theory, states that he used this rule himself. He justifies his choice on a psychological basis: “My intention was to minimize my future regret. So I split my contributions fifty-fifty between bonds and equities” (Zweig, 1998).

Benartzi and Thaler (2001) are able to show that a significant proportion of investors follow the “1/n strategy”: they divide their contributions evenly across the funds offered in the plan. Consistent with this naive notion of diversification, they find that the proportion invested in stocks depends strongly on the proportion of stock funds in the plan. In particular, Benartzi and Thaler provide evidence that in 401(k) plans, many people seem to use strategies as simple as allocating 1/n of their savings to each of the *n* available investment options, whatever those options are. Some evidence that people make their decisions according to this heuristic comes from the laboratory. Benartzi and Thaler ask subjects to make an allocation decision in each of the following three conditions: first, between a stock fund and a bond fund; next, between a stock fund and a balanced fund, which invests 50 percent in stocks and 50 percent in bonds; and finally, between a bond fund and a balanced fund. They find that in all three cases, a 50:50 split across the two funds is a popular choice, although of course this leads to very different effective choices between stocks and bonds: the average allocation to stocks in the three conditions was 54 percent, 73 percent and 35 percent, respectively. The 1/n diversification heuristic and other similar naive diversification strategies predict that in 401(k) plans which predominantly offer stock funds, investors will allocate more to stocks. Benartzi and Thaler test this in a sample of 170 large retirement savings plans. They divide the plans into three groups based on the fraction of offered stock funds (low, medium, and high). The allocation to stocks increases across the three groups, from 49 percent to 60 percent to 64 percent, confirming the initial prediction (Barberis and Thaler, 2003).

There are two ways in which such naive diversification behavior could be costly compared to an optimal asset allocation strategy. First, investors might choose a portfolio that is not on the efficient frontier according to the mean-variance portfolio selection theory (Markowitz, 1952). Secondly, they might pick the wrong point along the frontier.

The cost of the first type of error is usually quite small. Even the very naive 1/n strategy will end up with a portfolio that is reasonably close to some point on the frontier. For example, Canner, Mankiw and Weil (1997) show that the popular advice of financial planners, which is inconsistent with traditional models of portfolio selection, results in portfolios that are only 20 basis points below the efficient frontier. Particularly, their paper examines popular advice on portfolio allocation among cash, bonds, and stocks. It documents that this advice is inconsistent with the mutual-fund separation theorem (Tobin, 1958), which states that all investors should hold the same composition of risky assets. In contrast to the theorem, popular advisors recommend that aggressive investors hold a lower ratio of bonds to stocks than conservative investors. The paper explores various possible explanations of this puzzle. It concludes that the portfolio recommendations can be explained if popular advisors base their advice on the unconditional distribution of nominal returns. It also finds that the cost of this money illusion is small, as measured by the distance of the recommended portfolios from the mean-variance efficient frontier. In contrast to Canner, Mankiw and Weil (1997), Brennan and Xia (1998) are able to show that the variation in the ratio of bonds to stocks recommended by financial advisors is quite consistent with a model of portfolio optimization in a dynamic context. The reason for the violation of the separation principle is that bonds are not just any risky asset but have the particular property that their returns covary negatively with expectations about future interest rates. This covariation is important for an investor with a multi-period horizon, as the classic paper of Merton (1973) recognizes. Thus the apparent inconsistency between the Tobin Separation Theorem (Tobin, 1958) and the advice of popular investment advisors, as pointed out by Canner et al. (1997), is shown to be explicable in terms of the hedging demands of optimising long-term investors in an environment in which the investment opportunity set is subject to stochastic shocks.

On the contrary, the second inefficiency, picking an inappropriate point on the efficient frontier, can potentially be quite significant. Brennan and Torous (1999) consider the following example in their paper. They consider an individual with a coefficient of relative risk aversion of two, which is consistent with the empirical findings of Friend and Blume (1975). Then they estimate the loss of welfare from selection of portfolios, which do not match the assumed risk preferences. Assuming an investment horizon of twenty years, an investor who switched from an equity dominated investment strategy (80 percent in stocks) to a bond-rich plan (only 30 percent in stocks) would suffer a utility loss of 25 percent. If the horizon is increased to 30 years then the welfare loss can be up to 40 percent. This indicates significant costs for the investor.

2.2.1.2 Representativeness Heuristic

As mentioned before, heuristics enable decision making with the advantage of reduced mental effort but they may lead to systematic biases and errors in judgment. Many of the decisions people make are based on beliefs concerning the chances of uncertain events. They may be confronted with such events every day. For example, they are asked to predict the outcome of an election, to determine the guilt of a defendant, anticipate the future value of the Canadian dollar, assess the likelihood of a person being thought disordered, etc. In answering such questions, people typically rely on the representativeness heuristic. This mental strategy involves examining the degree to which a person, object, or process matches an ideal or representative model. When an event is judged to have several features which are similar to an "ideal" or prototype model, the mind concludes that the event is more likely to be member of this particular group. Thus under the representativeness heuristic, things are judged as being similar based on how closely they resemble each other using prima facie, superficial qualities rather than essential characteristics.[5]

The representativeness heuristic was first identified by Daniel Kahneman and Amos Tversky (1974). They document that when people try to determine the probability that a data set A was generated by a model B, or that an object A belongs to a class B, they often use the representativeness heuristic. This means that they evaluate the probability by the degree to which A reflects the essential characteristics of B. In many cases representativeness is a helpful heuristic, but it can generate some severe biases (see, for example, Beck-Bornholdt and Dubben, 2001 and 2003, who strikingly describe many real-life applications):

---

[5] Representativeness heuristic (2004, November 29). *Wikipedia: The Free Encyclopedia.* Retrieved December 21, 2004, from URL: http://en.wikipedia.org/wiki/Representativeness_heuristic

- People are insensitive to prior probability of outcomes. Thus they ignore the preexisting distribution of categories or base rate frequencies.
- People are insensitive to sample size. This implicates that they draw strong inferences from a small number of cases.
- People have a misconception of chance, also known as the Gambler's Fallacy. That is why they think chance will "correct" a series of "rare" events.
- People have a misconception of regression. So they misjudge "rare" events since they deny chance as a factor causing extreme outcomes.

The first mentioned bias is base rate neglect. To illustrate, Kahneman and Tversky (1974) present this description of a person named Linda:

> *Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.*

When people were asked whether "Linda is a bank teller" (statement A) or "Linda is a bank teller and is active in the feminist movement" (statement B) is more likely, subjects typically assign greater probability to B. According to probability theory this is incorrect. Representativeness provides a simple explanation for this misjudgment. Since the description of Linda sounds like the description of a feminist, in other words, she is representative for a feminist, subjects tend to pick B rather than A. In contrast, Bayes' theorem (see equation 3) states that people judge incorrectly since they put too much weight on p(description | statement B), which captures representativeness, and too little weight on the base rate, p(statement B).

$$p(\text{statement B} \mid \text{description}) = \frac{p(\text{description} \mid \text{statement B})\, p(\text{statement B})}{p(\text{description})} \tag{3}$$

Bayes' theorem is named after the Reverend Thomas Bayes (1702-61). Bayes worked on the problem of computing a distribution for the parameter of a binomial distribution. His work was edited and presented posthumously (1763) by his friend Richard Price, in "An

Essay towards solving a Problem in the Doctrine of Chances." Bayes' results were replicated and extended by Laplace in an essay of 1774, who apparently was not aware of Bayes' work.[6]

Representativeness also leads to another bias, namely sample size neglect. Human judges tend to underrate or completely ignore base rate information and focus instead on single case information or untested personal experiences. Thus, when judging the likelihood that a data set was generated by a particular model, people often fail to take the size of the sample into account. Since a small sample can be just as representative as a large one, six tosses of a coin resulting in three heads and three tails are as representative of a fair coin as 500 heads and 500 tails are in a total of 1000 tosses. Representativeness states that people will consider the two sets of coin tosses equally informative about the fairness of the coin, although the second set is much more significant.

Sample size neglect implies that in cases where people do not initially know the data-generating process, they will tend to infer too quickly on the basis of too few data points. For example they will believe that a financial analyst with four good stock picks is talented, although four successes are not representative of a bad or average analyst. It also generates a "hot hand" phenomenon. For example, sports fans become convinced that a basketball player who has made three shots in a row is on a hot streak and will score again, even though there is no evidence of a hot hand in the data (Gilovich, Vallone and Tversky, 1985). This belief that even small samples will reflect the properties of the parent population is sometimes known as the "law of small numbers" (Tversky and Kahneman, 1971; Rabin 1998 and 2002b). Assume that people will judge a coin to be unfair if heads comes up 80 percent of the time. It is easier to make this judgment based on five coin tosses than on 20 coin tosses. In summary, people will persistently use a representative strategy without evaluating the sample size that the information is drawn from.

When people do know the data-generating process in advance, the law of small numbers can lead to a Gambler's Fallacy effect. For example, if a fair coin generates five heads in a row, people will claim that "tails are due". Because they believe that even a short sample should be representative of the fair coin, there have to be more tails to balance out the large number of heads (Barberis and Thaler, 2003). Kahneman and

---

[6] Bayes' theorem (2004, November 10). *Wikipedia: The Free Encyclopedia.* Retrieved December 21, 2004, from URL: http://en.wikipedia.org/wiki/Bayes%27_theorem

Tversky (1974) also claim that people have difficulties in identifying truly random sequences. As people expect more alternations than are likely to exist in a random sequence, they expect micro-level instances of randomness to be representative of random sequences over a longer number of trials.

Recently, Griffiths and Tenenbaum (2004) have outlined a framework for understanding the rational basis of the human ability to find structure embedded in noise, viewing this inference in terms of the statistical problem of model selection. Solving this problem for small datasets requires two ingredients: strong prior beliefs about the hypothetical mechanisms by which the data could have been generated, and a rational statistical inference by which these hypotheses are evaluated.

Griffiths and Tenenbaum state that people are extremely good at finding structure embedded in noise. This sensitivity to patterns and regularities is at the heart of many of the inductive leaps characteristic of human cognition, such as identifying the words in a stream of sounds, or discovering the presence of a common cause underlying a set of events. These acts of everyday induction are quite different from the kind of inferences normally considered in machine learning and statistics: human cognition usually involves reaching strong conclusions on the basis of limited data, while many statistical analyses focus on the asymptotics of large samples. The ability to detect structure embedded in noise has a paradoxical character: while it is an excellent example of the kind of inference at which people excel but machines fail, it also seems to be the source of errors in tasks at which machines regularly succeed. For example, a common demonstration conducted in introductory psychology classes involves presenting students with two binary sequences of the same length, such as HHTHTHTT and HHHHHHHH, and asking them to judge which one seems more random. When students select the former, they are told that their judgments are irrational: the two sequences are equally random, since they have the same probability of being produced by a fair coin. In the real world, the sense that some random sequences seem more structured than others can lead people to a variety of erroneous inferences, whether in a casino or thinking about patterns of births and deaths in a hospital (Kahneman and Tversky, 1972).

When assessing the randomness of binary sequences which involves comparing random and regular sources, people's beliefs about the nature of regularity can be expressed in terms of probabilistic versions of simple computing machines. Different machines capture regularity when sequences are presented simultaneously and when their elements are presented sequentially, and the differences between these machines provide insight into the cognitive processes involved in the task. Analyses of the rational basis of human inference typically either ignore questions about processing or introduce them as relatively arbitrary constraints. Griffiths and Tenenbaum (2004) are able to give a rational characterization of process as well as inference, evaluating a set of alternatives that all correspond to restrictions of Kolmogorov complexity to simple general-purpose automata.

2.2.1.3 Conservatism

Several psychologists, including Edwards (1968), have identified a cognitive bias known as conservatism. Conservatism states that individuals are slow to change their beliefs in the face of new evidence. In other words, they anchor on the way things have usually been (Ritter, 2003).

While representativeness (see section 2.2.1.2) leads to an underweighting of base rates, conservatism leads to overemphasized base rates relative to sample evidence. In the experiment run by Edwards (1968), there are two urns, one containing 3 blue balls and 7 red ones, and the other containing 7 blue balls and 3 red ones. A random draw of 12 balls, with replacement, from one of the urns yields 8 reds and 4 blues. When people are asked for the probability the draw was made from the first urn, most people estimate a number around 0.7 (while the correct answer is 0.97). Thus people apparently overweight the base rate of 0.5. At first sight, the evidence of conservatism stands in contrast to the representativeness heuristic. Barberis and Thaler (2003) state that there may be a natural way in which both fit together. It appears that if a data sample is representative of an underlying model, then people overweight the data. On the other hand, if the data is not representative of any salient model, people react too little to the data and rely too much on their priors. In Edwards' experiment, the draw of 8 red and 4 blue balls is not particularly representative of either urn, possibly leading to an overreliance on prior information. According to Ritter (2003) people might underreact because of the conservatism bias. Conversely, if there is a long enough pattern, then they will adjust to it and possibly overreact, underweighting the long-term average.

In particular, Edwards benchmarks a subject's reaction to new evidence against that of an idealized rational Bayesian in experiments in which the true normative value of a piece of evidence is well defined. In his experiments, individuals update their posteriors in the right direction, but by too little in magnitude relative to the rational Bayesian benchmark. This finding of conservatism is in fact more pronounced the more objectively useful the new evidence is. In Edwards' (1968, p. 359) own words: "It turns out that opinion change is very orderly, and usually proportional to numbers calculated from the Bayes theorem - but it is insufficient in amount. A conventional first approximation to the data would say that it takes anywhere from two to five observations to do one observation's worth of work in inducing a subject to change his opinions."

Conservatism is extremely suggestive of the underreaction evidence described above. Individuals subject to conservatism might disregard the full information content of an earnings (or some other public) announcement, perhaps because they believe that this number contains a large temporary component, and still stick at least partially to their prior estimates of earnings. As a consequence, they might adjust their valuation of shares only partially in response to the announcement. Edwards would describe such behavior in Bayesian terms as a failure to properly aggregate the information in the new earnings number with an investors' own prior information to form a new posterior earnings estimate. In particular, individuals tend to underweight useful statistical evidence relative to the less useful evidence used to form their priors. Alternatively, they might be characterized as being overconfident about their prior information (Barberis, Shleifer and Vishny, 1998).

A different explanation comes from social psychology. There is a wealth of research in psychology demonstrating that agents process information with the aid of categories. The distinguished social psychologist Gordon Allport memorably noted, "the human mind must think with the aid of categories. We cannot possibly avoid this process. Orderly living depends upon it." Ideas of categorical thinking and stereotyping have been at the forefront of social psychology for five decades (Macrae and Bodenhausen, 2002; Markman and Gentner, 2001), but their potential has yet to be realized in economics. Fryer and Jackson (2003) introduce a model of how experiences are sorted into categories and how categorization affects decision making. Then they show that specific cognitive biases emerge from categorization. This intuition has even been imported into economics by

Mullainathan (2001), who shows that such categorization models can lead to biased estimates of probabilities. He presents a formal model that neatly reconciles the evidence on underweighting sample information with the evidence on overweighting sample information.

Recently Wilson (2004) introduced a model which explores the connection between bounded memory and biases in information processing. She shows that the optimal memory rule may perform very poorly in the short run, and can explain several biases that psychologists have observed. As a result, the agents appear to display a confirmatory bias, which is the tendency to ignore information that does not support their impressions (see also section 2.2.2.2), and an overconfidence/underconfidence bias with the tendency to infer too much from ambiguous information, too little from precise information.

2.2.1.4 Anchoring and Adjustment

Kahneman and Tversky (1974) state that, when forming estimates, people often start with some initial, possibly arbitrary value, and then adjust away from this reference point. Additionally, Slovic and Liechtenstein (1971) provide experimental evidence that the adjustment is often insufficient. They demonstrate that, in forming numerical estimates of uncertain quantities, adjustments in assessments away from (arbitrary) initial values are typically inadequate. Thus people “anchor” too much on the initial value. Kahneman and Tversky (1974, p. 1128) present the following example:

> *“Subjects were asked to estimate various quantities, stated in percentages (for example the percentage of African countries in the United Nations). For each quantity, a number between 0 and 100 was determined by spinning a wheel of fortune in the subjects’ presence. The subjects were instructed to indicate first whether that number was higher or lower than the value of the quantity, and then to estimate the value of the quantity by moving upward or downward from the given number. Different groups were given different numbers for each quantity, and these arbitrary numbers had a marked effect on estimates. For example, the median estimates of the percentage of African countries in the United Nations were 25 and*

> *45 for groups that received 10 and 65, respectively, as starting points. Payoffs for accuracy did not reduce the anchoring effect."*

In this experiment subjects' subsequent estimates were significantly affected by the initial random number. Those who were asked to compare their estimate to 10, subsequently estimated 25 percent, while those who compared to 60, estimated 45 percent (Barberis and Thaler, 2003).

While this example is to some extent artificial, Kahneman and Tversky (1974) point out that anchoring can occur as a natural part of the assessment process itself. If an individual is asked to construct a probability distribution for the level of the Dow Jones, her likely beginning point would be to estimate a median level. This value would then serve as an anchor or reference point for her further probability assessments. In contrast, if the subject would be asked to construct the probability assessments by stating the likelihood of the Dow Jones exceeding a pre-specified value, she would rather anchor on this value. Thus the two procedures lead to different predictions, with the first procedure yielding a probability distribution more concentrated around the median than the second one (Rabin, 1998).

2.2.1.5 Availability

When judging the probability of an event people tend to search their memories for relevant information. While this is a perfectly sensible procedure, it can produce biased estimates because not all memories are equally retrievable or "available", in the language of Kahneman and Tversky (1974). More recent, salient and vivid events will weigh more heavily and distort the estimate. Furthermore, people tend to correlate events that occur closely together. Tversky and Kahneman (1973) discuss, for example, how salience may distort clinicians' assessment of the relationship between severe depression and suicide. According to the availability heuristic incidents in which patients commit suicide are much more likely to be remembered than are occurrences where patients do not commit suicide. This may lead to an exaggerated assessment of the probability that depressed patients will commit suicide (for further examples, see also Combs and Slovic, 1979).

Ritter (2003) summarizes that people tend to underweight long-term averages and are predisposed to put too much weight on recent experience. As an example, when equity returns have been high for many years (such as they were in 1982-2000 in the US and Western Europe), many people begin to believe that high equity returns are “normal”.

2.2.1.6 Hindsight Bias

One of the most extensively studied biases in the judgment literature is the hindsight bias (see for example Hawkins and Hastie, 1990; Christensen-Szalanski and Willham, 1991; Villejoubert, 2004). Fischhoff (1975) first proposed this bias by observing that

- reporting an outcome’s occurrence increases its perceived probability of occurrence and
- people who have received outcome knowledge are largely unaware of the effect that it has changed their perceptions.

Combining these, the evidence on the hindsight bias shows that people exaggerate the degree to which their beliefs before an informative event would be similar to their current beliefs. For example, after a politician wins elections, people label it as inevitable. Thus they tend to believe that they always thought it was inevitable (Rabin, 1998).

Fischhoff’s (1975) original demonstration of this effect was a historical passage, regarding British intrusion into India and military interaction with the Gurkas of Nepal, read to subjects. Without being told the outcome of this interaction, some subjects were asked to predict the likelihood of each of the four possible outcomes: British victory, Gurka victory, a peace settlement, or military stalemate without a peace settlement. While the latter is the *real* true outcome, four other groups of subjects were each told that a different one of the four outcomes was the true one. For each reported outcome, when compared to a control group, which was not told any outcome, subjects’ average ex post guesses of their hypothetical ex ante estimates were fifteen percent higher than those of the control group. People don’t sufficiently “subtract” information they currently have about an outcome in imagining what they would have thought without the information. Moreover, evidence provided by Hawkins and Hastie (1990) suggests that subjects have the tendency to think that even other people “should have known” as well.

Barberis and Thaler (2003) recapitulate that hindsight bias is the tendency of people to believe, after an event has occurred, that they predicted it before it happened. If people think they predicted the past better than they actually did, they may also believe that they can predict the future better than they actually can (see also section 2.2.2.1 on overconfidence).

2.2.1.7 Recognition Heuristic

One view of heuristics is that they are imperfect versions of optimal statistical procedures considered too complicated for ordinary minds to be carried out. In contrast, Goldstein and Gigerenzer (2002) consider heuristics to be adaptive strategies that evolved in parallel with fundamental psychological mechanisms. Rather than starting with a normative process model Goldstein and Gigerenzer start with fundamental psychological mechanisms. They design and test computational models of heuristics which are:

- ecologically rational (i.e. heuristics that exploit structures of information in the environment)
- founded in evolved psychological capacities such as memory and the perceptual system
- fast, frugal, and simple enough to operate effectively when time, knowledge, and computational resources are limited
- precise enough to be modeled computationally
- powerful enough to model both good and poor reasoning.

Goldstein and Gigerenzer (2002) introduce this program of “fast and frugal” heuristics with perhaps the simplest of all heuristics: the recognition heuristic. The recognition heuristic, arguably the most “frugal” of all heuristics, makes inferences from patterns of missing knowledge. This heuristic exploits a fundamental adaptation of many organisms: the vast, sensitive, and reliable capacity for recognition.

In the statistical analysis of experimental data, missing data are an additional difficulty. According to Gigerenzer and Hoffrage (1995) outside of experimental designs, which means that data are collected by natural sampling instead of systematic sampling, missing knowledge can be used to make intelligent inferences. They asked about a dozen Americans and Germans, “Which city has a larger population: San Diego or San Antonio?” Approximately two thirds of the Americans correctly responded that San Diego is larger. Despite (or, according to the recognition heuristic, due to) a considerable lack of knowledge 100 percent of the Germans answered the question correctly. A similar surprising outcome was obtained when 50 Turkish students and 54 British students made forecasts for all 32 English Football Associations Cup third round soccer matches (Ayton and Önkal, 1997). The Turkish participants had very little knowledge about English soccer teams in comparison to the British participants. Nevertheless, the Turkish forecasters were nearly as accurate as the English ones (63 percent vs. 66 percent correct).

Intuitively these results seem to be an error. How could more knowledge be worse than significantly less knowledge? A look at what the less knowledgeable groups knew may provide an answer. All of the German participants tested had heard of San Diego but about half of them did not recognize San Antonio. All made the inference that San Diego is larger. Similarly, the Turkish students recognized some of the English soccer teams (or the cities that often make up part of English soccer team names) but not others. Among the pairs of soccer teams in which they rated one team as completely unfamiliar and the other as familiar to some degree, they chose the more familiar team in 627 of 662 cases (95 percent). In both these demonstrations, people used the fact that they did not recognize something as the basis for their predictions, and it turned out to serve them well (Goldstein and Gigerenzer, 2002).

The strategies of the German and Turkish participants can be modeled by what Goldstein and Gigerenzer (2002) call the recognition heuristic. The task to which the heuristic is suited is selecting a subset of objects that is valued highest on some criterion. For example, Borges, Goldstein, Ortmann, and Gigerenzer (1999) use the recognition heuristic as a successful device for selecting stock portfolios in a bullish market environment as described in their article “Can Ignorance Beat the Stock Market?” They use corporate name recognition for selecting a subset of stocks from Standard and Poor’s 500 Index, with profit as the criterion. In their empirical study Borges et al. compared the performance of a buy and hold portfolio, constructed only of stocks from companies with a high level of name recognition by either laypeople (pedestrians) or experts, with several benchmarks (mutual funds, market indices, “dartboard” portfolios and unrecognized stocks). The “ignorance” based portfolio constructed by pedestrians unexpectedly outperformed its touchstones and generated striking returns. This also substantiates the hypothesis that heuristics or “bounded rationality” can lead to accurate inferences in real-world domains (see section 2.4.3 on trading strategies).

In these laboratory experiments, the focus lies on the case of selecting one object from two. This task is known as paired comparison or two-alternative forced choice and represents a stock-in-trade of experimental psychology and an elementary case to which many other tasks such as multiple choice are reducible. The recognition heuristic is useful when there is a strong correlation between recognition and criterion. For simplicity, Goldstein and Gigerenzer (2002) assume that the correlation is positive. For two-alternative choice tasks, they describe the recognition heuristic as follows: If one of two objects is recognized and the other is not, then the recognized object is inferred to have the higher value with respect to the criterion. The recognition heuristic will not always apply, nor will it always make correct inferences. For example, the Americans and English in the experiments reported could not apply the recognition heuristic since they knew too much. It is also easy to think of instances in which an object may be recognized for having a small criterion value. Even in such cases the recognition heuristic still predicts that a recognized object will be chosen over an unrecognized object. The recognition heuristic works exclusively in cases of limited knowledge, i.e., when only some objects, not all, are recognized. The effectiveness of the apparently simplistic recognition heuristic depends on its ecological rationality: its ability to exploit the structure of the information in natural environments. The heuristic is successful when ignorance, specifically a lack of recognition, is systematically rather than randomly distributed, particularly when the object is strongly and positively correlated with the criterion. On the contrary, when this correlation is negative, the heuristic leads to the inference that the unrecognized object has the higher criterion value. The direction of the correlation between recognition and the criterion can be learned from experience, or it even can be genetically coded. Galef (1987) and Galef, McQuoid, and Whiskin (1990) provide evidence for the latter. They document an experiment with wild Norway rats that learned to recognize a novel diet by smelling it on their neighbors’ breath. A week later, the rats were fed with this diet in addition with a novel poisoned diet and hence the rats became ill. Next time presented with the two diets, the rats avoided the diet that they did not recognize from their neighbors’ breath. This recognition mechanism works regardless of whether the neighbor rat is healthy or not while its breath is smelt. It may seem unusual that an animal would eat the food its sick neighbor had eaten. Thus rats seem to follow recognition without taking further information, such as the health of the neighbor, into account (Goldstein and Gigerenzer, 2002).

Goldstein and Gigerenzer (2002) review literature that substantiates the view that the mechanism for distinguishing between the novel and the recognized is specialized and robust. For example, recognition memory often remains intact when other types of memory become impaired. Elderly people suffering from memory loss (Craik and McDowd, 1987; Schonfield and Robertson, 1966) and patients suffering from certain kinds of brain damage (Schacter and Tulving, 1994; Squire, Knowlton, and Musen, 1993) have problems saying what they know about an object or even where they have encountered it. Nevertheless they often know that they have encountered the object before. Such is the case with R. F. R., a 54 year old policeman, who developed such severe amnesia that he had great difficulty identifying people he knew, even his wife and mother. One might say that he had lost his capacity for recognition. In a test in which he was shown pairs of photographs consisting of one famous and one non famous person, he could point at the famous persons as accurately as a healthy control group could (Warrington and McCarthy, 1988). His ability to distinguish between the unrecognized people (who he had never seen before) and the recognized people (famous people he had seen in the media) remained intact. Yet his ability to recall anything about the people he recognized was impaired. Laboratory research has demonstrated that memory for mere recognition encodes information even in divided-attention learning tasks that are too distracting to allow more substantial memories to be formed (Jacoby, Woloshyn, and Kelley, 1989). Because recognition continues to operate even under adverse conditions, and it can be impaired independently from the rest of memory, Goldstein and Gigerenzer view it as a primordial psychological mechanism (for cases involving a selective loss of recognition, see Delbecq-Derousné, Beauvois, and Shallice, 1990).

Goldstein and Gigerenzer (2002) investigate recognition as it concerns proper names. Proper name recognition is of particular interest because it constitutes a specialized domain in the cognitive system that can be impaired independently of other language skills (McKenna and Warrington, 1980; Semenza and Sgaramella, 1993; Semenza and Zettin, 1989). Because an individual's knowledge of geography comprises an incomplete set of proper names, it is ideal for recognition studies. In their recent article, Goldstein and Gigerenzer focus on two situations of limited knowledge: Americans' recognition of German city names and Germans' recognition of city names in the United States. The American students they have tested over the years recognized about one fifth of the 100 largest German cities, and the German students recognized about one half of the 100 largest U.S. cities. An additional reason why cities were used to study proper name recognition is because of the strong correlation between city name recognition and population. Evidently the recognition heuristic does not apply everywhere. The recognition heuristic is domain specific, i.e. it works only in environments in which recognition is correlated with the criterion being predicted. In cases of inference, the criterion is not immediately accessible to the organism. A possible way how correlations between recognition and inaccessible criteria can arise is via "mediators". There are mediators in the environment that have the dual property of reflecting (but not revealing) the criterion and being accessible to the senses. For example, a person may have no direct information about the endowments of universities, because this information is usually confidential. The endowment of a university may be reflected in how often it is mentioned in the newspaper. The more often a name occurs in the newspaper, the more likely it is that a person will hear of this name. Because the newspaper serves as a mediator, a person may infer the inaccessible endowment criterion.

Goldstein and Gigerenzer (2002) report an interesting counterintuitive implication of the recognition heuristic: the less-is-more effect. They provide the following thought experiment: Three Parisian sisters receive the bad news that they have to take a test on the hundred largest German cities at their lycée. The test will consist of randomly drawn pairs of cities, and the task will be to choose the more populated city. The youngest sister has never heard of Germany (nor any of its cities) before. The middle recognizes half of the hundred largest cities from what she has overheard in the family salon. The elder sister has been furiously studying for her baccalaureate and has heard of all of the hundred largest cities in Germany. The city names the middle sister has overheard belong to rather large cities. In fact, the 50 cities she recognizes are larger than the 50 cities she does not recognize in about 80 percent of all possible pairs (the recognition validity α is 0.8). The middle and elder sisters not only recognize the names of cities but also have some knowledge beyond recognition. When they recognize two cities in a pair, they have a 60 percent probability of correctly choosing the larger one (the knowledge validity β is 0.6, whereas a β of 0.5 would mean that they have no useful further knowledge). Which of the three sisters will score highest on the test if they all rely on the recognition heuristic whenever they can? The youngest sister can do nothing but to guess on every question. The oldest sister relies on her knowledge (β) on every question and scores 60 percent correct. Surprisingly, the middle sister, who has half of the knowledge of her older sister, scores the greatest proportion of correct inferences (nearly 68 percent correct) since she is the only one who can use the recognition heuristic. Furthermore, she can make the most of her ignorance because she happens to recognize half of the cities, and this allows her to use the recognition heuristic most often. The recognition heuristic leads to a paradoxical situation in which those who know more exhibit lower inferential accuracy than those who know less. Goldstein and Gigerenzer call this the less-is-more effects. Furthermore, they conclude that there will be a less-is-more effect whenever $\alpha > \beta$, that is, whenever the accuracy of mere recognition is greater than the accuracy achievable when both objects are recognized.

Goldstein and Gigerenzer present a further application of the recognition heuristic. They showed students in Germany and the United States lists of pairs of German and American cities. Each group was equally skilled at picking the city with the larger population of each pair, despite differences in familiarity in the other country's geography. The researchers created an exhaustive list of city pairings, having the participants make forced-choice judgments of which city in each pair was larger. They found that when a participant recognized only one city in a pair, she judged that city as larger about 90 percent of the time. Thus the students were able to pick the right answers most of the time because they used the recognition heuristic, according to which the more recognizable city had a higher value (Gigerenzer and Goldstein, 1996; Goldstein and Gigerenzer, 1999 and 2002). Interestingly, in a different study a less-is-more effect was demonstrated with German students who scored higher when tested on American cities than on German ones (Hoffrage, 1995; see also Gigerenzer, 1993).

Although the recognition heuristic has earned enormous scientific attention it is still seen as very controversial. For example, Kahneman and Frederick (2002) describe how the recognition heuristic is a challenge to dual-process theories of judgment, because it doesn’t fit well into most prominent classification schemes for heuristic processing. Furthermore, Oppenheimer (2003) argues that by using the thirty largest German cities as their sample, Goldstein and Gigerenzer (2002) conflated recognition with the knowledge that the city was large. Thus most people not only recognize the city Berlin, but they also know that Berlin is one of the largest cities in Germany. Therefore, it is impossible to determine whether the judgments were due to mere recognition or rather from knowledge that the recognized cities were large. Under these conditions almost any model of judgment, simple or complex, would predict similar results. Since Goldstein and Gigerenzer (1999) clearly assert that the level of recognition is not important in using the recognition heuristic ("the distinction relevant for the recognition heuristic is that between unrecognized objects and everything else") and they discuss the "inconsequentiality of further knowledge" as an essential feature to maintain the frugality of the heuristic. Accordingly, an individual using the recognition heuristic should judge a recognized city as larger than an unknown one even if the recognized city is known to be small. To test this counterintuitive prediction Oppenheimer (2003) asked fifty participants to judge populations of local cities that were known to be small, as compared to made-up cities (which, by virtue of being fictional, were unrecognizable). Oppenheimer's results did not support the predictions of the recognition heuristics as reported by Goldstein and Gigerenzer (1999) and he concludes that although individuals do use recognition as a cue for size estimations, they do so in a more complicated manner. Individuals appear to make attributions about their mental state of recognition, and perform some kind of Bayesian discounting based upon that attribution. Thus the recognition heuristic may simply not be as fast or frugal as it was originally postulated.

Newell and Shanks (2004) document two experiments where they sought to distinguish between the claim that recognition of an object is treated simply as a cue among others for the purpose of decision making in a cue-learning task and the claim that recognition is attributed with a special status of fundamental, non compensatory properties. Results of both experiments supported the former interpretation. When recognition had a high predictive validity, the majority of participants relied solely on it. When other cues in the environment had higher validity, recognition was ignored, and these other cues were used. The results provide insight into when, where, and why recognition is used in decision making and also question the elevated status assigned to recognition in some frameworks as the one by Goldstein and Gigerenzer (2002).

Newell, Weston and Shanks (2003) generally criticize the “fast-and-frugal” heuristics approach to decision making under uncertainty advocated by Gigerenzer and colleagues (for example Gigerenzer and Goldstein, 1996). They state that this framework has achieved great popularity despite a relative lack of empirical validation and report two experiments that examine the use of one particular heuristic (“take-the-best”). In both experiments the majority of participants adopted frugal strategies, but only one-third (33 percent) behaved in a manner completely consistent with take-the-best’s decision rules. Furthermore, a significant proportion of participants in both experiments adopted a non-frugal strategy in which they accumulated more information than was predicted by the theory. These results provide an insight into the conditions under which different heuristics are used, and question the predictive power of the fast-and-frugal approach.

2.2.2 Beliefs

Belief is a representational mental state that takes the form of a propositional attitude. Belief is considered propositional in that it is an assertion, claim or expectation about reality.[7] Cognitive psychologists have documented many patterns regarding how people form their expectations. Furthermore this is a crucial component of any model of financial markets.

2.2.2.1 Overconfidence

A huge amount of evidence shows that people are overconfident in their judgments. In their summary of the “microfoundations” of behavioral finance, DeBondt and Thaler (1995) assert that “perhaps the most robust finding in the psychology of judgment is that

[7] Belief (2004, November 28). *Wikipedia: The Free Encyclopedia.* Retrieved December 21, 2004, from URL: http://en.wikipedia.org/wiki/Belief

people are overconfident". Evidence of overconfidence has been investigated in different contexts.

Studies of the calibration of subjective probabilities found that people tend to overestimate the precision of their knowledge (see Lichtenstein, Fischhoff, and Phillips (1982) for a review of the calibration literature). First, when people are asked to forecast quantities, for example the level of the Dow Jones stock market index, they tend to assign too narrow confidence intervals to their estimates. Alpert and Raiffa (1982) report that their 98 percent confidence intervals include the true quantity only about 60 percent of the time. Second, people are poorly calibrated when estimating probabilities: according to Fischhoff, Slovic and Lichtenstein (1977) events they think are certain to occur actually occur only around 80 percent of the time, and events they consider impossible occur approximately 20 percent of the time. Odean (1998b) reports that overconfidence has been observed in many professional fields. Clinical psychologists (Oskamp, 1965), physicians and nurses (Christensen-Szalanski and Bushyhead, 1981; Baumann, Deber, and Thompson, 1991), investment bankers (Staël von Holstein, 1972), engineers (Kidd, 1970), entrepreneurs (Cooper, Woo, and Dunkelberg, 1988), lawyers (Wagenaar and Keren, 1986), negotiators (Neale and Bazerman, 1990), and managers (Russo and Schoemaker, 1992) have all been observed to exhibit overconfidence in their judgments. The best established finding in the calibration literature is that people tend to be overconfident in answering questions from moderate to extreme difficulty (Fischhoff, Slovic, and Lichtenstein (1977); Lichtenstein, Fischhoff, and Phillips, 1982; Yates, 1990; Griffin and Tversky, 1992). Exceptions to overconfidence in calibration are that people tend to be underconfident when answering easy questions, and they tend to be well-calibrated when predictability is high and when performing repetitive tasks with fast, clear feedback. For example, expert bridge players (Keren, 1987), race-track bettors (Dowie, 1976; Hausch, Ziemba, and Rubinstein, 1981) and meteorologists (Murphy and Winkler, 1984) tend to be well calibrated.

Miscalibration is only one manifestation of overconfidence. Researchers also find that people overestimate their ability to do well on tasks and these overestimates increase with the personal importance of the task (Frank, 1935). People are also unrealistically optimistic about future events. They expect good things to happen to them more often than to their peers (Weinstein, 1980; Kunda, 1987). They are even unrealistically optimistic about pure chance events (Marks, 1951; Irwin, 1953; Langer and Roth, 1975). For example, Ito (1990) reports evidence that participants in foreign exchange markets are more optimistic about how exchange rate moves will affect them than how they will affect others. Over two years the Japan Center for International Finance conducted a bi-monthly survey of foreign exchange experts in 44 companies. Each was asked for point estimates of future yen/dollar exchange rates. The experts in import-oriented companies expected the yen to appreciate (which would favor their company), while those in export-oriented companies expected the yen to fall (which would again favor their company).

Furthermore people tend to have unrealistically positive self-evaluations (Greenwald, 1980). Most individuals see themselves as better than the average person and most individuals see themselves better than others see them (Taylor and Brown, 1988). They rate their abilities and their prospects higher than those of their peers. In the social psychology literature this is a prominent stylized fact, also known as the better-than-average effect. For example, when a sample of U.S. students with an average age of 22 years assessed their own driving safety, 82 percent judged themselves to be in the top 30 percent of the group (Svenson, 1981). Also a modest 51 percent of a group of older Swedish students with an average age of 33 years placed themselves in the top 30 percent of their group. According to Cooper, Woo, and Dunkelberg (1988) 81 percent of 2994 new business owners thought their business had a 70 percent or better chance of succeeding but only 39 percent thought that any business like theirs would be this likely to succeed. Thus when individuals assess their relative skills, they tend to overstate their competence relative to the average (Larwood and Whittaker, 1977; Svenson, 1981). Alicke et al. (1995) present research in which people compare themselves with an average peer. The results have consistently shown that people evaluate themselves more favourably than they evaluate others. Seven studies were conducted to demonstrate that the magnitude of this better-than-average effect depends on the level of abstraction in the comparison. These studies showed that people were less biased when they compared themselves with an individualized target than when they compared themselves with a non-individualized target, for example the average college student. The better-than-average effect was reduced more when the observer had personal contact with the comparison target than when no personal contact was established. Differences in the magnitude of the better-than-average effect could not be attributed to the contemporaneous nature of the target's presentation, communication from the target, perceptual vividness, implied evaluation, or perceptions of similarity. This effect extends to economic decision-making in experiments (Camerer and Lovallo, 1999).

Additionally people overestimate their own contributions to past positive outcomes, recalling information related to their successes more easily than that related to their failures. Fischhoff (1982) writes that “they even misremember their own predictions so as to exaggerate in hindsight what they knew in foresight" (see also section 2.2.1.6 on the hindsight bias). And when people expect a certain outcome and the outcome then occurs, they often overestimate the degree to which they were influential in bringing it about (Miller and Ross, 1975). Taylor and Brown (1988) argue that exaggerated beliefs in one's abilities and unrealistic optimism may lead to “higher motivation, greater persistence, more effective performance, and ultimately, greater success." Moreover, they report that these illusions appear to promote other criteria of mental health, including the ability to care about others, the ability to be happy or content, and the ability to engage in productive and creative work. These strategies may succeed, in large part, because both the social world and cognitive-processing mechanisms impose filters on incoming information that distort it in a positive direction. Thus negative information may be isolated and represented as unthreatening as possible. These positive illusions may be especially useful when an individual receives negative feedback or is otherwise threatened and may be especially adaptive under these circumstances. Nevertheless these beliefs can also lead to biased judgments.

Barberis and Thaler (2003) note that overconfidence may in part be related or stem from two other biases, namely self-attribution bias and hindsight bias. Self-attribution bias refers to a people’s tendency to ascribe any success they have in some activity to their own talents, while blaming failure on bad luck, rather than on their incompetence. When doing this repeatedly, this will lead people to the pleasing but wrong conclusion that they are very talented. For example, according to Gervais and Odean (2001) investors might become overconfident after several quarters of investing success. As described in section 2.2.1.6, hindsight bias is the tendency of people to believe, after an event has occurred, that they already predicted this event before it happened. Barberis and Thaler (2003) conclude that when people think they predicted the past better than they actually did, they may also believe that they can predict the future better than they actually can.

Overconfidence can lead to increased trading activity. Odean (1999) provides evidence that trading volume is in general too high. Odean argues that some investors trade too much because they are overconfident. A particular striking finding is that when predictability of future events is low (as in financial markets), overconfidence seems to be even higher for experts than for novices. Another interesting finding is that men tend to be more overconfident than women. This manifests itself in many ways, including trading behavior. Barber and Odean (2001) recently analyzed the trading activities of people with discount brokerage accounts. They found that the more people traded, the worse they did, on average. Additionally, men traded more, and did worse, than women investors. However, Hirshleifer (2001) states that behavioral finance seems to have focused primarily on miscalibration of private information (overestimation of the precision, i.e. too high mean, too narrow variance).

2.2.2.2 Belief Perseverance and Confirmatory Bias

There is a range of research suggesting that once people have formed strong hypotheses, they are often inattentive to new information contradicting their hypotheses. According to Lord, Ross and Lepper (1979) people are attached to their initial opinions too tightly and for too long. At least two effects appear to be at work. First, people are reluctant to search for evidence that contradicts their beliefs. Secondly, even if they find such evidence, they treat it with excessive skepticism. For example, once someone is convinced that one investment strategy is more profitable than another, she may not be sufficiently attentive to evidence suggesting that the strategy is flawed.

Perkins (1981) provides support for the perspective that ”fresh” thinkers may be better at seeing solutions to problems than people who have meditated at length on the problems, because the fresh thinkers are not overwhelmed by the “interference” of old hypotheses. Psychological evidence reveals a stronger and more provocative phenomenon, known as confirmatory bias: People tend to misread or misinterpret evidence as additional support for initial hypotheses. For example if a teacher initially believes that one student is smarter than another, she has the tendency to confirm that hypotheses when interpreting later performance (Rabin, 1998; Barberis and Thaler, 2003).

Substantiation for the confirmatory bias is a series of experiments, which confront people, who initially differ in their beliefs on a topic, with the same ambiguous information. The counterintuitive result is polarization, i.e. their beliefs move further apart (Lord, Ross and Lepper, 1979; Darley and Gross, 1983; Plous, 1991). Lord, Ross and Lepper (1979) selected twenty-four proponents and twenty-four opponents to the death penalty based on questionnaire information. Then each of the groups were asked to judge the merits of randomly selected studies on the deterrent efficacy of the death penalty, and to state whether a given study provided evidence for or against the deterrence hypothesis. Both groups thought that most of the relevant research supported their own beliefs. Lord, Ross and Lepper found that proponents of the death penalty became on average more in favor of the death penalty believing more in its deterrent efficacy, while opponents became even less in favor of the death penalty and believed even less in its deterrent efficacy.

Rabin and Schrag (1999) explore the confirmatory bias in a symmetric model in which exactly one of two hypotheses is true. They are able to show that the confirmatory bias induces overconfidence: Given any probabilistic assessment by an agent that one of the hypotheses is probably true, the appropriate beliefs should deem it less likely to be true. When the agent believes relatively weakly in a hypothesis after receiving extensive information, the hypothesis in which he believes, may be more likely to be wrong than right. If the confirmatory bias is strong enough, with positive probability the agent may eventually come to believe with close certainty in a false hypothesis even after receiving an infinite amount of information.

2.2.2.3 Optimism and Wishful Thinking

According to Weinstein (1980) most people display unrealistically rosy views of their abilities and prospects. Typically, over 90 percent of those surveyed think they are above average in such domains as driving skill (Svenson, 1981), ability to get along with people and sense of humor (see also section 2.2.2.3 on overconfidence). Furthermore people display a systematic planning fallacy: they predict that tasks (such as writing papers) will be completed much sooner than they actually are (Buehler, Griffin and Ross, 1994; Barberis and Thaler, 2003).

2.2.3 Preferences

This section discusses the psychological research on how people modify their utility functions, which are used as a concept by economists. The assumptions about investor preferences, or about how investors evaluate risky gambles, are a crucial feature of any model trying to understand asset pricing and trading behavior. The vast majority of models assume that investors evaluate gambles according to the expected utility framework introduced by von Neumann and Morgenstern (1944). Von Neumann and Morgenstern are able to show theoretically that if preferences satisfy a number of reasonable axioms, in particular completeness, transitivity, continuity, and independence, then they can be represented by the expectation of a utility function (Barberis and Thaler, 2003).

Unfortunately, related experimental research in the decades after von Neumann and Morgenstern (1944) has provided convincing evidence that people systematically violate expected utility theory when they are forced to choose among risky gambles. As a reaction to this, there has been an explosion of research on so-called non-expected utility theories, with the goal to develop functions which better match the experimental evidence. Prominent research models include weighted-utility theory (Chew and MacCrimmon, 1979; Chew, 1983), implicit expected utility (Chew, 1989; Dekel, 1986), disappointment aversion (Gul, 1991), regret theory (Bell, 1982; Loomes and Sugden, 1982), rank-dependent utility theories (Quiggin, 1982; Segal, 1987, 1989; Yaari, 1987), and prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992).

According to Barberis and Thaler (2003) expected utility theory may be a good approximation to people's evaluation of a risky gamble like the stock market, even if it does not explain attitudes to the kinds of gambles studied in experimental settings. Nevertheless, the difficulties that the expected utility approach has encountered in trying to explain basic facts about the stock market suggests that it is fruitful to look at the experimental evidence. Moreover, recent work in behavioral finance has stated that some of the lessons which one can learn from violations of expected utility are central to the understanding of a number of financial phenomena. Thus financial economists should be more than interested in any of these alternatives to expected utility.

2.2.3.1 Prospect Theory

Prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) is one of the most promising non-expected utility theories for financial applications. Furthermore, it is the most successful in capturing the experimental evidences. Most of the other non-expected utility models may be called quasi-normative theories, since they try to capture some of the anomalous experimental evidence by weakening the von Neumann and Morgenstern (1944) axioms. The difficulty of this approach is that in trying to achieve two goals, to be normative and descriptive, the models end up being unsatisfactory in both criteria. Quite the opposite, prospect theory has no ambitions as a normative theory since it simply tries to capture people's attitudes towards risky gambles as parsimoniously as possible. In fact, Tversky and Kahneman (1986) argue convincingly that normative approaches are doomed to fail, since people routinely make choices that are impossible to justify on normative grounds, because they violate dominance or invariance.

Kahneman and Tversky (1979) design the original version of prospect theory for gambles with at most two non-zero outcomes. They propose that when offered the gamble "get outcome x with probability p, outcome y with probability q", as defined in equation (4), with $x \leq 0 \leq y$ or $y \leq 0 \leq x$, people assign it a value as defined in equation (5).

$$(x, p;\; y, q) \tag{4}$$

$$\pi(p)\upsilon(x) + \pi(q)\upsilon(y) \tag{5}$$

When choosing between different gambles, people pick the one with the highest value. The results of Kahneman and Tversky's (1979) experiments are shown in figure 3 and 4.

Kahneman and Tversky's formulation has a number of important characteristics. First, utility is defined over gains and losses rather than over final wealth positions, an idea first proposed by Markowitz (1952). In the words of the authors, one of the predictions of prospect theory is that "a person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwise" (Kahneman and Tversky, 1979, p. 287). This naturally fits with the way gambles are often presented and discussed in real life.

Overwhelming evidence shows that humans are often more sensitive to how their current situation differs from some reference level than to the absolute characteristics of the situation (Helson, 1964). For example, the same temperature that feels cold when one is adapted to hot temperatures may appear hot when one is adapted to cold temperatures. Understanding that people are often more sensitive to changes than to absolute levels suggests that economists have to incorporate additional factors into their utility analysis, such as habitual levels of consumption. For example, instead of utility at time t, which depends solely on present consumption c_t, utility may additionally depend on a reference level r_t, which is determined by factors like past consumption or expectations of future consumptions (Rabin, 1998).

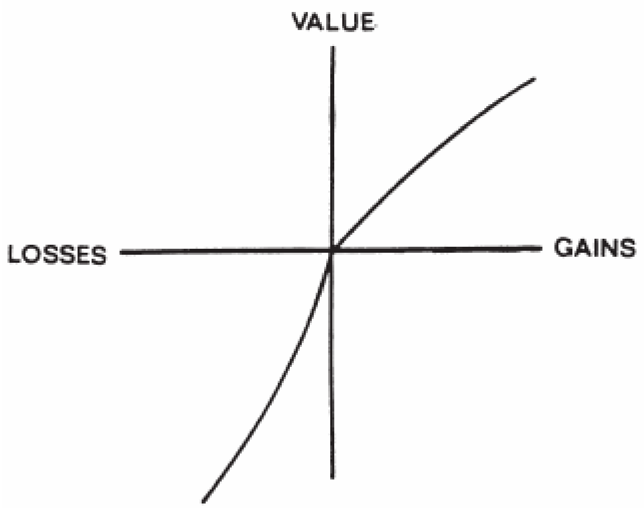


Figure 3: Kahneman and Tversky's (1979) empirically evaluated value function v (from Barberis and Thaler, 2003). v is concave in the domain of gains, while convex and steeper in the domain of losses. This indicates risk seeking behavior for losses and risk averse behavior for gains.

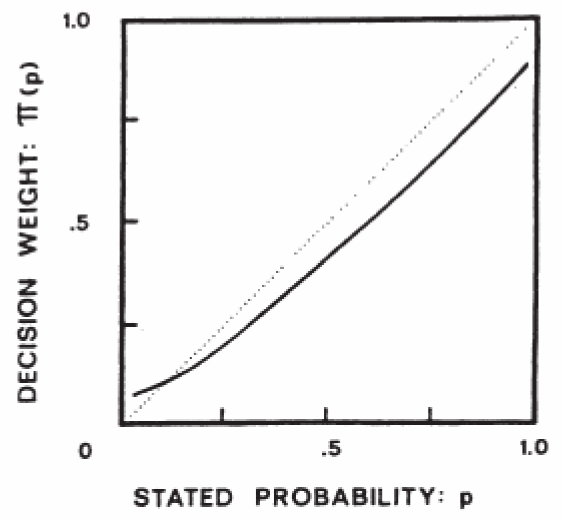


Figure 4: Kahneman and Tversky's (1979) empirically evaluated probability weighting function π (from Barberis and Thaler, 2003). Small probabilities are overweighted, since people are more sensitive to differences in probabilities at higher probability levels. In particular, people place much more weight on outcomes that are certain relative to outcomes that are merely probable, a feature sometimes known as the "certainty effect".

Kahneman and Tversky's (1979) formulation of prospect theory is in general consistent with the way people perceive attributes such as brightness, loudness, or temperature relative to earlier levels of cognition, rather than in absolute terms. Furthermore Kahneman and Tversky (1979) describe the following violation of expected utility as evidence that people focus on gains and losses. In their experiments subjects are asked:[8]

In addition to whatever you own, you have been given 1000 (with probability 0.5). Now choose between

A = (1000, 0.5)
B = (500, 1).

B was the more frequent choice. The same subjects were then again asked:

---

[8] All the experiments are conducted in terms of Israeli currency; the authors note that at the time of their research, the median monthly family income was about 3000 Israeli lira.

In addition to whatever you own, you have been given 2000. Now choose between

$C = (-1000, 0.5)$

$D = (-500, 1).$

Now C was more popular. Both problems lead to identical final wealth positions, but people choose differently. The subjects apparently focus more on gains than losses. Nevertheless, when they are not given any information about prior winnings, they choose B over A and C over D.

According to Barberis and Thaler (2003) the second important feature of prospect theory is the shape of the value function υ since it exhibits concavity in the domain of gains and convexity in the domain of losses, both measured relative to the point of reference (figure 3). The value function is also steeper in the loss domain. Thus Kahneman and Tversky's (1979) value function indicates risk seeking behavior for losses and risk averse behavior for gains. Simple evidence for this outcome is derived from the already mentioned fact that in the absence of any information about prior winnings a statistically significant larger fraction of subjects preferred B to A and C to D. Furthermore the value function υ shows a kink at the origin, which indicates a greater sensitivity to losses than to gains, a feature known as loss aversion. Loss aversion is introduced to capture aversion to bets in the form of (see also equation 4):

$E = (110, ½; -100, ½)$

Game E should be read as "a statistically significant fraction of Kahneman and Tversky's subjects preferred the bet to win 110, with probability ½; G1 to G2."Even to understand attitudes to games as simple as E, it is necessary to depart from the expected utility framework of von Neumann and Morgenstern (1944). In his remarkable paper, Rabin (2000) shows that if an expected utility maximizer rejects gamble E at all wealth levels, then he will also reject game F:

$F = (20000000, ½; -1000, ½)$

This is an absolutely questionable prediction. Rabin (2000) provides a theorem showing that expected utility theory in general is an utterly implausible explanation for appreciable risk aversion over modest stakes: Within expected utility theory, for any concave utility function, even very little risk aversion over modest stakes implies an absurd degree of risk aversion over large stakes. The intuition is simple: if a smooth, increasing, and concave utility function defined over final wealth has sufficient local curvature to reject E over a wide range of wealth levels, it must be an extraordinarily concave function, making the investor extremely risk averse over large stake gambles.

The final important feature of prospect theory is the nonlinear probability transformation. Small probabilities are overweighted, so that the probability weighting function equals $\pi(p)>p$. This is deduced from Kahneman and Tversky's (1979) finding that a statistically significant fraction of subjects preferred the game (5000, 0.001) to (5, 1) and (-5, 1) to (-5000, 0.001) together with the earlier assumption that the value function υ is concave (convex) in the domain of gains (losses). Furthermore, people are more sensitive to differences in probabilities at higher probability levels. Thus a 20 percent jump in probability from 0.8 to 1 is more striking to people than a 20 percent jump from 0.2 to 0.25. In particular, people place much more weight on outcomes that are certain, relative to outcomes that are merely probable, a feature sometimes known as the "certainty effect".

Kahneman and Tversky's (1979) are not just able to capture experimental evidence. Prospect theory is also able to explain preferences for insurance and for buying lottery tickets. Even though the concavity of the value function υ in the region of gains generally produces risk aversion, for example, lotteries, which offer a small chance of a large gain, the overweighting of small probabilities (probability weighting function π) dominates, which leads to risk seeking behavior. On the other hand, while the convexity of the value function υ in the region of losses typically leads to risk seeking, the same overweighting of small probabilities induces risk aversion over gambles which have a small chance of a large loss (Barberis and Thaler, 2003). In more recent research Tversky and Kahneman (1992) also present a generalization of prospect theory which can be applied to gambles with more than two outcomes. Furthermore, they use experimental evidence to estimate a coefficient of loss aversion (λ), a measure of the relative sensitivity to gains and losses.

Over a wide range of experimental contexts λ has been estimated to be equal to approximately 2 (or λ=2.25, as suggested by Tversky and Kahneman, 1992).

To fully describe the decision making process, prospect theory often needs to be combined with an understanding of "mental accounting" (Thaler, 1985, 1990 and 1999; see section 2.2.3.3). One needs to understand when individuals faced with separate gambles treat them as separate gains and losses and when they treat them as one, pooling them to produce one gain or loss (Mullainathan and Thaler, 2000).

2.2.3.2 Framing

People often lack stable preferences that are robust to different ways of eliciting those preferences. The most prominent set of research that provides an interpretation of this type of choice behavior is related to framing effects: Two logically but not transparently equivalent statements of a problem lead decision makers to choose different options (Rabin, 1998).

An important feature of prospect theory is that it can accommodate the effects of problem description or framing. There are numerous demonstrations of a 30 percent to 40 percent shift in preferences depending on the presentation of a problem. No normative theory of choice can accommodate such behavior since a first principle of rational choice is that choices should be independent of the problem description or representation (Barberis and Thaler, 2003).

Framing refers to the presentation of a problem to the decision maker. In many actual choice contexts the decision maker also has flexibility in how to think about the problem. For example, suppose that a gambler goes to the race track and wins $200 in his first bet, but then loses $50 on his second bet. The gambler can now code the outcome of the second bet as a loss of $50 or as a reduction in his recently won gain of $200. In other words, is the utility of the second loss $\upsilon(-50)$ or $\upsilon(150) - \upsilon(200)$? The process by which people formulate such problems for themselves is called "mental accounting" (Thaler, 1999a and 2000). Mental accounting matters since in prospect theory, υ is nonlinear.

Ritter (2003) documents interesting evidence where framing makes a big difference. He starts out with a simple valuation question. Then he lists some specific assumptions about a hypothetical firm, and the final question to answer is, ‘‘How much is the equity of this firm worth?’’

> *Assumptions: The inflation rate is 6%, and the equity risk premium is zero, so the nominal cost of capital is 10% (a real cost of capital of 4%). The firm wants to keep the real value of its debt unchanged, so it must increase the nominal amount of debt by 6% each year. There is no real growth, and all free cash flow (if any) is paid out in dividends.*

Revenue	*$1,200,000*
Cost of goods sold	*$600,000*
Administrative expenses	*$400,000*
Interest expense	*$200,000*
Taxes	*$0*
After-tax profits	*$0*
Debt	*$2,000,000*
Book equity	*$1,500,000*
Shares outstanding	*10,000*
Interest rate on debt	*10%*

According to Ritter (2003), with inflation at 6 percent and $2 million in debt, the firm must issue $120,000 more debt next year to keep the real value of its debt constant. This cash can be used to pay dividends, which would be equal to $12 per share. Equation (6) gives the stock price assuming a growing perpetuity. Using equation (6) with interest rate r=10 percent and growth g=6 percent, P=$12 / (0.10 - 0.06)=$300 per share.

$$P = \frac{DIV_1}{(r-g)} \tag{6}$$

Thus the equity is worth $3 million, or the mentioned $300 per share. Earnings are zero because the accountants treat nominal interest expense as a cost, but they do not treat the inflation induced decrease in the real value of debt as a benefit to equity holders. In other words, the real economic earnings are higher than the accounting earnings, because accountants measure the cost of debt financing, but not the benefit to equity holders, when there is inflation.

This is an example of where framing makes a difference. Nominal interest expense appears on the income statement. The decrease in the real value of nominal liabilities due to inflation does not appear on the income statement. Because it does not appear, investors do not take it into account and therefore undervalue equities when inflation is high. If the market makes this mistake, then stocks become riskier, because they fall more than they should when inflation increases, and they rise more than they should when inflation decreases. Over a full inflation cycle, these two effects balance out, which is why stocks are less risky in the long run than they are in the short run (Siegel, 1998).

Modigliani and Cohn (1979) argued that the US stock market was grossly undervalued in the mid and late 1970s because investors had irrational beliefs about earnings, given the high inflation at that time. Ritter and Warr (2002) investigate the effect of declining inflation on the bull market of 1982-1999. In their work they conduct an out-of-sample test of the Modigliani-Cohn hypothesis. Their results suggest that part of the bull market of the 1980s was attributable to a recovery from the undervaluation. Furthermore they argue that the continued stock market rise in the 1990s was an overshooting and the stock market became overvalued. They also predicted that 2000-2002 would have low stock returns. Ritter (2003) states that, fortunately, he believed in his own research, and had much of his retirement assets in inflation-indexed bonds for the last three years. These have been the best-performing asset class.

2.2.3.3 Mental Accounting

> *"A former colleague of mine, a professor of finance, prides himself on being a thoroughly rational man. Long ago he adopted a clever strategy to deal with life's misfortunes. At the beginning of each year he establishes a target donation to the local United Way charity. Then, if anything untoward happens to him during the year, for example an undeserved speeding ticket, he simply deducts this loss from the United Way account. He thinks of it as an insurance policy against small annoyances."*
>
> Richard H. Thaler (1999a)

The preceding anecdote illustrates the cognitive process called mental accounting. People act as if they associate different classes of risky assets to different mental accounts (for example bonds and stocks, see Thaler, 1990 and 1999).

Tversky and Kahneman (1981) define a mental account quite narrowly as "an outcome frame which specifies (i) the set of elementary outcomes that are evaluated jointly and the manner in which they are combined and (ii) a reference outcome that is considered neutral or normal" (typically the reference point is the status quo). According to this definition, a mental account is a frame for evaluation. Thaler (1999a) uses the term "mental accounting" to describe the entire process of coding, categorizing, and evaluating events. In their later work Kahneman and Tversky (1984) propose three ways that outcomes might be framed: in terms of a minimal account, a topical account, or a comprehensive account. Comparing two options using the minimal account entails examining only the differences between the two options, disregarding all their common features. A topical account relates the consequences of possible choices to a reference level that is determined by the context within which the decision arises. A comprehensive account incorporates all other factors including current wealth, future earnings, possible outcomes of other probabilistic holdings, etc. Thus economic theory generally assumes that people make decisions using the comprehensive account. The following example from Tversky and Kahneman (1981) illustrates that mental accounting is topical:

> *Imagine that you are about to purchase a jacket for ($125)[$15] and a calculator for ($15)[$125]. The calculator salesman informs you that the calculator you wish to buy is on sale for ($10)[$120] at the other branch of the store, located 20 minutes drive away. Would you make the trip to the other store?*

When two versions of this problem are presented (one with the figures in parentheses, the other with the figures in brackets), most people say that they will travel to save the $5 when the item costs $15 but not when it costs $125. If people use a minimal account frame they would be just asking themselves whether they are willing to drive 20 minutes to save $5, and would give the same answer in either version. Interestingly, a comparable analysis applies in the comprehensive account frame. If W equals the existing wealth and W* equals existing wealth plus the jacket and calculator minus $140 then the choice comes down to the utility of W* plus $5 versus the utility of W* plus 20 minutes. This example illustrates an important general point: the way a decision is framed will not alter choices if the decision maker is using a comprehensive and wealth based analysis. Thus framing does alter choices in the real world because people make decisions gradually and are influenced by the context of the choice (Thaler, 1999a).

One important feature of mental accounting is narrow framing, which is the tendency to treat individual gambles separately from other portions of wealth. For example, when offered a gamble, people often evaluate it as if it is the only gamble they face in the world, rather than merging it with already presented bets to see if the new bet is a worthwhile addition (Barberis and Thaler, 2003).

Shefrin and Statman (2000) apply mental accounting to asset allocation and present their notion of a behavioral portfolio theory. The result of investors who connect different classes of risky assets to different mental accounts is the construction of a layered portfolio. According to Shefrin and Statman behavioral based portfolios resemble layered pyramids, where layers are associated with objectives. Thus when such a portfolio is segregated into multiple mental accounts, covariances among mental accounts are overlooked and the investors can be simultaneously risk averse and risk seeking: they buy both bonds and lottery tickets.

2.2.3.4 Ambiguity Aversion

The experimental evidence described so far has focused on understanding how people act when the outcomes of gambles have known objective probabilities. In reality, probabilities are rarely objectively known. Therefore Savage (1954) developed a counterpart to expected utility known as Subjective Expected Utility. Under certain axioms of Subjective Expected Utility, preferences can be represented by the expectation of a utility function, weighted by an individual's subjective probability assessment. Nevertheless, experimental evidence in the last few decades did not substantiate the predictions of Subjective Expected Utility.

The classic experiment by Ellsberg (1961) describes the so called Ellsberg paradoxes, which suggest that people are averse to ambiguity, causing them to make irrational choices. In Ellsberg's experiments (1961) there are two urns, numbered with one and two. Urn two contains a total of 100 balls, 50 red and 50 blue. Urn one also contains 100 balls, again a mix of red and blue, but the subject does not know the proportion of each. Subjects are now asked to choose one of the following two gambles a_1 and a_2, each of which involves a possible payment of $100, depending on the color of a ball drawn at random from the relevant urn.

a_1 : a ball is drawn from Urn 1, $100 if red, $0 if blue,
a_2 : a ball is drawn from Urn 2, $100 if red, $0 if blue.

Subjects are then also asked to choose between the following two gambles b_1 and b_2:

b_1 : a ball is drawn from Urn 1, $100 if blue, $0 if red,
b_2 : a ball is drawn from Urn 2, $100 if blue, $0 if red.

a_2 is typically preferred to a_1, while b_2 is chosen over b_1. These choices are inconsistent with Subjective Expected Utility since the choice of a_2 implies a subjective probability that less than 50 percent of the balls in urn one are red, while the choice of b_2 implies the opposite. Thus the experiment suggests that people do not like situations where they are uncertain about the probability distribution of a gamble. Such situations are known as situations of ambiguity, and the general dislike for them is known as ambiguity aversion. Interestingly, Knight (1921) provides an early discussion of this aversion. He defines risk as a gamble with known distribution and uncertainty as a gamble with unknown distribution, and suggests that people dislike uncertainty more than risk. Subjective Expected Utility does not allow agents to express their degree of confidence about a probability distribution and therefore cannot capture such aversion.

Ambiguity aversion appears in a wide variety of contexts and has been confirmed in market experimental settings. It seems to reflect a more general tendency for emotions such as fear to affect risky choices (Peters and Slovic, 1996). Camerer (1995) suggests that ambiguity aversion may improperly increase risk premia improperly when new financial markets are introduced, because of the layering of uncertainty of both the structure of the economic environment and the structure of the resulting outcomes. A possible explanation for ambiguity aversion is that the absence of an identifiable parameter of the decision problem may often be associated with higher risk and the possibility of hostile manipulation. Heath and Tversky (1991) argue that in the real world ambiguity aversion has much to do with how competent an individual feels at assessing the relevant distribution. Ambiguity aversion over a bet can be strengthened by highlighting individuals' feelings of incompetence, either by showing them other bets in which they have more expertise, or by mentioning other people who are more qualified to evaluate the bet (Fox and Tversky, 1995). Further evidence that supports the competence hypothesis is that in situations where people feel especially competent in evaluating a gamble, the opposite of ambiguity aversion, namely a "preference for the familiar", has been observed. Just as with ambiguity aversion, such behavior cannot be captured by the theory of Subjective Expected Utility (Hirshleifer, 2001; Barberis and Thaler, 2003).

According to Daniel, Hirshleifer, and Teoh (2002) investors often do not participate in certain asset and security categories. In the absence of transaction costs, mean-variance optimization implies participating in all asset and security markets. For many years prior to the rise of mutual funds and defined contribution retirement plans, participation in the U.S. stock market was very incomplete (see for example Blume and Friend, 1975). Even now, many investors entirely neglect major asset classes such as commodities, stocks, bonds, real estate, and omit many individual securities within each class. Investors are strongly biased toward investing in stocks based in their own home country (see for example

Cooper and Kaplanis, 1994; Lewis, 1999). Mutual funds tend to invest locally, and earn higher returns on their local investments (Coval and Moskowitz, 2001), which is consistent with either rational processing of private information or with limited ability to process public information. Investors with more social ties are more likely to participate certain asset and security categories (Hong, Kubik and Stein, 2001). A possible source of non-participation is aversion to ambiguity, as reflected in the Ellsberg paradox. For example, Sarin and Weber (1993) experimentally find that graduate business students and bank executives were averse to gambles with “ambiguous” probabilities relative to equivalent lotteries, and that this aversion affected market prices. Employees tend to invest in their own firm’s stocks and perceive this stock as low risk (Huberman, 1999). The degree to which they invest in their employer’s stock does not predict the stock’s future returns (Benartzi, 2001), suggesting that the investment is not based on superior inside knowledge of their own firm.

2.2.3.5 Loss Aversion and the Disposition Effect

A pervasive feature of reference dependence (see section 2.2.3.2), which is evident in a wide variety of domains, is that people are significantly more averse to losses than they are attracted to same-sized gains (Kahneman, Knetsch, and Thaler, 1990). As mentioned in section 2.2.3.1 on prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), such loss aversion is evident in preferences over wealth levels. Tversky and Kahneman (1991) suggest that in the domain of money people value modest losses roughly twice as much as equal sized gains. In addition to loss aversive behavior individual investors engage in mental accounting (Thaler, 1985), which can drive them to confuse the unpleasantness of experiencing an economic loss with the unpleasantness of realizing the loss. Both cognitive biases are related to the disposition effect, as confirmed by several studies of behavior in field and experimental markets, since investors are more prone to realize gains than losses (Rabin, 1998; Daniel, Hirshleifer, and Teoh, 2002).

The disposition effect is the individual investor’s tendency to hold losers too long and sell winners too soon (Shefrin and Statman, 1985). In particular, Odean (1998a) shows that the individual investors who trade through a large discount brokerage firm tend to be more likely to sell their winners than their losers. Furthermore Odean shows that the stocks that investors choose to sell subsequently outperform the stocks that investors retain. A substantial amount of the underperformance of the losers relative to the winners derives from the momentum effect, but momentum does not explain all of the underperformance of these investors. Interestingly, the individual investor behavior that Odean observes goes against the commonly known investing maxim: “Ride your winners and sell your losers”. This investing maxim may be designed as a corrective to individual biases. An open question is who is taking the opposite side of these individual investors’ transactions. There is some evidence consistent with institutional investors (for example mutual funds) buying high momentum stocks and selling low momentum stocks, but until now there is no direct evidence that links the sales of individual investors to the purchases of mutual funds. One relevant fact is that there are large flows into mutual funds which have experienced good past performance. Home sellers also appear to be loss averse in the way that they set prices. Their reluctance to sell at a loss relative to past purchase price may help to explain the strong positive correlation of volume with price movements (Genesove and Mayer, 2001).

Therefore, the disposition effect shows up in aggregate stock trading volume. During a bull market, trading volume tends to grow. If the market then turns down, trading volume tends to fall. For example, trading volume in the Japanese stock market fell by over 80 percent from the late 1980s to the mid-1990s. The fact that volume tends to fall in bear markets results in an increased level of systematic risk taken by the commission business of brokerage firms. An exceptional case is represented by the U.S. bear market beginning in April 2000, when aggregate stock market volume did not drop. This is apparently due to increased trading by institutions, since stock trading by individuals has, in fact, declined. The significant drop in transaction costs associated with the move to decimalization and technological advances partly accounts for this effect. Another reason is that many firms split their shares in late 1999 and the first half of 2000, which would have resulted in higher trading volume. The drop in commission revenue derived from individuals, as predicted by the disposition effect, has resulted in revenue declines for retail-oriented brokerage firms (Ritter, 2003).

2.2.3.6 Social Preferences

> *"It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard for their own interest. We address ourselves not to their humanity, but to their self-love, and never talk to them of our necessities, but of their advantage."*
>
> Adam Smith (1776)

There is not much to disagree with Smith's poetic analysis of the motivation driving most market behavior. Yet pure self-interest is far from a complete description of human motivation, and realism suggests that economists should move away from the presumption that people are solely self-interested. Massive experimental evidence (see for example Dawes and Thaler, 1988) makes clear that preferences depart from pure self-interest in non-trivial ways: subjects contribute to social goods more than can be explained by pure self-interest; they often share money when they could readily grab it for themselves; and they often sacrifice money to retaliate for unfair treatment (Rabin, 1998).

A simple hypothesis for how people care about others' well being has the longest history in economics: altruism, the positive concern for others as well as oneself. Altruism can be either "general" or "targeted". One may care about all others' well being, or maybe selected others' (friends, family) well being. Most often the more a sacrifice helps somebody the more likely one is willing to make this sacrifice. This is predicted by simple altruistic preferences, which assume that people weight others' utility positively in their own utility function. In this sense, assuming simple altruism provides insight into departures from self-interest. But such simple altruism is not adequate for understanding many behaviors. Two other aspects of social preferences show up prominently in psychological and recent experimental economic evidence. First, people care about the fairness and equity of the distribution of resources, beyond ways that just increase direct total well being. Secondly, people care about intentions and motives and want to reciprocate the good or bad behavior of others. The literature that identifies the nature of social preferences is among the most active areas of research in experimental economics (Rabin, 2002a). Recently the research group of Ernst Fehr at the Institute for Empirical

Research in Economics at the University of Zurich[9] has attracted a lot of interest for their work in this field (see for example Fehr and Gächter, 2002).

Charness and Rabin (2002) design a range of simple experimental games that test these theories more directly than existing experiments. Their experiments show that subjects are more concerned with increasing social welfare, sacrificing to increase the payoffs for all recipients, especially low-payoff recipients, than with reducing differences in payoffs. Furthermore, they find that subjects are also motivated by reciprocity: they withdraw willingness to sacrifice in order to achieve a fair outcome when others also are unwilling to sacrifice. Moreover, unfair behavior is sometimes punished. Charness and Rabin provide the following examples involving decisions regarding how much money (either pennies in Berkeley, California, or pesetas in Barcelona, Spain) to allocate between two anonymous parties. The first example involves Party C choosing between two different allocations for two other anonymous parties, A and B:

C chooses between (A, B) allocations : ($7.50, $3.75) vs. ($4.00, $4.00)

They find an approximate allocation of 50:50. A natural interpretation of these findings (consistent with other experimental evidence) is that C may want to help these parties, but cares about both social efficiency and "equality" producing a desire to help the party that is worse off. Those who care relatively more about social efficiency choose the higher total-surplus outcome ($7.50, $3.75), while those caring more about helping the "worse off" choose ($4.00,$4.00).

Now Charness and Rabin consider the same situation, except that B, as one of the two interested parties, is making the choice.

B chooses between (A, B) allocations : ($7.50; $3.75) vs. ($4.00, $4.00)

Charness and Rabin find that subjects select a 40:60 allocation on average. B does indeed seem to have similar preferences as the neutral party C, though B is a bit less willing to choose the allocation that is good for A and bad for herself. This difference, which in these

---

[9] URL: http://www.iew.unizh.ch/home/fehr/

cases and by replication is small but statistically significant, may be because B is self-interested, or because she is envious of coming out behind A. The previous two examples illustrate how parties might assess the attractiveness of different allocations in a "reciprocity-free" context. That is, one party is making a decision that affects one or more other parties who have not themselves behaved nobly.

To see how reciprocation of the behavior of others might affect choice, now suppose that B makes the same choice as in the previous example, but chooses after A has created her choice by rejecting ($5.50, $5.50). A's decision to forego an allocation of ($5.50, $5.50) in favor of trying to get B to choose ($7.50, $3.75) is clearly selfish and unfair behavior, since it involves a small increase in total surplus while leading to an unequal allocation. Charness and Rabin's findings are as follows:

B chooses between (A; B) allocations : ($7:50; $3:75) vs: ($4:00; $4:00)

They find an approximate allocation of 10:90. They conclude that B is much less willing to sacrifice in order to give the good allocation to A, following the obnoxious choice by A. B's choice in the previous two examples is identical in terms of outcomes. And yet here, and in many related examples, players in games behave systematically differently as a function of previous behavior by other players. This shows that people care not just about outcomes, but also how they arrived at those outcomes. The fact that preferences cannot solely be defined over outcomes can be reconciled with preference theory, but requires an expansion of the notion of what enters the utility function. These additional complications appear necessary to do justice to economic models and are crucial for understanding the nature of retaliation and reciprocal altruism. Fehr and Gächter (1998) even state the existence of a "Homo Reciprocans". According to their work, reciprocity can account for a wide range of empirical phenomena, since it

- is a powerful effort elicitation device
- explains why employers refuse to hire underbidders and, hence why wages are downwardly rigid
- gives rise to non-compensating wage differentials and to a positive correlation between profits and wages
- provides a rationale for the absence of explicit financial incentives

- is a key force that sustains social norms

Nevertheless, they also found a non-negligible fraction of individuals who do not reciprocate and behave completely selfish. The coexistence of reciprocal and selfish types raises exciting questions about their possible interaction behavior.

Recently Quervain et al. (2004) investigated the underlying neural mechanisms of altruistic punishment by using Positron Emission Tomography (PET) in an economic exchange experiment. Subjects could punish defection either symbolically or effectively. Symbolic punishment did not reduce the defector's economic payoff, whereas effective punishment did reduce the payoff. Quervain et al. scanned the subjects' brains while they learned about the defector's abuse of trust and determined the punishment. Effective punishment, as compared with symbolic punishment, activated the dorsal striatum (a brain area), which has been related to the processing of rewards that accrue as a result of goal-directed actions. Moreover, subjects with stronger activations in the dorsal striatum were willing to incur greater costs in order to punish. The results support the hypothesis that people derive satisfaction from punishing norm violations and that neuronal activation in the dorsal striatum reflects the anticipated satisfaction from punishing defectors.

2.2.4 Emotions

"Of all the ways of defining man, the worst is the one which makes him out to be a rational animal."
Anatole France

Emotions are a neglected topic, especially in the field of economics. This is surprising since economics is concerned with the best ways of promoting human satisfaction in a world of scarce resources. With one exception, all human satisfaction comes in the form of emotional experiences. The exception is the hedonic satisfaction produced by the senses, for example the taste of sweetness on the tongue or the feeling of wind on one's face after a long climb. Such sensations differ from emotions in that no prior cognition is necessary to produce them. Thus one does not have to recognize the wind as wind to enjoy the sensation. On the contrary, to get angry when an Albanian host offers a cup of tea by passing it under his left arm one has to know that in Albania this is considered an insult. For infants, sensations may be the most important source of satisfaction. For most adults, they take second place. If one grants the truth of that claim, or even of the weaker claim that emotional experiences are important sources of human satisfaction, one would expect economists to have studied the ways in which people organize their life to maximize emotional satisfaction. Furthermore economists should have tried to identify sources of suboptimal emotion-seeking behavior and should suggest ways of improving this behavior (Elster, 1996).

Contemporary economic work focuses exclusively on the role of emotions in sustaining (or preventing) cooperative interactions (Hirshleifer, 1987; Frank, 1988). No economist has considered emotions in their main role as providers of pleasure, happiness, satisfaction, or utility. To put it crudely, economists have totally neglected the most important aspect of their subject matter. No doubt there are reasons for this neglect. One is the lack of a metric. If you ask someone whether he prefers shame or grief, whether he would rather be caught cheating at an exam or have his girlfriend leave him, he would probably be at a loss for an answer. Emotions themselves interfere with our ability to observe them (Montaigne cites Petrarch to the effect that “He who can describe how his heart is ablaze is burning on a small pyre”). Another reason may be the lack of good theories of how emotions are triggered and transformed in encounters with the world. A further reason may be that emotional satisfaction is largely (but not only) derived from encounters with other people rather than from material goods and that these are encounters not mediated by the market. A final reason may be an unclear insight that people do not usually try to maximize their good emotional experiences, and that they are likely to fail if they try. Economists may be deterred from studying emotions simply because people do not seem to manage their emotional life very rationally (Elster, 1996).

Emotions, like beliefs and desires, can be conceived either as current mental events or as dispositions for such events to occur. Whereas emotions are only to a small extent under the control of the will, dispositions can be consciously shaped to a larger extent. A straightforward characterization of the emotions might be that they go together with physiological arousal. The arousal need not be very strong, and may arguably be absent altogether, as in the puzzling case of the aesthetic emotions (Elster, 1996).

One feature that distinguishes emotions from pain is that emotions have an intentional object. In that respect, they are like beliefs and desires. Psychologists argue that this feature may also be absent, for example in free-floating anxieties. Although one could think of such cases as dispositions to feel anxious about many individual occurrences. In most cases emotions are intentional. As Hume warned, one should take care not to confuse the object of an emotion with its cause: If someone, for instance, receives bad news in the mail, he might react by getting angry at his family. Also unlike pain, emotions usually have a cognitive antecedent (for some exceptions see Goleman, 1995). Before we can react emotionally to a situation, we have to process it cognitively. We must decide whether the person stamping on my foot on the subway did so intentionally, whether the person who got the job I wanted obtained it by immoral means, etc. Often, as we shall see, the emotions have cognitive consequences as well, i.e. they may cause a reassessment of the situation that caused them in the first place. When pain keeps us awake at night, we want it to end. When we are kept awake by love, we want the emotion to continue. Unlike pain and unlike emotions such as grief or guilt, love is a highly desirable disposition. In the language of psychology, it has positive valence. Other emotions such as the ones just mentioned have negative valence. Thus we would rather not have them (but we might welcome the disposition to have them). Some emotions may be neutral in this respect, such as a bittersweet feeling of nostalgia that is caused by a contrast effect and an endowment effect that exactly offset each other (Tversky and Griffin, 1991).

In addition to arousal, intentionality, cognitive antecedents, and valence, most emotions are associated with a characteristic action tendency. The action tendency of shame is to hide or disappear; that of guilt is reconciliation and repairs; that of anger, to strike; that of fear, to run; that of joy, to dance. But not all emotions have such action tendencies: Sadness and grief, and the aesthetic emotions, do not seem to have any. Although spontaneous emotional urges are largely outside the control of the will, we can refrain from acting on them (see Goleman, 1995). Furthermore, emotions tend to have visible physiological expression: turning red, turning pale, smiling, baring one's teeth, crying, blushing, fainting, frowning, etc. These expressions are related both to arousal and to action tendency, although different from both. Unlike arousal, the expression of an emotion is to some extent within the control of the will (but see Ekman, 1992). Unlike action tendencies, expressions are not intentional (Elster, 1996).

How are emotions related to decision making? Damasio (1994) and LeDoux (1996) argue that emotions improve decision making in two respects. First, they enable us to avoid procrastination, i.e. to make a decision when it matters rather than making the optimal decision. Thus emotions allow us to decide among options none of which is rationally superior to the others (De Sousa, 1987). Second, in some cases emotions can actually help us making the best decision. In both cases, it is assumed that decision guided by emotions and reason is better than what can be achieved by rational deliberation alone (Elster, 1998).

The earliest documented statement of the first problem was by the philosopher Ronald de Sousa (1987). He observed that rational-choice theory is indeterminate in many situations, since it does not allow us to identify the uniquely optimal action. Winter (1964, p. 252) observed that the attempt to reduce satisficing (Simon, 1955) to a form of maximizing gives rise to an infinite regress, because the "choice of a profit-maximizing information structure itself requires information, and it is not apparent how the aspiring profit maximizer acquires this information, or what guarantees that he does not pay an excessive price for it." De Sousa (1987, p. 195) states that the "role of emotion is to supply the insufficiency of reason […] For a variable but always limited time, an emotion limits the range of information that the organism will take into account, the inferences actually drawn from a potential infinity, and the set of live options from which it will choose."

Damasio (1994) argues that emotional responses enhance our capacity to make good decisions, not by guiding us to the best possible decision, but by ensuring that we make *some* decision in situations where procrastination is likely to be disastrous. The implicit premise of this interpretation is that rationality has an "addiction to reason" (Elster, 1989, p. 117). Some people do indeed seek to make all decisions on the basis of sufficient reasons. But that makes them irrational rather than rational. A rational person would know that under certain conditions it is better to follow simple decision rules (heuristics) than to use more sophisticated procedures with higher opportunity costs. Thaler (1980) argues that neglect of opportunity costs and excessive focus on out-of-pocket expenses is a frequent source of cognitive irrationality. Furthermore the neglect of opportunity costs that are created by the fact that decision making takes time is also an important and persistent source of irrationality.

For the second problem Damasio (1994) provides evidence based on work with patients with brain lesions. Although he speculates that "Reduction in emotion may constitute an [...] important source of irrational behavior" (p. 53), his work only supports the weaker conclusion that "The powers of reason and the experience of emotion decline together" (p. 54). Thus Damasio only proves the correlation that brain-lesioned patients are both emotionally flat and unable to make decisions, but not the causation. His studies show that people use "somatic markers" to make decisions in largely indeterminate and complex situations. Somatic markers are referred to as gut feelings that are not available to the emotionally disabled, who for that reason tend to procrastinate indefinitely. Somatic markers may also help us to form rational beliefs. First, many pieces of information that we possess are not consciously acknowledged. Secondly, the cognitive basis of the emotions includes unconscious knowledge (Elster, 1996).

2.2.5 Discussion

One weakness of behavioral finance, as viewed by critics, is that there obviously exist competing behavioral explanations for particular empirical facts. It is sometimes said that the long list of cognitive biases offers behavioral modelers so many degrees of freedom that anything can be explained. For example Gigerenzer (1996) presents a profound critique of Kahneman and Tversky's so-called heuristics-and-biases approach. He states that according to their approach, judgments of probability or frequency are sometimes influenced by what is similar (representativeness), comes easily to mind (availability), and comes first (anchoring). Gigerenzer argues that the problem with these heuristics is that they at once explain too little and too much: too little, because one does not know when they work and how; too much, because, post hoc, one of them can be fitted to almost any experimental result. For example, base rate neglect is commonly attributed to representativeness. However, the opposite result, overweighting of base rates ("conservatism"), is as easily "explained" by saying the process is anchoring (on the base rate) and adjustment. Furthermore Gigerenzer argues that there are two major obstacles to understanding these cognitive processes. The first is that the norms for evaluating reasoning have been too narrowly drawn, with the consequence that judgments deviating from these norms are mistakenly interpreted as cognitive illusions. The second is that vague heuristics have directed attention away from detailed models of cognitive processes and toward post-hoc accounts of alleged errors. Gigerenzer suggests that in place of plausible heuristics that explain everything and nothing (not even the conditions which trigger one heuristic rather than another), models which make falsifiable predictions and are able to reveal the mental processes that explain both valid and invalid judgment are needed (see Gigerenzer, 1996; Gigerenzer, 2000).

Nevertheless Barberis and Thaler (2003) concede that there are numerous degrees of freedom, but note that rational modelers have just as many options to choose from. As Arrow (1986) has forcefully argued, rationality per se does not yield many predictions. The predictions in rational models often come from auxiliary assumptions.

In a recent review Hirshleifer (2001) states that many psychological biases can be viewed as outgrowths of heuristic simplification, self-deception, and emotion-based judgments. Heuristic simplification may explain many different documented biases, such as salience and availability effects (heavy focus on information that stands out or is often mentioned), framing effects (wherein the description of a situation affects judgments and choices), money illusion (wherein nominal prices affect perceptions), and mental accounting (tracking gains and losses relative to arbitrary reference points). Self-deception can explain overconfidence (a tendency to overestimate ones ability or judgment accuracy), and dynamic processes that support overconfidence such as biased self-attribution (a tendency to attribute successes to one’s own ability and failure to bad luck or other factors), confirmatory bias (a tendency to interpret evidence as consistent with one’s preexisting beliefs), hindsight bias (a tendency to think that you “knew it all along”), rationalization (straining to come up with arguments in favor of one’s past judgments and choices), and action-induced attitude changes of the sort that motivate cognitive dissonance theory (becoming more strongly persuaded of the validity of an action or belief as a direct consequence of adopting that action or belief; see Cooper and Fazio, 1984). Feeling or emotion-based judgments can explain mood effects (such as the effects of irrelevant environmental variables on optimism), certain kinds of attribution errors (attributing good mood to superior future life prospects rather than to immediate variables such as sunlight or a comfortable environment), and problems of self-control (such as difficulty in deferring immediate consumption and the effects of feelings such as fear on risky choices).

According to Barberis and Thaler (2003) economists are wary of this body of experimental evidence related to cognitive biases because they believe that

- people, through repetition, will learn their way out of biases

- experts in a field, such as traders in an investment bank, will make fewer errors
- with more powerful incentives, the effects will disappear.

While all these factors can weaken biases to some extent, there is little evidence that they wipe them out altogether. The effect of learning is often muted by errors of application: When the bias is explained, people often understand it, but then immediately proceed to violate it again in specific applications. Even expertise is often a burden rather than a help: Experts, armed with their sophisticated models, have been found to exhibit more overconfidence than laymen, particularly when they receive only limited feedback about their predictions. Finally, in a review of dozens of studies, Camerer and Hogarth (1999, p. 7) conclude that while incentives can sometimes reduce the biases people display, “no replicated study has made rationality violations disappear purely by raising incentives”.

2.3 Market Anomalies

Anomalies are empirical results that seem to be inconsistent with maintained theories of asset-pricing behavior. They indicate either market inefficiency (profit opportunities) or inadequacies in the underlying asset-pricing model. At a fundamental level, anomalies can only be defined relative to a model of “normal” return behavior. Fama (1970) noted this fact early on, pointing out that tests of market efficiency also jointly test a maintained hypothesis about equilibrium expected asset returns. Thus, whenever someone concludes that a finding seems to indicate market inefficiency, it may also represent evidence that the underlying asset-pricing model is inadequate (Schwert, 2003).

Nevertheless, in recent years a body of evidence on security returns has presented a sharp challenge to the traditional view that securities are rationally priced to reflect all publicly available information. According to Daniel, Hirshleifer, and Subrahmanyam (1998) some of the more pervasive anomalies can be classified as follows:

- Event-based return predictability, i.e. public-event-date average stock returns are of the same sign as average subsequent long-run abnormal performance (see for example Bernard and Thomas, 1989).

- Short-term momentum, i.e. positive short term autocorrelation of stock returns, for individual stocks and the market as whole (see for example Jegadeesh and Titman, 1993).
- Long-term reversal, i.e. negative autocorrelation of short-term returns separated by long lags or "overreaction" (see for example De Bondt and Thaler, 1985).
- High volatility of asset prices relative to fundamentals (see for example Shiller, 1981).
- Post-earnings announcement drift: Earnings announcements set off a stock price movement in the direction indicated by the earnings surprise in the short-run, but abnormal stock price performance drifts in the opposite direction of long-term earnings changes (see for example De Bondt, and Thaler, 1987; Lakonishok, Shleifer and Vishny, 1994).

Disagreement over the scientific interpretation of the above mentioned evidence remains. One possibility is that these anomalies are chance deviations to be expected under market efficiency (Fama, 1998). Daniel Hirshleifer, and Subrahmanyam (1998) state that the evidence does not comply with this viewpoint since some of the return patterns are strong and regular. Out-of-sample tests in time and location have established several of these patterns as regularities. For example the size, book-to-market, and momentum effects are present both internationally and in different time-periods. Furthermore, the patterns mentioned in event-based return predictability correspond with the results of the great majority of event studies.

Alternatively these patterns could represent variations in rational risk premia. Due to the high Sharp ratios (relative to the market), which are apparently achievable with simple trading strategies (MacKinlay, 1995), any asset pricing consistent with these patterns would have to have extremely variable marginal utility across states. Campbell and Cochrane (1999) find that a utility function with extreme habit persistence is required to explain the predictable variation in market returns. To stay consistent with cross-sectional predictability research on size, book-to-market, and momentum, a model would presumably require even more extreme variation in marginal utilities. Furthermore, the model would require that marginal utilities covary strongly with the returns on the size, book-to-market and momentum portfolios. No such correlation is obvious in examining the data. Given this evidence, it is more than reasonable to consider explanations for the observed return patterns based on bounded rationality (Daniel, Hirshleifer, and Subrahmanyam, 1998).

2.3.1 The Equity Premium and Myopic Loss Aversion

The stock market has historically earned a high excess rate of return. For example, using annual data from 1871-1993, Campbell and Cochrane (1999) document evidence that the average log return on the Standard and Poor’s 500 index is 3.9 percent higher than the average log return on a short-term commercial bond. This has been known as the equity premium puzzle since the work of Mehra and Prescott (1985). Although these facts are widely agreed on they remain a little controversial since research literature has argued that the equity premium is overstated due to survivorship bias (Brown, Goetzmann and Ross, 1995).

According to Barberis and Thaler (2003) the core of the equity premium puzzle is that even though stocks appear to be an attractive asset, they have high average returns and a low covariance with consumption growth. Thus investors appear very unwilling to hold them. In particular, they appear to demand a substantial risk premium in order to hold the market supply. Behavioral finance has approached this puzzle based on preferences, in particular with applying the predictions of prospect theory.

One of the earliest works which links prospect theory to the equity premium is Benartzi and Thaler (1995). They study how an investor with prospect theory type preferences allocates his financial wealth between treasury bills and the stock market. Prospect theory argues that when choosing between gambles, people compute the gains and losses for each one and select the one with the highest prospective utility. In a financial context, this suggests that people may choose a portfolio allocation by computing, for each allocation, the potential gains and losses in the value of their holdings, and then taking the allocation with the highest prospective utility.

In order to implement this model, Benartzi and Thaler specify how often investors evaluate their portfolios. To see why this matters Barberis and Thaler (2003) compare two investors: energetic Nick who calculates the gains and losses in his portfolio every day, and laid-back Dick who looks at his portfolio only once per decade. Since, on a daily basis, stocks go down in value almost as often as they go up, the loss aversion built into the value function υ (see section 2.2.3.1 on prospect theory) makes stocks appear unattractive to Nick. In contrast, loss aversion does not have much effect on Dick's perception of stocks since, at ten year horizons, stocks offer only a small risk of losing money.

Rather than simply picking an evaluation interval Benartzi and Thaler (1995) calculate how often investors would have to evaluate their portfolios to make them indifferent between stocks and treasury bills. This calculation can be thought of as asking what kind of equity premium might be sustainable in equilibrium: how often would investors need to evaluate their gains and losses so that even in the face of the large historical equity premium, they would still be happy to hold the market supply of treasury bills. Benartzi and Thaler find that for the parametric forms of the probability weighting function π and the value function υ estimated in experimental settings (see section 2.2.3.1) the answer is one year. They conclude that this is indeed a natural evaluation period for investors to use. The way people frame gains and losses is plausibly influenced by the way information is presented to them. Since we receive our most comprehensive mutual fund reports once a year, and do our taxes once a year, it is not unreasonable that gains and losses might be expressed as annual changes in value. The calculation of Benartzi and Thaler therefore suggests a simple way of understanding the high historical equity premium. If investors get utility from annual changes in financial wealth and are loss averse over these changes, their fear of a major drop in financial wealth will lead them to demand a high premium as compensation. Benartzi and Thaler (1995) call the combination of loss aversion and frequent evaluations myopic loss aversion.

2.3.2 The Volatility Puzzle

Stock returns and price-dividend ratios are both highly variable. In the same data set, the annual standard deviation of excess log returns on the Standard and Poor's 500 is 18 percent, while the annual standard deviation of the log price-dividend ratio is 0.27. These facts are called the volatility puzzle since they are hard to rationalize in a simple consumption-based model (Campbell, 1999).

To understand the volatility puzzle, Barberis and Thaler (2003) note that in a simple economy, both discount rates and expected dividend growth are constant over time. A direct application of the present value formula implies that the price-dividend (P/D) ratio is also constant. The standard deviation of log returns will therefore only be as high as the standard deviation of log dividend growth, namely 12 percent. The particular volatility puzzle illustrates a more general point, first made by Shiller (1981) and LeRoy and Porter (1981), namely that it is difficult to explain the historical volatility of stock returns with any model in which investors are rational and discount rates are constant. Since the volatility of log dividend growth is only 12 percent, the only way for a model to generate an 18 percent volatility of log returns is to introduce variation in the P/D ratio. But if discount rates are constant, the present-value formula shows that the only way to do that is to introduce variation in investors' forecasts of the dividend growth rate: A higher forecast raises the P/D ratio, a lower forecast brings it down. If investors are rational, their expectations for dividend growth must, on average, be confirmed. In other words, times of higher (lower) P/D ratios should, on average, be followed by higher (lower) cash-flow growth. Unfortunately, price-dividend ratios are not reliable forecasters of dividend growth, neither in the USA nor in most international markets (Campbell, 1999).

Shiller's and LeRoy and Porter's results shocked the profession when they first appeared. At the time, most economists felt that discount rates were close to constant over time, apparently implying that stock market volatility could only be fully explained by appealing to investor irrationality. Today, it is well understood that rational variation in discount rates can help explain the volatility puzzle, although models with irrational beliefs also offer a plausible way of thinking about the evidence (Barberis and Thaler, 2003).

For example Barberis, Huang and Santos (2001) show that their model can explain both the equity premium and volatility puzzles based on bounded rational preferences. They appeal to experimental evidence about dynamic aspects of loss aversion. This evidence suggests that the degree of loss aversion is not the same in all circumstances but depends on prior gains and losses. In particular, Thaler and Johnson (1990) find that after prior gains, subjects take on gambles they normally do not, and that after prior losses, they refuse gambles that they normally accept. The first finding is sometimes known as the "house money effect", reflecting gamblers' increasing willingness to bet when they are ahead. One interpretation of this evidence is that losses are less painful after prior gains because they are cushioned by those gains. Nevertheless, after being burned by a painful loss, people may become more wary of additional setbacks.

To capture these ideas, Barberis, Huang and Santos (2001) modify the utility function in their model, that the investors' sensitivity to losses is no longer constant, but is determined in a way that reflects the experimental evidence described above. A model of this kind can help to explain the volatility puzzle. For example, if one supposes that there are positive cash flow news, these news push the stock market up, generating prior gains for investors, who are now less scared of stocks: any losses will be cushioned by the accumulated gains. As a consequence investors discount future cash flows at a lower rate, pushing prices up still further relative to current dividends and adding to return volatility (Barberis and Thaler, 2003).

2.3.3 Predictability in Returns

"The reaction of one man can be forecast by no known mathematics;
the reaction of a billion is something else again."
Isaac Asimov

Stock returns are forecastable. By using monthly, real, equal-weighted New York Stock Exchange returns from 1941-1986, Fama and French (1988) are able to show that the dividend-price ratio (see section 2.3.2) can explain 27 percent of the variation of cumulative stock returns over the subsequent four years. Although these facts are widely agreed on they are not uncontroversial since a large body of literature has debated the statistical significance of a time series' predictability.

The predictability puzzle is closely related to the volatility puzzle (see section 2.3.2), since in any model with a stationary P/D ratio, a resolution of the volatility puzzle is simultaneously a resolution of the predictability puzzle. This holds because any model which captures the empirical volatility of returns must involve variation in the P/D ratio. Moreover, for a model to be a satisfactory resolution of the volatility puzzle, it should not make the counterfactual prediction that P/D ratios forecast subsequent dividend growth. Now suppose that the P/D ratio is higher than average. The only way it can return to its mean is if the cash flows of dividends (D) subsequently go up, or if prices (P) fall. Since the P/D ratio is not allowed to forecast cash flows, it must forecast lower returns, thereby explaining the predictability puzzle (Barberis and Thaler, 2003).

2.3.4 Herding Behavior

> *"Fashion is the great governor of this world; it presides not only in matters of dress and amusement, but in law, physics, politics, religion, and all other things of the gravest kind; indeed, the wisest of men would be puzzled to give any better reason why particular forms in all these have been at certain times universally received, and at others universally rejected, than that they were in or out of fashion."*
>
> Henry Fielding

Certain classes of investors and their agents change their behavior in parallel. This phenomenon, called herding, is consistent with rational responses to new information, agency problems or conformity bias. According to Daniel, Hirshleifer, and Teoh (2002) herding behavior has been documented in the trading decisions of institutional investors (see for example Grinblatt, Titman, and Wermers, 1995), in recommendation decisions of stock analysts (Welch, 2000), and in investment newsletters (Graham, 1999). The tendency of analysts to follow the prevailing consensus is not stronger when that consensus proves to be correct than when it proves to be wrong (Welch, 2000).

Devenow and Welch (1996) state that imitation and mimicry are among our most basic instincts. Herding can be found in fashion and fads, just as in such simple decisions as how to best commute and what research to work on. There is a prominent belief, not only among practitioners but also among financial economists, that investors are influenced by the decisions of other investors and that this influence is a first-order effect.

Devenow and Welch argue that it is difficult to precisely define "herding". In its most general form, herding could be defined as behavior patterns that are correlated across individuals. On the other hand, if many investors are purchasing "hot stocks", it could just be due to correlated information arrival in independently acting investors. The notion of herding which Devenow and Welch focus on is one which can lead to systematic and erroneous (i.e. sub-optimal relative to the best aggregate choice) decision-making by entire populations. In this sense, herding is closely linked to such distinct phenomena as imperfect expectations, inconsistent changes without much new information, bubbles, fads, and frenzies. Furthermore herding does require a coordination mechanism. This mechanism can be either a widely spread rule to coordinate based on some signal (for example a price movement), or based on a direct ability to observe other decision-makers (for example observing a colleague's investments).

There are two polar views of herding: the non-rational and the rational views. The non-rational views focus on investor psychology and state that agents behave like lemmings, following one another blindly and foregoing rational analysis. Less crazy investors are assumed to be able to profit generously from them. The rational views focus on externalities, optimal decision making being distorted by information difficulties, or incentive issues. The intermediate view holds that decision makers are near rational, economizing on information processing or on information acquisition costs by using heuristics, and that rational activities by third-parties cannot eliminate this influence (Devenow and Welch, 1996).

A psychological and neurological explanation is provided by Prechter (2001). He states that human herding behavior results from impulsive mental activity in individuals responding to signals from the behavior of others. Impulsive thought originates in the basal ganglia and limbic system. In emotionally charged situations, the limbic system's impulses are typically faster than rational reflection performed by the neocortex. Experiments with a small number of naive individuals as well as statistics reflecting the behavior of large groups of financial professionals provide evidence of herding behavior. Herding behavior, while appropriate in some primitive life-threatening situations, is inappropriate and counterproductive for success in financial situations. Unconscious impulses that evolved in order to attain positive values and avoid negative values encourage herding behavior, making rational independence extremely difficult to exercise in group settings. A negative feedback loop develops because stress increases impulsive mental activity, and impulsive mental activity in financial situations, by inducing failure, increases stress. The interaction of many minds in a collective setting produces macro behavior that is patterned according to the survival-related functions of the primitive portions of the brain. As long as the human mind comprises the construction and its functions, patterns of herding behavior will remain immutable. This is the psychological basis of financial market trends and patterns and may be exploited by momentum based trading strategies.

Furthermore, Hirshleifer and Teoh (2003) assume that there are many patterns of convergent behavior and fluctuations in capital markets that do not obviously make immediate sense in terms of traditional economic models. They provide examples such as fixation on poor projects, stock market crashes, sharp shifts in investment and unemployment, bank runs. Such behavioral convergence often appears even in the face of negative payoff externalities. Although other factors, such as payoff externalities, can lock-in inefficient behaviors, the rational social learning theory and especially cascades theory differ in that they imply pervasive but fragile herd behavior. This occurs because the accumulation of public information slows down or blocks the generation and revelation of further information. This idiosyncratic feature of cascades and rational observational learning models cause the social equilibrium to be unstable with respect to seemingly modest new shocks.

According to Hirshleifer and Teoh (2003) rational observational learning theory suggests that in many situations, even if payoffs are independent and people are rational, decisions tend to converge quickly but tend to be idiosyncratic and fragile. Convergence arises locally or temporally upon a behavior, and can suddenly shift into convergence on the opposite behavior. The required assumptions, primarily discreteness or boundedness of possible action choices, are mild and likely to be present in many realistic settings. This suggests that the effects of observational learning and herding are likely to affect behavior in and related to capital markets. This includes both herding by firms and actions by firms such as financing, disclosure and reporting policies that can potentially be managed to exploit investors that herd. Similarly, perhaps the special skill that some hedge fund and mutual fund managers seem to have is in exploiting the herding behavior of imperfectly rational investors.

Models of reputation-based herding do not typically share the fragility feature of rational observational learning theory. Reputation-based models have much to offer. This includes explanation of those herds that seem stable and robust. As another example, the reputation approach helps explain dispersion as well as herding, and when one or the other will occur. Reputation models also offer a rich set of implications about the extent of herding in relation to characteristics of the agency problem and the manager. In most instances herding in capital markets likely involve mixtures of reputational effects, informational effects, direct payoff interactions, preference effects, and imperfect rationality. For example, to explain predictability in securities markets, some imperfect rationality is likely to be needed. Integration of the different effects will lead to better theories of capital market behavior (Hirshleifer and Teoh, 2003).

2.3.5 The Shaping Hypothesis

There is some evidence that specific anomalies become less frequent in repeated experimental markets. Some of this evidence shows a particularly interesting feature: anomalies are eroded when individuals' preferences or valuations are elicited in repeated markets, but not when they are elicited by other mechanisms including repetition, incentives and feedback (Cox and Grether, 1996; Shogren et al., 2001).

According to Loomes, Starmer, and Sugden (2003) one interpretation is that individuals, particularly marginal traders, are learning to act on underlying preferences which satisfy standard assumptions. The mechanisms to explain this type of behavior are the refining and market discipline hypotheses. Both theories assume that each agent has true preferences that are independent of the mechanisms via which they are revealed or elicited. Both hypotheses assume that a repeated elicitation mechanism filters some or all extraneous errors and biases without affecting the preferences themselves. The assumption that preferences are "mechanism independent" is crucial if such hypotheses are to justify conventional economic theory as an explanation of behavior in real repeated markets.

An alternative interpretation, the "shaping" hypothesis, states that individuals' preferences are adjusting in response to cues given by market prices. The shaping hypothesis says that, in repeated auctions in which prices represent no informational content, there is a tendency for agents to adjust their bids towards the price observed in the previous market period. If there is some element of common value in an auction, such an adjustment rule may be consistent with Bayesian updating of agents' beliefs about the value of the good for which they are bidding. But the shaping hypothesis also applies to cases in which values are entirely private. The intuition behind the hypothesis is that, prior to her involvement in a specific market, an agent may not have well-articulated preferences waiting to be "discovered". Instead, values may be only partially formulated and/or imprecise, so that when confronted by an elicitation mechanism, responses are generated using heuristics in which market prices act as cues.

The experimental results of Loomes, Starmer, and Sugden suggest that systematic shaping effects do occur. If behavior in markets is indeed influenced by shaping, the validity of repeated market mechanisms as means of eliciting individuals' preferences is questioned. If such mechanisms are to be used for this purpose, they need to be designed in the light of an understanding of the dynamics of shaping. It is possible that shaping, like loss aversion, is itself a bias which market experience eventually eliminates. But Loomes, Starmer, and Sugden's results provide grounds for scepticism on this interpretation. Their conjecture is that shaping is associated with preference imprecision: the less sure a person is about what her preferences really are, the more susceptible she is to external cues such as information about market prices. If that is right, one should expect any erosion of shaping effects to be associated with a reduction in the stochastic component of individuals' preferences. But Loomes, Starmer, and Sugden's results give little support to the idea that preference imprecision declines with market experience.

2.3.6 Discussion

Puzzling questions emerge if researchers investigate the aggregate stock market and the aggregate market participants' behavior. While the behavior of the aggregate stock market is not easy to understand from the rational point of view, promising rational models have nonetheless been developed and can be tested against behavioral alternatives. Empirical studies of the behavior of individual stocks have discovered a set of facts that are altogether more frustrating for the rational paradigm. Many of these facts are about the cross-section of average returns: They document that one group of stocks earns higher average returns than another. These facts have come to be known as anomalies because they cannot be explained by the even simplest and most intuitive model of risk and return in the financial economist's toolkit, the Capital Asset Pricing Model. According to Barberis and Thaler (2003) three of the most striking market anomalies are:

- The Equity Premium, i.e. the stock market has historically earned a high excess rate of return (see Mehra and Prescott, 1985).

- High volatility of asset prices relative to fundamentals (see for example Shiller, 1981).
- Predictability of stock returns (see for example Fama and French, 1988).

All three of these facts can be labelled puzzles since they are hard to rationalize in a simple consumption-based model. The first fact has been known as the equity premium puzzle since the work of Mehra and Prescott (1985). Campbell (1999) calls the second fact the volatility puzzle, and Barberis and Thaler (2003) refer to the third fact as the predictability puzzle. Although these facts are widely agreed on, they are not completely uncontroversial. For example, a vast literature has debated the statistical significance of the time series predictability, while others have argued that the equity premium is overstated due to survivorship bias (see for example Brown, Goetzmann and Ross, 1995).

Both the rational and behavioral approaches to finance have made progress in understanding these puzzles (Barberis and Thaler, 2003).

2.4 Trading Strategies

The next day they passed the same place and the one economist said,
“See! I told you there was no $100 there!”

Is it possible to profit on psychological biases or the resulting financial market anomalies? While early economic literature suggests that financial markets are efficient, more recent evidence supports the view that markets can not be fully efficient because of the cost of collecting and analyzing information, the cost of trading, and limits on the capital available to arbitrageurs. Nowadays both academics and practitioners share the view that pockets of inefficiency exist within broad market efficiency.

Singal (2004) provides a comprehensive review of related research and gives insights on successful application of profitable trading strategies. He states that most but not all anomalies are expected to generate tradable profits. Furthermore, if generating profits is not possible, information about an anomaly will help the practitioner to better understand the mispricing and modify her trading behavior to avoid being hurt by the negative effects of the market anomaly.

The next sections will review promising approaches to trading strategies in financial markets.

2.4.1 Momentum and Contrarian Investing

Simple trading strategies have attracted attention since the early days of stock exchanges. Probably the most obvious strategies are trading strategies based on the past return patterns of stocks. Many of these are about the cross-sectional patterns of average returns: They document that one group of stocks earns higher average returns than another. These cross-sectional patterns can be exploited by momentum or contrarian strategies, depending on return continuation or reversals in the subsequent investment horizon. A momentum (contrarian) strategy is based on a simple rule: buy stocks that performed best (worst) and sell stocks that performed worst (best) in the recent past (Swinkels, 2004).

Every month from January 1963 to December 1989, Jegadeesh and Titman (1993) group all stocks traded on the New York Stock Exchange into deciles based on their prior six month return and compute average returns of each decile over the six months after portfolio formation. They find that the decile of biggest prior winners outperforms the decile of biggest prior losers by an average of 10 percent on an annual basis.

On the other hand, De Bondt and Thaler (1985) report on long-term reversals. For every three years from 1926 to 1982, they rank all stocks traded on the New York Stock Exchange by their prior three year cumulative return and form two portfolios: a “winner” portfolio of the 35 stocks with the best prior record and a “loser” portfolio of the 35 worst performers. Then they measure the average return of these two portfolios over the three years subsequent to their formation. They find that over the whole sample period, the average annual return of the loser portfolio is higher than the average return of the winner portfolio by almost 8 percent per year. De Bondt and Thaler (1985) argue that results from Kahneman's and Tversky's (1974) research on judgment under uncertainty, in particular the representativeness heuristic (see section 2.2.1.2), could explain overreaction in financial markets. Thus contrarian strategies may be an appropriate way to exploit this type of anomaly (“buy losers, sell winners”). Lo and MacKinlay (1990) provide evidence against overreaction as the *only* source of contrarian profits. Their argument is based on correlations across stocks and the fact that returns of large stocks lead those of small stocks. They further argue that if returns on some stocks systematically lead or lag those of others, a portfolio strategy that sells winners and buys losers can produce positive expected returns, even if stock returns are not negatively autocorrelated as models of overreaction imply.

Comparing the results of Jegadeesh and Titman (1993) to De Bondt and Thaler's (1985) study of prior winners and losers illustrates the crucial role of the length of the prior ranking period. In one case, prior winners continue to win, while in the other, they perform poorly. A challenge to both behavioral and rational approaches is to explain why an extension of the formation period switches the results in this way.

According to Barberis and Thaler (2003) there is some evidence that tax-loss selling creates seasonal variation in the momentum effect. Stocks with poor performance during the year may later be subject to selling by investors keen to realize losses that can offset capital gains elsewhere. This selling pressure means that prior losers continue to lose, enhancing the momentum effect. At the turn of the year, the selling pressure eases off, allowing prior losers to rebound and to weaken the momentum effect. A careful analysis by Grinblatt and Moskowitz (1999) finds that on net, tax-loss selling may explain part of the momentum effect, but by no means all of it. In any case, while selling a stock for tax purposes is rational, a model of predictable price movements based on such behavior is not. Roll (1983) calls such explanations "stupid" since investors would have to be stupid not to buy in December if prices were going to increase in January.

Momentum is stronger in small than in large firms (Jegadeesh and Titman, 1993; Grinblatt and Moskowitz, 1999), in growth than in value firms (Daniel and Titman, 1999), and in firms with low rather than high analyst following (Hong, Lim and Stein, 2000). These tendencies are potentially consistent with limits to attention, reducing the extent to which investors are able to take advantage of momentum. Also, it suggests that smart investors may be more deterred by transaction costs than foolish investors. Both industry and non-industry components of momentum help to predict future returns (Grundy and Martin, 2001; Moskowitz and Grinblatt, 1999). Moskowitz and Grinblatt find that the profitability of industry momentum comes mainly from winners, but the profitability of individual stock momentum strategies is stronger for losers. As already mentioned, at long horizons momentum reverses (see De Bondt and Thaler, 1985).

In their recent article Cooper et al. (2004) test overreaction theories of short-run momentum and long-run reversal in the cross section of stock returns. According to them, momentum profits depend on the state of the market, as predicted. From 1929 to 1995, the mean monthly momentum profit following positive market returns is 0.93 percent, whereas the mean profit following negative market returns is 0.37 percent. The up-market momentum reverses in the long-run. Their results are robust to the conditioning information in macroeconomic factors. Moreover, they find that macroeconomic factors are unable to explain momentum profits after simple methodological adjustments to take account of microstructure concerns.

2.4.2 The Daylight Saving Anomaly and a SAD Stock Market Cycle

There is evidence that environmental factors that influence mood are correlated with stock price movements (see for example Hirshleifer, 2001). For instance, a stochastic variable, cloud cover in New York, can be associated with low daily U.S. stock market returns (Saunders, 1993). Hirshleifer and Shumway (2003) provide evidence that sunny weather is associated with higher stock returns. Their paper examines the relationship between morning sunshine in the city of a country's leading stock exchange and daily market index returns across 26 countries from 1982 to 1997. Sunshine is strongly significantly correlated with stock returns (the coefficient of the simple pooled regression equals -0.011, with a t-statistic of -4.49). After controlling for sunshine, rain and snow are unrelated to returns. Substantial use of weather-based strategies is optimal for a trader with very low transactions costs. Because these strategies involve frequent trades, the gains are fairly eliminated by modest costs.

Kamstra, Kramer, and Levi (2000) find that a deterministic variable, changes to and from daylight savings time, which disrupts sleep, is related to stock returns. Empirical studies show that weekend effects (a significantly negative return from Friday closing to Monday opening prices) are also reflected in stock market returns. In addition there is empirical evidence for an even stronger negative effect on stock market returns on spring and fall daylight-saving weekends. Kamstra, Kramer, and Levi report that the magnitude of the daylight-saving effect is roughly 200 to 500 percent of the regular weekend effect, which makes it both statistically and economically significant. Furthermore the effect applies in several international financial markets. A possible strategy could be to go short or sell on Fridays and buy on Mondays, in particular at these two special weekends in spring and fall (see Kamstra, Kramer, and Levi, 2000 and 2002). Singal (2004) suggests an abnormal return based on past evidence of the weekend effect of about 0.20 percent per weekend. He states that it is not easy to capture the "normal" weekend effect with current financial instruments because the trading costs can be large. Singal rather recommends investors to change their trading patterns accordingly. Thus investors should recognize the weekend effect and avoid buying stocks on Fridays and selling on Mondays.

According to Kamstra, Kramer, and Levi (2003) the seasonal variation in the length of day (daylight exposure) can influence the mood of people (clinically diagnosed as seasonal affective disorder or SAD, meaning a form of depression) and therefore affect the risk taking behavior of market participants (increased risk aversion with decreased daylight). This translates into the empirical supported evidence of seasonal variation of equity returns. A trading strategy could be to buy in months with relatively low returns (people are risk averse) and sell in months with higher returns (people have recovered). Instead one could alternately invest in stock markets located in different hemispheres (long during northern fall and winter then transfer the money to a southern market during southern fall and winter). The strategy should also consider a possible asymmetric effect between fall and winter which results in lower returns in fall and higher returns in winter months. Kamstra, Kramer, and Levi provide an example of a trading strategy. They select the countries Sweden and Australia, because each is one of the most extremely located in its hemisphere. As a benchmark they use a "neutral" portfolio allocation strategy in which investors place 50 percent of their portfolio in the Swedish index and 50 percent in the Australian index. From the early 1980's up to 2003, the average annual return to this neutral strategy equals 13.20 percent. On the contrary, a SAD portfolio allocation strategy, in which the investor reallocates 100 percent of her portfolio twice a year at fall and spring equinox, placing her money in the Swedish market during the Northern Hemisphere's fall and winter, then moving it into the Australian market for the Southern Hemisphere's fall and winter, would lead to an amazing average annual return of 21.10 percent.

2.4.3 Can Ignorance Beat the Stock Market?

Borges, Goldstein, Ortmann, and Gigerenzer (1999) use the recognition heuristic (see section 2.2.1.7) as a device for selecting stock portfolios in a bullish market environment. Furthermore, this represents an example for an integrated markets based trading strategy, i.e. a strategy that involves both the consumer and the financial market. The empirical study, as described by Borges et al., introduces the notion of a buy and hold portfolio strategy that relies on the ignorance-based cognitive decision making mechanism, called the recognition heuristics. They use corporate name recognition for selecting a subset of stocks from Standard and Poor’s 500 index in a bullish market environment. Borges et al. compared the performance of a portfolio constructed only of stocks from companies with a high level of name recognition by either laypeople (pedestrians) or experts with several benchmarks (mutual funds, market indices, “dartboard” portfolios and unrecognized stocks). The portfolio, consisting solely of stocks from companies recognized by laypeople, unexpectedly outperformed its touchstones and generated striking returns (the ten German stocks most recognized by American laypeople outperformed the market by 23 percent).

Boyd (2001) presents a study that replicates recent tests of the recognition heuristic as a device for selecting stock portfolios. The heuristic represents a lower limit to the search for information, since simple name recognition is the least one can know about anything. Gigerenzer and others conducted original experiments in this field at the Max Planck Institute for Human Development's Center for Adaptive Behavior and Cognition (the "ABC Research Group"). The ABC Group's tests support the use of the heuristic in a bull market environment. Boyd’s study, conducted in a down market, reaches a different conclusion: Not only can a high degree of company name recognition lead to disappointing investment results in a bear market, it can also be beaten by pure ignorance. Virtually the only finding of the ABC Group's study that Boyd can match is that Americans are not very good at picking American stocks to outperform the market.

2.4.4 Discussion

Ritter (2003) concludes that it is very difficult to find trading strategies that reliably make money. But this does not imply that financial markets are informationally efficient since low-frequency misvaluations may be large, without presenting any opportunity to make money. Furthermore the forces of arbitrage, which work well for high-frequency events, may work very poorly for low-frequency events.

Discovered market anomalies must be viewed with caution and scepticism, as spurious mispricings can surface for a variety of reasons, such as errors in defining normal return, data mining, survivorship bias, small sample bias, selection bias, nonsynchronous trading, and misestimation of risk. Although anomalies should disappear in a close to efficient market, they may persist because they are not well understood, arbitrage is too costly, the profit potential is insufficient, trading restrictions exist, and behavioral biases exist (see also chapter 2.1 on limits to arbitrage). Documented and valid anomalies may still be unprofitable because the evidence is based on averages and may therefore include a large fraction of losers. Furthermore, the conditions responsible for the anomaly may change, and trading by informed investors may cause the anomaly to disappear (Singal, 2004).

Nevertheless Singal (2004) provides profitable trading strategies. He states that if generating profits is not possible, information about an anomaly will help the practitioner to better understand the mispricing and modify her trading behavior to avoid being hurt by the negative effects of the market anomaly.

3 Agent-Based Computational Economics

Agent-based computational economics is the study of computationally simulated economies modeled as evolving systems of autonomous interacting agents. The systems are composed of self-contained actors, who interact according to a fixed set of rules. Starting from initial conditions, specified by the modeler, the computational economy evolves over time as its constituent agents repeatedly interact with each other and learn from these interactions. Agent-based computational economics is therefore a bottom-up approach to the study of economic systems (Tesfatsion, 2003).

Conventional models of financial markets based on assumptions of rational choice and market efficiency are extremely elegant in form. Nevertheless, no single standard model to date has proven to be capable of explaining the basic empirical features of real financial markets, including fat-tailed asset return distributions, high trading volumes, persistence and clustering in asset return volatility, and cross-correlations between asset returns, trading volume, and volatility. Due to these well known difficulties, financial markets have become one of the most active research areas for agent-based computational economic modelers. Agent-based computational financial market models have been able to provide possible explanations for a variety of observed regularities in financial data (see for example Lux, 1998; Farmer and Lo, 1999; Lux and Marchesi, 2000; Hommes, 2002). Several of the earliest agent-based financial market studies are surveyed in detail in LeBaron (2000), including the highly influential Santa Fe artificial stock market study by Arthur et al. (1997), who developed a dynamic theory of asset pricing based on heterogeneous stock market traders updating their price expectations individually and inductively by means of classifier systems (Tesfatsion, 2003).

Agent-based systems address phenomena that are generated through individual interactions, rather than aggregate behavior (Tesfatsion, 2002). One branch of agent-based computational economics is concerned with using simulated economic systems as laboratories in which economic theory can be tested. These economic laboratories occupy a niche between analytic theoretical models and empirical research. Artificial economic systems are typically more complex than allowed by analytic theory, but simpler than real systems. They therefore provide the opportunity to test theories in a more realistic setting than closed-form analytic models, while retaining the ability to examine and understand the resultant behavior.

Another branch of agent-based computational economics focuses on understanding the emergence of global behaviors based on local interactions. While global behaviors can be observed in empirical data, it is difficult to conclusively show the reason why such behaviors occur. If the same behaviors are shown to exist in a simpler simulated system, then at a minimum it can be said that the actors included in the simulation are sufficient to induce the observed behavior.

These two branches of agent-based computational economics are not exclusive. The generation of known global behaviors is an important way of validating a simulation study. A simulation that can not replicate known global behaviors in the domain of interest cannot be trusted in studying new behaviors. Similarly, replication of known global behaviors gives the modeler some confidence that the dynamics inherent in the simulation are reasonable. Of course, one must still be cautious in the subsequent exploration of new phenomena.

A central feature of agent-based economic simulations is the ability to include a mixture of agent types. Global behaviors caused by interaction between heterogenous agents are beyond the scope of many analytic models, and constitute a major area of interest for agent-based computational economics models. Another significant feature of agent-based models is the ability to explicitly model “boundedly rational” agents (Simon, 1982). These are agents with explicit limitations on their memory, knowledge or computational abilities. Finally, agent-based models can simulate long-term effects such as learning and adaptation which are difficult to include in analytic models.

It is important to emphasize that agent-based technologies are well suited for testing behavioral theories. They can answer two key questions that should be asked of any behavioral structure. First, how well do behavioral biases hold up under aggregation, and second, which types of biases will survive in a co-evolutionary struggle against others. Therefore, the connections between agent-based approaches and behavioral approaches will probably become more intertwined as both fields progress (LeBaron, 2004).

4 Social Consumer Agents in an Integrated Markets Model

Neoclassical economic theory is based on the assumption of rationally acting individuals, who are able to consider all available information in the decision-making process. As an early critic of economic agents with unlimited information processing capabilities, Herbert Simon (1955, 1982) suggested the term “bounded rationality” to describe a more realistic approach to cover human problem solving. Indeed, the complexity of human behavior suggests that a choice model should explicitly capture uncertainty. Real economic agents are restricted at least in their cognitive (knowledge) and computational abilities (Mullainathan and Thaler, 2000).

Enriched by a social network perspective, which states that most behaviors are also closely embedded in networks of interpersonal relations, an additional focus lies in the relationships among interacting units. According to Wassermann and Faust (1994) a social network is a set of people or groups of people (“actors” or agents) with certain pattern of interactions (“ties”) between them. Central concepts are:

- actors and their actions are viewed as interdependent
- relationships among actors are channels for transfer of resources
- the network structure provides constraints and opportunities for individual action
- lasting patterns of relations are conceptualized as structure.

Recent work on social networks has focused on distinctive features of network structure (Newman et al., 2002). One of these is the “small world” effect first described by Milgram (1967). His experiment involved letters that were passed between pairs of apparently distant people. Milgram found that the typical chain from acquaintance to acquaintance only has a length of about six persons (popularly known as “Six Degrees of Separation”). Since then dozens of academic studies have revealed that many networks have related “small-world” properties (see for example Watts and Strogatz, 1998). Usually the topology of a (social) network is assumed to be either completely regular or completely random. However, many biological, technological and social networks lie between these two extremes. These systems are highly clustered, like a regular lattice, but have small path lengths, like random graphs, and are named “small-world” networks. From a social systems perspective this means that it only takes a small number of well-connected people to make a world small (Collins and Chow, 1998).

In this chapter an agent-based computational economic model, which incorporates boundedly rational agents embedded in a social network structure is introduced. Computational economic models bridge the gap between theoretical and empirical economics. They can represent a testbed which enables one to investigate the predictions of a theory under conditions which are too complex to be addressed analytically. Hence computational models can be used to gain insights into complex systems and furthermore suggest new hypotheses to be tested in empirical studies (for a review of agent-based computational economics see Tesfatsion, 2002).

A considerably extended version of the integrated markets model, introduced by Sallans et al. (2002, 2003), is presented. The model spans two markets: a consumer market and a financial equities market. The consumer market consists of production firms offering a good for sale, and customer agents who can purchase the good. The financial equities market consists of stock traders who can buy and sell shares in the production firms. The new model focuses on a more life-like model of consumer agents. The new agents are embedded in a social structure based on “small-world” principles and incorporate an enhanced cognitive decision structure related to the consumat approach presented by Janssen and Jager (2000). Since in real life people do not behave in a systematic manner (see for example Gintis, 2000) a rational agent approach can not account for behavioral dynamics such as habits, imitation and social comparison. To explore how such behavioral dynamics affect the evolution of an economic system, it is practical to apply a more sophisticated approach in the integrated modeling context. The main contribution of this approach is that it increases the psychological richness and possibilities of validation of the simulated behavioral dynamics since it introduces behavioral rules based on a conceptual meta-model of behavior. This will take account of certain types of behavior like imitation, social comparison and market dynamics like lock-in, loyalty and bandwagon or snob effects.

To evaluate the fruitfulness of the new approach it is useful to compare the model’s output (macro level, for example a firm agent’s market share) to known “stylized facts” in consumer and financial markets. Stylized facts are robust empirical phenomena, which characterize market dynamics (for example market anomalies) and have been observed in real markets. For the validation of the integrated markets model a well known and accepted stylized fact found in consumer markets, the Bass curve, is implemented. The Bass curve is described by the Bass diffusion model and was introduced by Frank M. Bass (1969) in his now classic paper.

The Bass Model summarizes in a simple mathematical form the key finding from over 4,000 diffusion studies: most people wait until they have witnessed peers having favorable experiences with a new technology or service before they adopt. The original Bass model makes adoption a function of innovation and imitation effects. For example most people are influenced by word of mouth or advertising. The effects of interpersonal communication in particular are thought to be a key factor for the speed and shape of the diffusion of an innovation (Rogers, 1983; Mahajan et al., 1990). Another explanation might be the bandwagon effect or herding behavior (section 2.3.4) since conspicuous consumption gives rise to a conformistic behavior (Leibenstein, 1950).

While such theories are not easy to implement in a neoclassical rational economic framework, the integrated markets model represents the ideal environment to analyze the Bass model and its complex underlying mechanisms which are based on an agent's bounded rational and social behavior.

4.1 The Integrated Markets Model

The model consists of two interacting markets, a consumer and a financial equities market. The consumer market simulates the manufacture of a product by *production firms*, and the purchase of the product by *consumers*. The financial market simulates trading of shares. The shares are bought and sold by *financial traders*. The two markets are coupled: The financial traders buy and sell shares in the production firms, and the managers of firms are concerned with their share price. The traders can use the performance of a firm in the consumer market in order to make trading decisions. Similarly, the production firms can potentially use positioning in product space and pricing to influence the decisions of financial traders (see figure 5).

The simulator runs in discrete time steps. Simulation steps consist of the following operations:

- Consumers make purchase decisions
- Firms receive an income based on their sales and their position in product space
- Financial traders make buy/hold/sell decisions. Share prices are set and the market is cleared
- Every N_p steps, production firms update their products or pricing policies based on performance in previous iterations

The integrated markets model is intended to be a generic model of the interaction between financial and consumer markets. It has been shown to reproduce a large range of empirical “stylized facts” including learning-by doing in the consumer market; low predictability, high kurtosis and volatility clustering; and correlations between volatility and trading volume in the financial market.

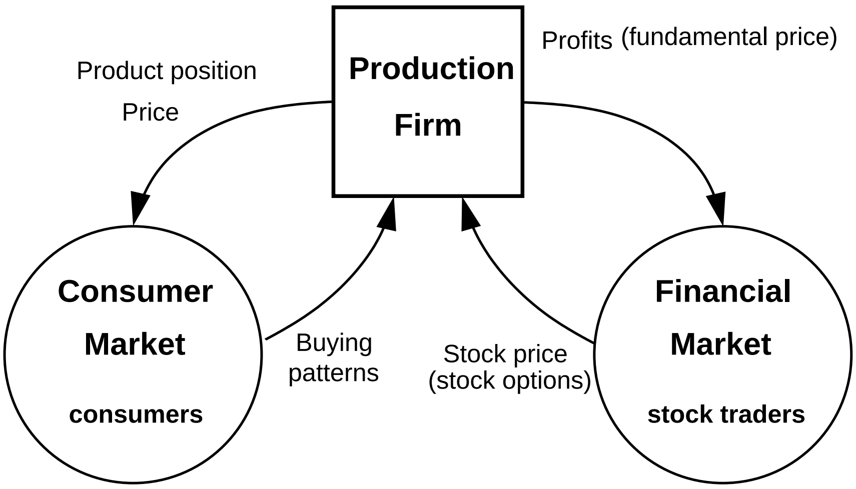


Figure 5: The Integrated Markets Model. Consumers purchase products, and financial traders buy and sell shares. Production firms link the consumer and financial markets, by selling products to consumers and offering their shares in the financial equities market (from Sallans et al., 2003).

4.1.1 The Consumer Market

The consumer market consists of firms which manufacture products, and consumers who purchase them. The consumers will re-purchase at regular intervals. The product space is represented as a two-dimensional simplex, with product features represented as real numbers in the range [0, 1]. Each firm manufactures a single product, represented by a point in this two-dimensional space. Consumers have fixed preferences about what kind of product they would like to purchase. Consumer preferences (individual needs) are also represented in the two-dimensional product feature space. There is no distinction between product features and consumer perceptions of those features. Each consumer agent is embedded in a social structure which influences its social needs and incorporates a cognitive decision structure which accounts for its committed behavior (repetition, imitation, social comparison, deliberation). Consumer agents react to their individual needs, social needs and the price of the produced products. Details of the consumer agents are described in section 4.2.

4.1.2 Production Firms

The production firms are adaptive learning agents. They adapt to consumer preferences and changing market conditions via a reinforcement learning algorithm (Sutton and Barto, 1998). In each iteration of the simulation the firms must examine market conditions and their own performance in the previous iteration, and then modify their product or pricing. A boundedly rational agent can be subject to several kinds of limitations. These limits manifest themselves in the firm's representation of its environment and its knowledge of its competitors. The firms do not have complete information about the environment in which they operate. In particular, they do not have direct access to consumer preferences. They must infer what the consumers want by observing what they purchase. Purchase information is summarized by performing “k-means” clustering on consumer purchases. K-means is a common clustering technique used in consumer market research. The number of cluster centers (N) is fixed at the start of the simulation. The current state information consists of the positions of the cluster centers in feature space, along with additional state information such as whether or not the previous action was profitable or boosted stock price, and where the competitors products are located. This information gives a summary of the environment at the current time step.

Firms make decisions based on a finite history of states of some length. This limited history window represents an additional explicit limit on the firm's knowledge. In each iteration the firms can take one of several actions. The actions include taking a random action, doing nothing, raising or lowering product price, or moving the product in feature space. The random action was included to allow the firm to explicitly choose to take a “risky” exploratory action. A firm's manager seeks to modify its behavior so as to maximize an external reward signal. This reward signal can be viewed as the managers compensation for its actions. The reward signal takes the form of a fixed reward, a variable amount based on the firm’s profitability, and a variable amount due to change in the value of the firms stock. The constant part of the reward signal can be interpreted as a fixed salary paid to the manager of the firm. The profit-based reward can be interpreted as a performance-based bonus given to the manager of the firm, and the stock-based reward as a stock grant or stock option. The model parameters α_ϕ and α_p trade off the relative importance of profits and stock price in the reward signal. They sum to unity and are fixed at the beginning of the simulation and held constant throughout.

Given the reward signal, the firm learns to make decisions using a reinforcement learning algorithm (Bertsekas and Tsitsiklis, 1996; Sutton and Barto, 1998). Reinforcement learning provides a way to estimate the action-value function from experience. Based on observations of states, actions and rewards, the learner can build up an estimate of the long term consequences of its actions. Intuitively the used learning rule minimizes the squared error between the action-value function and a bootstrap estimate based on the current reward and the future discounted return, as estimated by the action-value function. Given the reward signal at each time step, the learning agent attempts to act so as to maximize the total (discounted) reward received over the course of the task. The discounting indicates how “impatient” the manager is to receive its reward. It can also be related to the interest rate for a low-risk investment or the rate of inflation. The firms’ learning is triggered by the model parameters firm learning rate (ν) and reinforcement learning discount factor (γ). In order to get good learning, the firm learning rate (ν) should be rather low. If the discount factor (γ) is low, the firm focuses on the near-term, if it is high, it will focus on a long-term time horizon.

The implemented algorithm is designed to iteratively improve the firms' strategies, given the constraints on their knowledge and computational power.

4.1.3 The Financial Market

Our financial market represents a standard capital market model (see for example Arthur et al., 1997; Brock and Hommes, 1998; Dangl et al., 2001). Myopic investors maximize their next period's utility subject to a budget restriction. At each time step agents invest their wealth in a risky asset (a stock or index of stocks) and in bonds, which are assumed to be risk free. The risk free asset is perfectly elastically supplied and earns a risk free and constant interest rate. Investors are allowed to change their portfolio in every time step. As in many other heterogeneous agent models the existence of two kinds of investors is assumed: fundamentalists and chartists. The two types of investors differ in how they form expectations about future prices. Additionally investors have different time horizons which are modeled via the time length agents look back into the past. Fundamentalists determine their price expectations according to a model based on fundamental information, which in the integrated markets model are past dividends. They calculate a fair price and expect that the current price will gradually move towards it at some fixed rate. A fundamentalist assumes that the fair price is a linear function of past dividends. Chartists use the history of the stock prices in order to form their expectations. They assume that the future price change per period equals the average price change during the previous periods. The market uses a sealed-bid auction, where the clearance mechanism chooses the price at which trading volume is maximized. Note that there may be a range of prices that would maximize volume. The maximum price in this range is selected. If there are buy orders but no sellers then the share price is set to the maximum bid. If there are only sell orders then the price is set to the minimum ask. If there are no orders in a time period, then the price remains unchanged. Each trader specializes in a single firm, and only buys or sells shares in this firm. Each trader is initialized with a supply of shares in its firm of interest.

The timing of the events within the financial model can be described as follows. The first step is the formation of expectations. Based on past prices and dividends an investor forms its expectation about the distribution of the next period's price and dividend. The trading agent is then able to determine the demand function, which is submitted to the stock market via limit buy orders and limit sell orders. After the orders of all agents are submitted the stock market calculates this period's equilibrium price At the end of the period the current dividend is announced and becomes public information.

4.2 Social Consumer Agents

The integrated markets model, which already incorporates a validated consumer and financial market, will serve as a testbed for the new consumer market. The advantage of this approach is that one can profit from an already validated and realistic model while a part of the model (the consumer market) will be improved. It allows one to investigate the behavior of the new bounded rational and socially connected consumer agents in an integrated context.

The consumer market consists of product manufacturing firm agents and regularly re-purchasing consumer agents. During a simulation time step, each consumer must make an individual product purchase decision based on the following factors:

- its preference in product space (individual needs)
- the behavior of its social network and
- the current price of the offered products.

Furthermore the agents are able to commit to repetition, imitation, social comparison and deliberation behavior dependent on their cognitive state (satisfaction and uncertainty).

4.2.1 Consumer Preferences

The product features are represented in two dimensions as pairs of real numbers in the range [0, 1]. Each firm manufactures a single product with certain properties, which define the product’s position in feature space and are adaptable to the consumer’s demands. Each consumer agent is initialized with a random product preference in product feature space. There is no distinction between product features and consumer perceptions of those features.

The product preference *IN* represents the individual needs of an agent (equation 7). It is calculated at each simulation time step and is a function of the distance between the firms' manufactured products and the consumer agent's own preferences. The measure is computed as one minus the Euclidian distance between the position of the ideal preferred product of customer *c* (IP_c) and the position of the produced product *i* (PP_i) in the two-dimensional feature space (equation 7).

$$\mathrm{IN}_{c,i} = 1 - \sqrt{\frac{(PP_i - IP_c)^2}{2}} \qquad (7)$$

4.2.2 Social Networks

Every consumer agent is embedded in a social network structure which is randomly initialized regarding the number of neighbors and the topology of the network.

For a social network structure to have "small-world" topology it must exhibit certain properties. This can easily be described in a graphical example. Figure 6 shows three examples of networks with fifteen consumers, each with an average of four neighbors. Every vertex represents one consumer agent and an edge represents a bi-directional connection between two consumer agents. The left picture shows a completely regular graph (random connection probability per consumer which is henceforth denoted as proportion of clustering *PCLUS* is zero). The right graph represents a completely random connected topology (random connection rate or *PCLUS* is one). Although regular networks and random graphs are useful idealizations, many real networks lie between the extremes of order and randomness. For intermediate values of randomness (the middle picture consists of fifteen percent random connections or a *PCLUS* of 0.15) the graph can be interpreted as a small-world network. To construct small-world network topologies it is useful to start out with a completely regular graph. Then with a certain probability one can reconnect each edge to a randomly chosen vertex over the entire ring, with duplicate edges forbidden. The small-world networks are much more clustered than a random graph. Hence if consumer A is linked to B and B is linked to C, there is a greatly increased probability that A will also be linked to C, a property that is called transitivity (Wassermann and Faust, 1994). Despite the high clustering small-world networks have characteristic small path lengths, like random graphs (Watts and Strogatz, 1998; Strogatz, 2001).

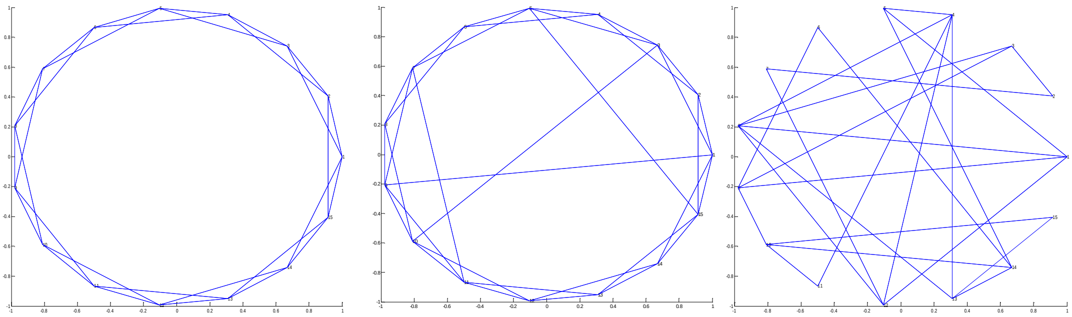


Figure 6: Example of a regular graph (left, *PCLUS* equals 0), small-world network (middle, *PCLUS* is 0.15) and a completely random graph (*PCLUS* is equal to 1). Each graph consists of fifteen consumers, all connected with on average four neighbors (adapted from Watts and Strogatz, 1998).

The "social" market share *SM* (equation 8) is defined to transform the social network into a relevant decision structure for an individual consumer agent *c*. It is represented by the quantity of the last purchases of product *i* in the consumer agent *c*'s social neighborhood ($LPP_{c,i}$) divided by the number of all purchases occurred in its neighborhood (products range from 1 to *n*).

$$\mathrm{SM}_{c,i} = \frac{\sum_{c} LPP_{c,i}}{\sum_{c}\sum_{n} LPP_{c,n}} \tag{8}$$

Intuitively, the social market share represents a measure of a product's popularity amongst a "clique" of socially connected people.

4.2.3 Cognitive States

According to the consumat approach (Janssen and Jager, 2000) two intrinsic cognitive states can account for different types of behavior and decision making. Dependent on their experienced level of satisfaction (*S*) and uncertainty (*U*) consumer agents are able to commit to repetition, imitation, social comparison and deliberation behavior.

It is defined that the consumer *c* experiences the following satisfaction level (*S*) regarding the purchase of product *i* (equation 9).

$$S_{c,i} = SM_{c,i} * SNW + IN_{c,i} * (1 - SNW) + (1 - \frac{P_i}{\max(P)}) * PSAT \tag{9}$$

Thus consumer agents can react to their individual needs (*IN*), social needs (*SM*) and the prices of the produced products (*P*) with modification of their cognitive parameter satisfaction (*S*). Furthermore satisfaction weighs the social market share (weight *SNW)* against individual needs (weight *1-SNW)* and the price of the offered product (weight *PSAT*).

A consumer agent's experienced uncertainty (*U*) is defined as the squared deviation of the actual level of satisfaction (S_t) from its expected level of satisfaction which equals the agent's last obtained satisfaction level (S_{t-1}, see equation 10).

$$U_t = (S_t - S_{t-1})^2 \tag{10}$$

To differentiate between possible actions threshold parameters for minimum satisfaction (S_{min}) and maximum uncertainty (U_{max}) are introduced. They also represent an agent's bias to commit to a certain category of action with a certain probability (table 1).

Table 1: Actions resulting from cognitive state variables of consumer agents (according to Janssen and Jager, 2000).

Cognitive state	Satisfied	Not Satisfied
Certain	Repetition	Deliberation
Uncertain	Imitation	Social Comparison

The agent's performed behavior and purchase decision is a result of its experienced levels of satisfaction and uncertainty:

- Repetition: if the agent experiences satisfaction ($S > S_{min}$) and is also certain about its choice (that means that its last choices nearly met its expectations, hence $U \leq U_{max}$) then it has no reason to change its last decision. Therefore the customer agent will consume exactly the same product which it purchased the last time step.
- Imitation: if a customer agent again feels satisfied ($S > S_{min}$) but it experiences uncertainty (its last choice deviated much from its expectations and $U > U_{max}$) then the customer will investigate its social neighborhood and give the product a try that is consumed most by its friends. If there is more than one product one will be randomly selected among the most purchased products.
- Deliberation: if a consumer is not satisfied ($S \leq S_{min}$) and it is certain (its expectations were met, thus $U \leq U_{max}$) it will purchase the product with the highest overall satisfaction value (according to equation 9). Again, if there is more than one candidate product, one will be randomly selected among the most satisfying products.
- Social comparison: if the consumer agent happens to be not satisfied ($S \leq S_{min}$) and uncertain ($U > U_{max}$) at the same time step, it will engage in a behavior called social comparison. This means that the agent will consider the product that is consumed the most in its social neighborhood (analogue evaluation of the social market share) but one that also exceeds or reaches its expectations for satisfaction (see equation 9) originating from its last consumption. If there is more than one candidate product, one will be randomly selected from the eligible products.

With this cognitive decision structure implemented and the agents' ability to relate their expectations to their social network the simulation results are validated against a complex behavioral phenomenon and an empirically stable stylized fact found in consumer markets.

4.3 Model Validation

Gaining crucial insights into underlying mechanisms of real markets is a major goal of agent-based economic modeling. Thus a useful model should be able to reproduce observable market behavior or so-called “stylized facts” capturing the dynamics of the investigated market. Therefore the model is validated against these empirical properties by using a recently introduced algorithm based on Markov chain Monte Carlo (MCMC) sampling (Sallans et al., 2003).

The MCMC sampling helps to focus computational power on parameter space areas where stylized facts are reproduced very well. The goal is to understand the impact of parameters on model behavior, especially in these interesting areas. The stylized facts of the consumer market which are analyzed in this work are the properties of the consumers’ social networks (sections 4.2.2 and 4.5) and the Bass diffusion model (sections 4.3.2). In order to quantify how well the market reproduces a stylized fact, an energy function is defined. The energy function represents a measure of the fit of the output of the model to the stylized fact. An energy function for a stylized fact is constructed such that low energy corresponds to good reproduction of the fact. For example, an energy function for the Bass diffusion model would generate low values if a Bass curve fits very well to a firm’s market share data (section 4.3.3).

The MCMC procedure first randomly changes model parameter values before a simulation run which generates one sample of model parameters. Then the quality of the generated parameter sample is evaluated based on a previously defined energy function, which is unique for each stylized fact. The sample is accepted or rejected based on the energy and the MCMC procedure starts over until an arbitrary number of parameter samples is drawn (a minimum of one thousand samples is chosen to get statistically significant results). The advantage of this method is that computational resources are distributed on what are probably the most interesting parameter combinations. The whole validation procedure works as follows:

- Selection of an empirically stable stylized fact
- Design of an adequate energy function for that fact
- MCMC simulation runs
- Analysis and perhaps repetition of simulation runs

The MCMC sampler which is used was recently introduced by Sallans et al. (2003) and is based on principles of the Metropolis algorithm (Metropolis et al., 1953). It has the property that samples are more likely to be drawn from low-energy areas. The sampler acts as a “directed” random walk through model parameter space, avoiding high-energy areas. In the limit, parameter samples are drawn according to the normalized probability distribution defined by the energy function. But even without theoretical guarantees on the distribution of sampled parameters, the sampler can find good model parameter settings, and reveal interesting correlations between model parameters. In practice, one is not able to generate Markov chains that are sufficiently long to reach the equilibrium distribution. Instead one can be content with one thousand samples drawn for each model run. While this is too short to allow for convergence, one can still examine the sample set to identify regions where stylized facts are well reproduced, and look for statistically significant correlations between parameters. Validation results for the Bass model runs are shown in section 4.4 and for the social networks in section 4.5.

4.3.1 Model Parameters

The focus of state of the art modeling techniques is not to cover every market phenomenon observed. Rather it lies on “noncritical” abstraction and careful parameter selection by gradually adding complexity once the previous model has been fully understood. This prevents the modeler from introducing ad hoc parameters to capture important causal relationships which might capture no market phenomenon at all.

The presented model is built on the foundation of a validated integrated markets model including consumer, firm and stock trading agents. Thus one starts out with the originally given parameter values (Sallans et al., 2003) which guarantee a well functioning integrated financial and consumer market. Despite the goal to keep the model as simple as possible, additional parameters were necessarily introduced to account for the social network functionality and the improved agents’ cognitive decision structure (table 2). All parameter values must be initialized before a model simulation is run. The column “value” of table 2 shows the start values used for the validation runs with the MCMC sampler. These values were found to be plausible based on evaluations of initial trial simulation runs (values in italics are given by the original model and were held fixed for all simulation runs).

Table 2: Model parameters for the integrated markets simulation.

Parameter	Description	Range	Value	Reference
NCons	Number of simulated consumer agents	N	100	Section 4.2.2
NNbs	Number of average neighbors per consumer agent	N	4	Section 4.2.2
PClus	Percentage of randomness of small-world network	[0, 1]	0.1	Section 4.2.2
SNW	Weight of social network for satisfaction	[0, 1]	0.5	Equation (9)
PSat	Weight of price for satisfaction	[0, 1]	0.5	Equation (9)
SAT_{min}	Threshold for minimum satisfaction of consumer agent	[0, 1]	0.5	Section 4.2.3
UNC_{max}	Threshold for maximum uncertainty of consumer agent	[0, 1]	0.5	Section 4.2.3
ν	Firm learning rate	$R \geq 0$	0.001	Section 4.1.2
γ	Reinforcement learning discount factor for firm	[0, 1]	0.83	Section 4.1.2
α_ϕ	Strength of profitability reinforcement to firm	[0, 1]	*0.47*	Section 4.1.2
α_p	Strength of stock price reinforcement to firm	[0, 1]	*0.53*	Section 4.1.2
N	Number of consumer cluster centers	N	*3*	Section 4.1.2
N_f	Proportion of fundamentalist traders	[0, 1]	*0.57*	Section 4.1.3
N_c	Proportion of chartist traders	[0, 1]	*0.43*	Section 4.1.3

The quality of reproduction of the stylized facts should simply depend on the characteristics of the model's behavior. The parameters, which account for different features of the integrated markets simulation, can be grouped as follows:

- Social network properties: These are described by the number of consumers (*NCONS*), the average number of neighbors (*NNBS*), and the proportion of clustering (*PCLUS*). *NCONS* and *NNBS* account for the dimension of the artificial consumer market. The proportion of clustering (*PCLUS*) accounts for the complexity of the social network structure. While a value of zero represents a completely regular graph with low complexity, a value of one indicates a completely random connected topology consisting of the highest possible structural complexity (see figure 6 in section 4.2.2). For values between these extremes, the consumers' social structure exhibits small-world properties.
- Consumers' cognitive behavior: The consumers' behavior and decisions are triggered by the parameter weight for social needs (*SNW*), individual needs (*1-SNW*) and product price (*PSAT).* These parameters account for the level of satisfaction and uncertainty

experienced by the consumer. Furthermore thresholds for minimum satisfaction (SAT_{MIN}) and maximum uncertainty (UNC_{MAX}) will influence the action a consumer agent commits to (section 4.2.3).

- Firms' learning behavior: The firms' learning is triggered by the firm learning rate (ν) and the reinforcement learning discount factor (γ). In order to get good learning, the firm learning rate (ν) should be rather low. If the discount factor (γ) is low, the firm focuses on the near-term, if it is high, it will focus on a long-term time horizon.
- Fixed parameters: The firm agent's parameters α_ϕ and α_p, which sum to unity, trade off the relative importance of profits and stock price in a firm agent's decision-making process (see section 4.1.2). N denotes the number of cluster centers as described in section 4.1.2. As mentioned in section 4.1.3 the stock market consists of fundamentalists and technical traders. The parameters proportion of fundamentalists (N_f) and proportion of chartists (N_c) maintain the heterogeneity of the market traders, which is necessary to preserve financial market liquidity and trading volume.

4.3.2 The Bass Diffusion Model

The seminal work of Frank M. Bass (1969) describes a simple mathematical model of market penetration of a new product or concept as a function of internal (for example word of mouth) and external influences (for example advertising). The model and its variations have been successfully applied by marketing scientists in many different areas for over 30 years. Examples include DirecTV (early 1990s), a satellite television service which forecasts new subscription rates, and RCA (mid 1980s), which effectively used an extension of the Bass model to forecast the sales of CDs as a function of the sales of CD players. Fields of application are usually the quantification of the speed of diffusion of durable and non durable products and the forecast of future consumer adoptions (see Van den Bulte, 2002, for a meta-analysis of research on different product types over different regions).

The diffusion of innovations is influenced by interpersonal and mass media communication. The effects of interpersonal communication in particular are thought to be a key factor for the speed and shape of the diffusion of an innovation (Rogers, 1983;

Mahajan et al., 1990). The theory of network externalities provides a related explanation and quantification of increasing consumer demand and S-shaped diffusion of network goods or service sales over time (Grajek, 2002). Positive network externalities are defined as utility, which consumers derive from consumption of a good or service. The utility increases with the number of other consumers. Economic literature usually distinguishes between direct and indirect network externalities (see for example Katz and Shapiro, 1985; Economides, 1996). Direct network externalities are related to physical networks (for example telecommunication technologies). The utility that consumers derive from using these technologies, undoubtedly depends on the number of other users. An obvious reason for a positive dependence is that a larger network allows consumers to satisfy more communication needs and may decrease the common costs of the service. Another explanation might be the bandwagon effect since conspicuous consumption gives rise to a conformistic behavior (Leibenstein, 1950). A negative dependence between network size and consumers' utility might be justified by congestion or by non-conformism of consumers (snob effect). Indirect network externalities apply if a good consists of two complementary components, for example, hardware and software. The latter exhibits supply-side economies of scale (see Katz and Shapiro, 1985). Obviously the amount of users of a hardware platform determines the size of the market for software and furthermore enhances the utility gained by use of the hardware.

The original Bass model makes adoption a function of innovation and imitation effects. The theory details the characteristic sigmoid pattern observed empirically which levels off to a maturity level (see figure 7). The spread of an innovation in a market can be characterized by the Bass formula as a discrete time model (equation 11; see also Morris and Pratt, 2003).

$$X_t = p(M - X_{t-1}) + q\frac{X_{t-1}}{M}(M - X_{t-1}) \tag{11}$$

X denotes the number of consumers who will adopt at time t and *M* represents the market potential or the maximum number of people who will use the innovation. The parameters *p* and *q* provide information about the speed of diffusion. The coefficient of innovation *p* describes the external influences and the coefficient of imitation *q* describes the internal influences. A high value for *p* indicates that the diffusion starts out quickly but also decreases fast. A high *q* represents a slow diffusion process at first, which accelerates quickly afterwards (for example take-off is slower for non durables and products with competing standards that require heavy investments). In the model, the different firms' market share time series are validated against a cumulative discrete standard Bass function which gives the absolute number of adopted consumers at a certain point in time (equation 12).

$$X_t = X_{t-1} + p(M - X_{t-1}) + q\frac{X_{t-1}}{M}(M - X_{t-1}) \tag{12}$$

The market potential parameter *M* was set to one, representing the maximum possible proportion of agents in the competitive consumer market environment. Examples of standard bass curves (values for *p*=0.03 and *q*=0.38 describe the basic Bass model) generated by the integrated markets simulation are shown in figure 7.

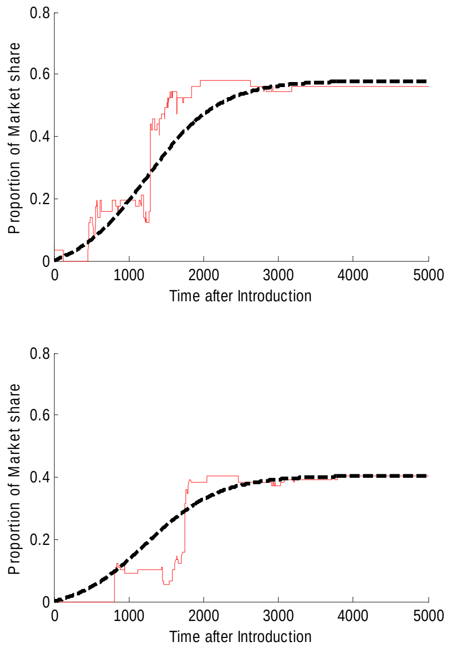


Figure 7: Examples of generated Bass curves in the artificial consumer market. The black dotted lines represents the standard Bass curve with the parameter values for *M*=1, *p*=0.03 and *q*=0.38.

4.3.3 The Energy Function for the Bass Model

To investigate which parameter settings have influence on the development of Bass curves in the artificial consumer market it is necessary to define an adequate energy function for the adapted Metropolis algorithm. The sampler acts as a "directed" random walk through model parameter space, avoiding high-energy areas. In the limit, parameter samples are drawn according to the normalized probability distribution defined by the energy function. The energy function is a measure presenting the optimal fit of a standard Bass curve on the consumer market share time-series. The measure should neither depend on where in the data the Bass curve is located (translation invariance with respect to time) nor on the scaling of the curve (scale invariance, see, for example, Bishop, 1995). The cross-correlation function (equation 13) represents a good solution to overcome these problems since the function is not sensitive to y-scaling (height) of the data, when comparing two different time series. To account for the x-scaling (time) the function is set up to compute the maximum correlation coefficient over *all* time lags (equation 14). Each data point of the sample (one discrete time step) equals a single simulation step.

$$\rho_{XY}(k) = \frac{\sigma_{XY}(k)}{\sqrt{\sigma_X \sigma_Y}} \tag{13}$$

$$\rho_{\max} = \max_k \{\rho_{XY}(k)\} \tag{14}$$

While X denotes the market share time series of a certain firm in the artificial consumer market, Y represents the time series of a standard Bass curve.

To find the optimal fitting standard Bass curve for X, one can set up a nonlinear optimization algorithm based on golden section search and parabolic interpolation (see, for example, Forsythe et al., 1976; Hagan et al., 1996). The algorithm fits standard Bass curve time series with different width (in Y) to X and minimizes the negative cross-correlation over all lags between X and Y. Hence the optimization algorithm varies standard bass curves by scaling until it finds the maximum cross-correlation coefficient (the best match). The energy for the MCMC sampler is then calculated as the reverse of the maximum correlation coefficient, since low energy corresponds to good reproduction of the stylized fact (equation 15).

$$E = \frac{1}{\rho_{\max}} \tag{15}$$

4.4 MCMC Validation Results

The market dynamics of the validated model that emerged and the set of parameter values identified, for which standard Bass curves could be reproduced very well, are presented in the following sections. First, the overall consumer market dynamics will be described, followed by a detailed analysis of the parameters and their relationships, grouped by their functionality. All the simulations were based on five firm agents (held fixed over all runs) acting in consumer markets initialized with one hundred consumer agents.

4.4.1 Overall Market Dynamics

Interestingly, the emerging market behavior of the simulation models is not restricted to the one investigated stylized fact (Bass curve) of a single firm. It is embedded in a realistic market scenario with oligopolic properties. Empirical investigations have shown that in real-life markets it is very frequent to find oligopolical industries that are characterized by a large range of different market shares, with no two firms having the same market share. Traditional economic models of quantitative competition oligopoly are not successful in explaining this stylized fact (Watt, 2002). The Cournot model predicts equal market shares for all competitors, while a generalized Stackelberg leader-follower oligopoly model with one leader predicts a larger market share for the leader and equal market shares for all followers (see Stackelberg, 1934; Sherali, 1984; Daughety, 1988).

The aggregate market dynamics of the simulated consumer market is represented by the market share of each firm and reflects the empirically found oligopoly market related stylized facts. For example, one or more firms attracted certain consumers by successful introduction and development of their products, while the others lost in market share or engaged in price wars. Bass curves solely emerged in low energy areas of the defined Bass model energy function (figure 8) while they did not emerge in high energy areas (figure 9). The squared correlation coefficient ρ^2 (equation 14) gives the proportion of variance explained by the fitted Bass curves with respect to the market share data. It is useful to compare the Bass curve reproduction quality of the different models.

Figure 8 presents samples where Bass curves were well reproduced. Run 1 (left column) shows the market share of two competing firms. The market leader (firm 5) increasingly looses market share to the market entrant (firm 1) which introduced its innovation at the beginning of the time period. Run 2 shows a similar dynamic with four competing firms. Here firm 5 introduces its innovation and competes against firms 2, 3 and 4. The competition ends with two market leaders, which basically divide the market. One is the former market leader, the other winner is the innovative newcomer. Run 3 also presents one firm with an emerged Bass curve in its market share (firm 2). It competes against firm 1, the former market leader, and firm 5. The scenario also ends in a rather stable oligopoly.

The runs in figure 9 present samples from a high energy area of the Bass energy function which did not generate any Bass curves. Run 1 shows initially four competing firms (firm 1 to 4). As the theory of Stackelberg (1934) suggests the competition ends in a stable oligopoly with one leader (firm 3) with higher market share and two followers with a nearly equal market share (firm 1 and 2). Run 2 shows four competing firms with a rather oscillating market share. First firm 4 seems to be the market leader soon beaten by the newcomer (firm 2). Then firm 1 fights against 2 until firm 3 gathers the whole market share. Afterwards firm 1, 3 and 4 compete until firm 3 wins again. Simulation run 3 shows a rather soft competition where firm 2 and 5 increasingly gain in market share while the firms 1 and 4 seem to have a decreasing trend in market share.

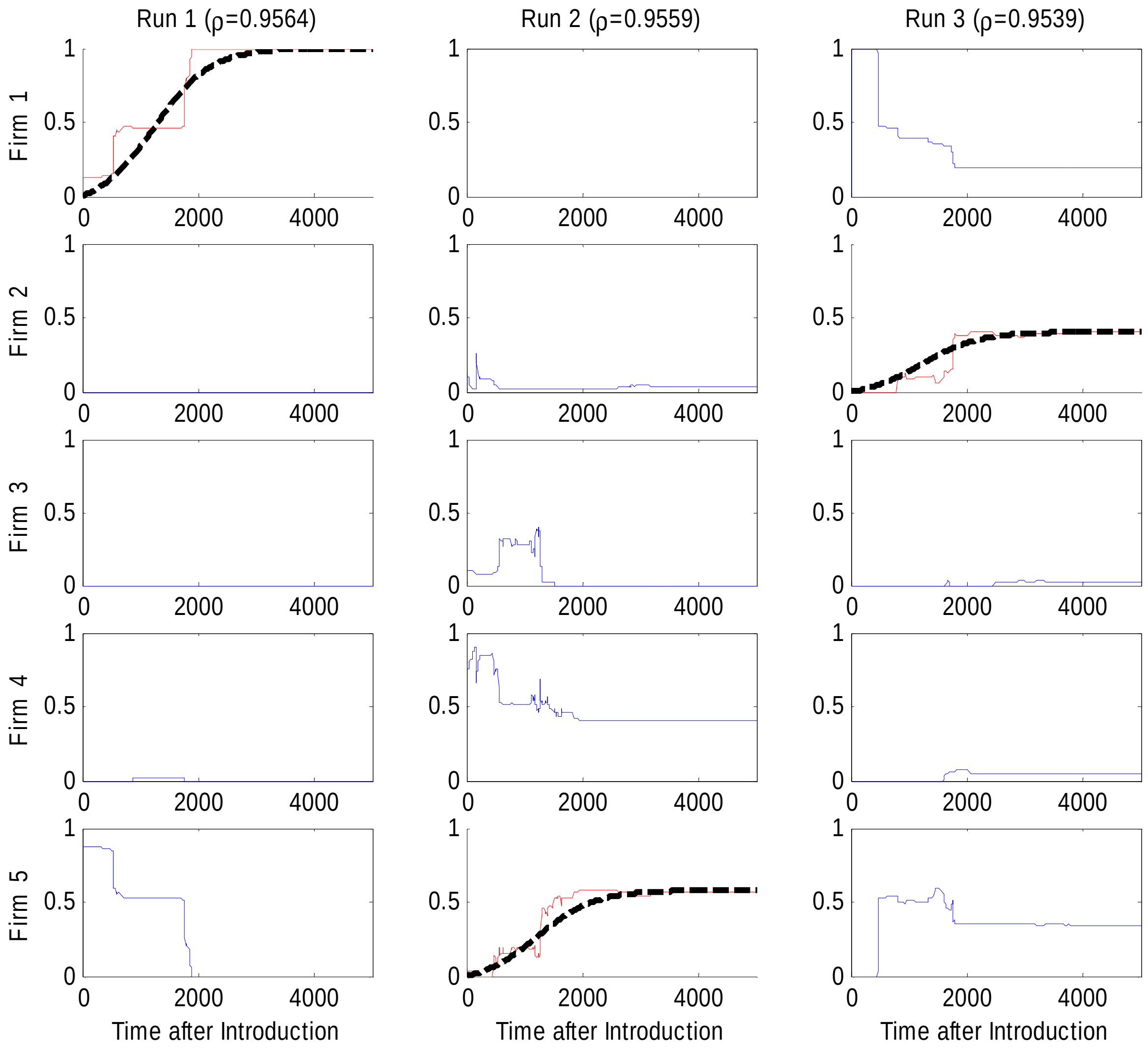


Figure 8: Three typical examples of simulation runs in low energy areas of the Bass energy function (section 4.3.3). Each column shows the emerging consumer market dynamics of an independent simulation run involving five firm agents and 69 (run 1), 57 (run 2), and 89 (run 3) consumer agents. The y-axis denotes the proportion of market share an individual firm agent could obtain at a specific point in time. The dotted line indicates the best fit of a standard Bass curve to the market share time series with a resulting correlation coefficient of ρ. The proportion of variance explained by the fitted Bass curves with respect to the market share data is given by a ρ^2 of 0.9147 (run 1), 0.9138 (run 2), and 0.91 (run 3).

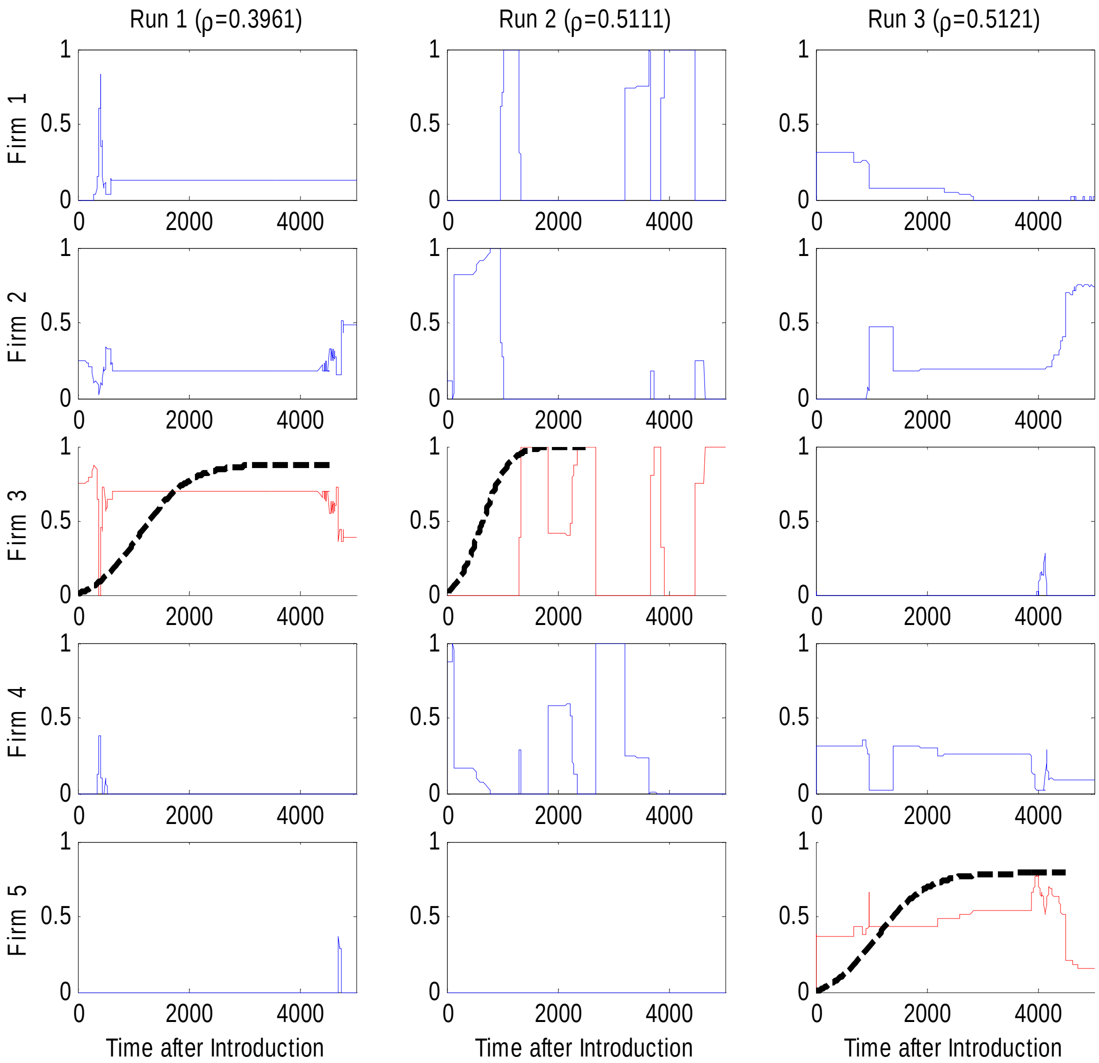


Figure 9: Three typical examples of simulation runs in high energy areas of the Bass energy function (section 4.3.3). Each column shows the emerging consumer market dynamics of an independent simulation run involving five firm agents and 74 (run 1), 75 (run 2), and 57 (run 3) consumer agents. The y-axis denotes the proportion of market share an individual firm agent could obtain at a specific point in time. The dotted line indicates the best fit of a standard Bass curve to the market share time series with a resulting correlation coefficient of ρ. The proportion of variance explained by the fitted Bass curves with respect to the market share data is given by a ρ^2 of 0.1569 (run 1), 0.2612 (run 2), and 0.2622 (run 3).

4.4.2 Ideal Model Parameters

The parameter values where standard Bass curves could be reproduced very well is presented in the form of histograms in figure 10. The “ideal” parameters do not take on extreme values, which is an indicator for the plausibility of the model. Furthermore, table 3 provides information about relationships between parameters. In the following sections the model parameters will be described and interpreted.

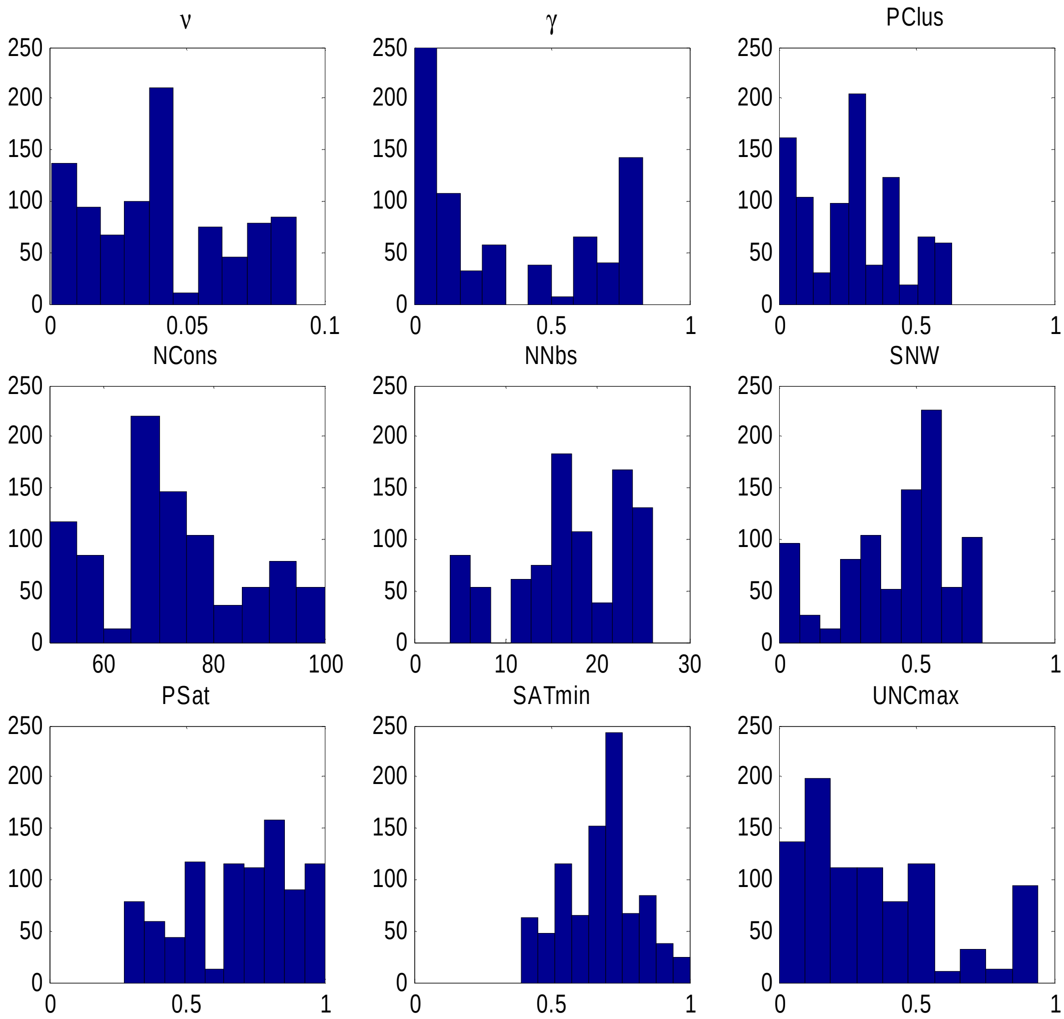


Figure 10: Histograms of parameter values from MCMC sampling for the Bass curve energy function. The histograms include the 90 % of samples with the lowest energy (equation 15). The x-axis denotes the parameter value and the y-axis denotes the number of sample runs.

Table 3: Correlation coefficients of the Bass validated integrated markets model based on 416 samples.[10] The measures social clustering coefficient (*SCC*) and social clustering length (*SCL*) are described in section 4.5.

	ν	NCons	γ	NNbs	PClus	SNW	PSat	SAT_{min}	UNC_{max}	SCC	SCL
ν	1.00										
Ncons	****-0.74**	1.00									
γ	0.11	-0.06	1.00								
NNbs	-0.09	0.03	-0.37	1.00							
Pclus	0.43	***-0.56**	-0.28	0.40	1.00						
SNW	-0.37	***0.52**	-0.11	-0.39	-0.28	1.00					
PSat	****-0.76**	0.49	0.01	0.21	-0.35	0.19	1.00				
SAT_{min}	****-0.75**	0.49	-0.38	-0.21	-0.29	***0.52**	***0.57**	1.00			
UNC_{max}	0.33	0.18	0.11	-0.37	-0.03	0.44	-0.44	-0.19	1.00		
SCC	-0.29	0.14	0.14	0.36	***-0.51**	-0.33	0.31	-0.15	-0.46	1.00	
SCL	-0.22	0.37	0.31	****-0.83**	****-0.65**	0.46	-0.14	0.27	0.37	-0.06	1.00

regression coefficients are significant at the *5 % level or at the **1 % level.

4.4.2.1 Firm Learning

The firms' learning behavior is dependent on the firm learning rate (ν) and the discount factor (γ). In order to get good learning, ν should be rather low. If the discount factor γ is low the firm focuses on near-term, if it is high it will focus on a long-term time horizon. For the Bass model an intermediate value of 0.04 for ν seems to be most appropriate (the initial value was set to 0.001, see figure 10). γ is initialized with a value of 0.83 and has it's peak around the rather low value of 0.06. Relationships between the firm learning and other parameters are interpreted as follows:

- Market complexity: The firm learning rate is negatively correlated with the number of consumers (*NCONS*) with a significant correlation coefficient of -0.74 (p=0.0015, see table 3), which indicates the necessity of better learning in a bigger and therefore more complex market environment (see figure 11, left picture).

[10] Significance was measured in the following way: first, the sequence of parameter values was subsampled such that autocorrelations were insignificant at the one percent confidence interval. Given this independent sample, the correlations between parameters could be measured with effective significance levels.

- Product price: A significant negative correlation of -0.76 (p=0.001) exists between ν and the price weight (*PSAT).* Since the firms can change their product's price or its features this means that the importance of the product price for consumers increases if firms are able to engage in more intelligent actions, for example by making necessary price adaptations (figure 11, right picture).
- Consumer satisfaction and adaptation: Another finding is that the overall consumer satisfaction and uncertainty seems to decrease with the learning rate ν. This is reflected by a negative correlation coefficient of -0.75 (p=0.0013) between ν and the threshold for minimum consumer satisfaction (SAT_{MIN}) and the positive trend (cc=0.33, but not significant with p=0.23) between ν and the threshold for maximum consumer uncertainty (UNC_{MAX}). Hence if the firms exhibit better learning (ν gets smaller) the consumers tend to be rather unsatisfied since the threshold for minimum satisfaction increased. They also get rather certain since the threshold value for uncertainty (UNC_{MAX}) increases. As an implication the consumer agents have a high probability to exhibit deliberative behavior, where they simply choose the maximum satisfying product (section 4.2.3). Thus they can react more sensitively to the firms' product price adaptations, a stylized fact that is strongly supported by the notion of network externalities (see section 4.3.2). Hence, intelligent firm agents seem to lead to better adapted consumer agents in a Bass curve reproducing market scenario.

So far a good model for Bass curves seems to involve firms which are rather good learners, operating in a market environment of stable complexity and small world network properties (see also parameter *PCLUS* in the next section).

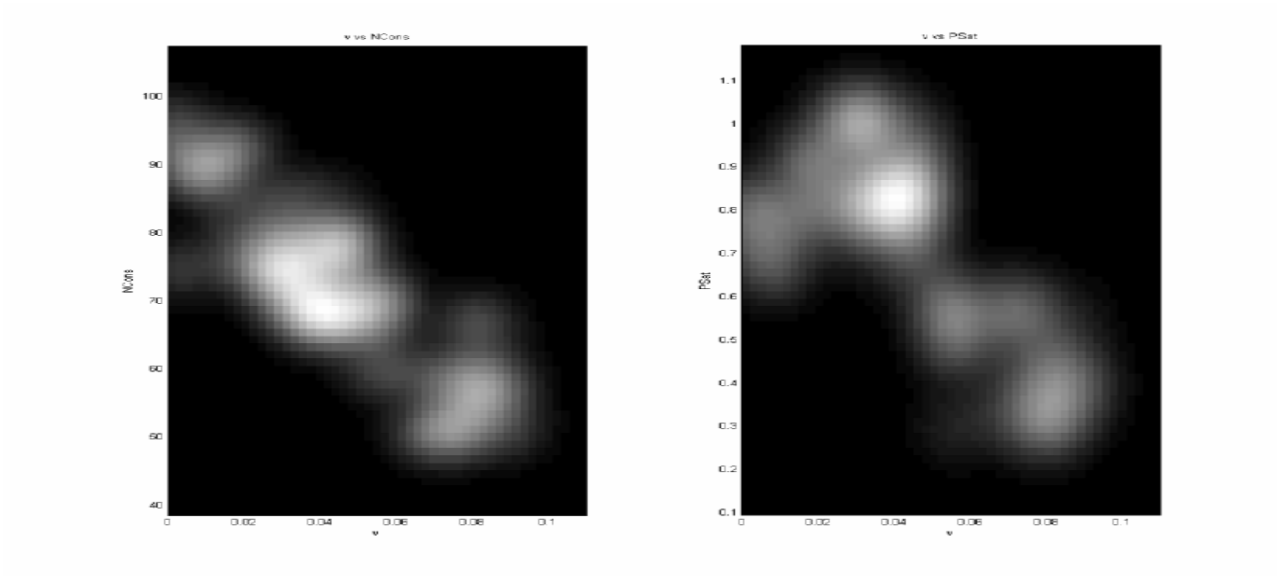


Figure 11: Negative correlation between firm learning rate (ν) and number of consumers (*NCONS*) and negative correlation between ν and price weight (*PSAT*). The plot shows the density[11] of samples for the different parameter values and includes the 90 % of samples with the lowest energy (equation 15).

4.4.2.2 Social Network Structure

The social network properties are described by the model parameters number of consumers (*NCONS*), average number of neighbors (*NNBS*), and proportion of clustering (*PCLUS*). *NCONS* and *NNBS* account for the dimension of the consumer market while *PCLUS* accounts for the complexity of the social network structure. As mentioned in section 4.2.2 a value of zero represents a completely regular graph with low complexity, while a value of one indicates a completely randomly connected topology. For values between these extremes, the consumers' social structure exhibits small-world properties (the exact properties of real-life social networks are described in section 4.5).

A "good" social network for Bass curves seems to be one with a moderate number of consumers (peak at 70), with each of them having around 16 neighbors on average (figure 10). Social network related parameters are interpreted as follows:

---

[11] The density plots were generated using the kernel density estimator for Matlab provided by C.C. Beardah at http://science.ntu.ac.uk/msor/ccb/densest.html (Beardah and Baxter, 1996).

- Small-world principles: The clustering rate *PCLUS* (see section 4.2.2) had two major peaks, a smaller one at zero and one at a value which lies around 0.3. While the first represents a network with a regular topology the latter is a strong indicator for the preference of a social network based on small-world principles. But there is more evidence on the importance of small-world properties for the occurrence of Bass curves. From the first half of samples only 1.2 % exhibit a proportion of clustering bigger than 0.5. This is in contrast to the second half of samples where already 27.5 % show a $PCLUS \geq 0.5$. Additionally it is found that all samples with a *PCLUS* between 0 and 0.1 had an average Bass energy correlation coefficient (equation 14) of 0.699 (equals 49 % of explained variance). Samples with a *PCLUS* between 0.6 and 0.7 exhibited an average Bass correlation coefficient of 0.823 (equals 68 % of explained variance).
- Balanced network structure: A balanced social structure seems to be necessary for the Bass curve stylized fact. This is substantiated by a negative correlation of -0.56 ($p=0.028$, table 3) between number of consumers and proportion of clustering in the consumer market. Since the proportion of clustering accounts for the complexity of the social network structure an increased number of consumers (increased dimension and complexity) interestingly leads to the preference of a lower proportion of clustering by the MCMC sampler.

Hence the consumer market seems more likely to reproduce Bass curves if the social network has balanced complexity and is structured like a small-world network.

4.4.2.3 Consumers' Cognitive States

In the integrated markets model the consumers' cognitive behavior and decisions are triggered by the parameter weight for social needs (*SNW*), individual needs (*1-SNW*), and product price (*PSAT).* Thresholds for minimum satisfaction (SAT_{MIN}) and maximum uncertainty (UNC_{MAX}) regulate the actions consumers will most likely commit to (see section 4.2.3). Consumer agents' cognitive parameters are interpreted as follows:

- Network externalities: The simulation results support the hypothesis of positive direct network externalities as an underlying mechanism for Bass curves (section 4.3.2). First as mentioned in section 4.4.2.1 consumer agents' sensitivity to price increases with more intelligent firm actions (price adaptations). In addition, the level of minimum satisfaction increases with the weight of price (cc=0.57, p=0.028, see table 3). Hence if the price gets more weight the consumers tend to be rather unsatisfied since the threshold for minimum satisfaction increases. This is another indication for positive network externalities since consumers seem to become more satisfied with a lower price (weight). Furthermore the top 30 % of samples (sorted by reproduction quality of Bass curves ρ_{max}, see equation 14) exhibit a mean correlation coefficient ρ_{max} of 0.75 and a mean consumer satisfaction proportion of 97 % (measured by the proportion of appearance of the consumer behaviors "imitation" and "repetition", see also section 4.2.3). In contrast the first 70 % of samples with a mean Bass reproduction correlation coefficient ρ_{max} of 0.67 show only 91 % consumers engaged in "satisfying" behavior. Hence there is a positive trend for increasing consumer satisfaction with quality of Bass curve reproduction. This can be explained by positive network externalities.
- Price vs. social needs: For the reproduction of standard Bass curves the best weighting factor for the social market share parameter (*SNW*) lies slightly above the initial value of one half (with a concentration around 0.6), while the price weight (*PSAT)* has its peak at a value of 0.8. This implies a normalized proportion for social needs of 0.33, individual needs of 0.22, and price of 0.44 (see equation 9). Thus for the satisfaction function of an individual agent the social market share slightly outweighed the individual needs, while the price seemed to be the most dominant factor. Although the latter finding is strongly supported by neoclassical economic theory which states that supply and demand both are functions of price, Bass curves seem to need some additional cognitive and social parameters to occur.
- Consumer satisfaction and uncertainty: The minimum satisfaction threshold (SAT_{MIN}) has a relatively high peak with the highest concentration around 0.7. This could mean that consumer agents in general are experiencing a very high level of satisfaction or they are rather committing themselves to a social comparison or deliberation decision style since the probability to be unsatisfied is rather high (section 4.2.3). The uncertainty threshold has its highest concentration at a rather low value (0.16), which

indicates that consumer agents experience a rather low level of uncertainty in general, or agents mostly engage in repetition or deliberative behavior. In order to distinguish between these possibilities the consumer agents conducted actions will now be analyzed.

- Consumer decisions: The Bass Model summarizes in a simple mathematical form the key finding from over 4,000 diffusion studies: most people wait until they have witnessed peers having favorable experiences with the technology or service before they adopt. Hence most people imitate rather than innovate. Analysis of the simulated consumer decisions reveals that the actual dominant consumer behavior is repetition with a proportion of 89.03 %, followed by deliberation (10.13 %), imitation (0.61 %) and social comparison (0.24 %) for the top 30 % of samples of the MCMC sampler (figure 12). Thus repetition behavior seems to be the most important mechanism for the emergence of Bass curves in the consumer market model. Since the consumer market of the integrated markets model is based on repeated purchases (every consumer purchases once at a simulation time step) the development of standard Bass curves over time heavily relies on consumer's repetition behavior. Hence the model shows behavior consistent with the Bass diffusion theory (section 4.3.2). Furthermore, repetition can be viewed as a type of imitation behavior since the consumer agent imitates its own last decision. For commitment of repetition behavior consumer agents must experience satisfaction and certainty (low levels of uncertainty). For deliberation behavior they need to be certain and unsatisfied. Despite the high threshold for minimum satisfaction the consumer market environment consists of rather certain agents, who are switching between repetition (when they are satisfied) and deliberation behavior (when they are unsatisfied).

Our results show that the emergence of Bass curves in consumer markets can be explained by the underlying consumer agents' repetition and imitation behavior which leads to increasing demand, and deliberation behavior which refers to positive network externalities and leads to increased price sensitivity.

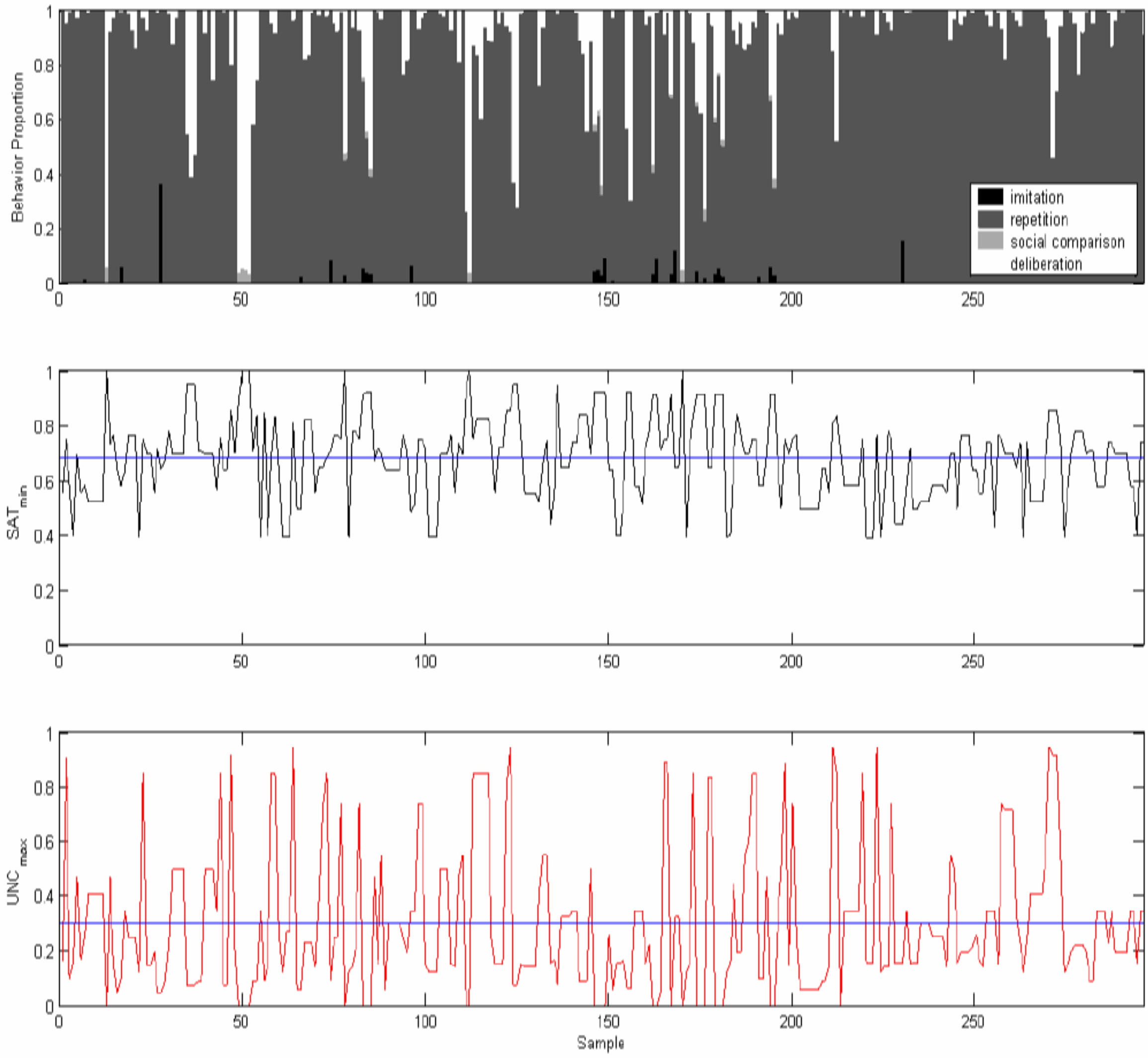


Figure 12: Consumer behavior separated into proportions of imitation, repetition, social comparison and deliberation behavior (upper diagram). The development of thresholds for minimum satisfaction (middle diagram) and maximum uncertainty (lower diagram) shows rather high levels for satisfaction and low levels for uncertainty as indicated by the mean of the values (the straight line). The plot includes the 30 % of samples with the lowest energy (equation 15).

4.5 Comparison to Real-Life Social Networks

In the previous section a small-world network topology was not explicitly imposed, but occurred because it led to the generation of Bass curves. Additional model runs were implemented, in which the simulation is forced to generate consumer markets with small-world structures related to a real-life social network. This new model is denoted as the "optimized" model. The "optimized" model should exhibit a more life-like social network structure in the consumer market. From the comparison of the "normal" Bass validated model as already described in sections 4.3 and 4.4 and the "optimized" model it is expected to gain supplementary insights into the role of the small-world properties for the consumer markets.

4.5.1 Social Clustering Coefficient and Characteristic Path Length

First, two additional estimators are introduced to better characterize the social network structure. One is the characteristic path length (*SCL*) and the other is the clustering coefficient (*SCC*, see Watts and Strogatz, 1998).

The characteristic path length measures the typical separation between two agents in the network (a global property) and is defined as the number of connections needed for the shortest path between two agents, averaged over all pairs of agents in the consumer market. To find the shortest path between two agents the Floyd-Warshall algorithm was applied (see, for example, Cormen et al., 2001).

The clustering coefficient measures the cliquishness of a typical neighborhood (a local property) and is the fraction of existing connections compared to all possible connections within an agent's neighborhood, again averaged over all consumer agents. Suppose that the consumer agent c has NN_c number of neighbors and NC_c actual connections between them. Then its clustering coefficient SCC_c is defined as follows (equation 16):

$$SCC_c = \frac{NC_c}{NN_c(NN_c - 1)\frac{1}{2}} \tag{16}$$

The characteristic path length (*SCL*) and the clustering coefficient (*SCC*) are both a function of the amount of randomness or complexity of the network structure (expressed by the parameter proportion of clustering *PCLUS*). Watts and Strogatz (1998) find that path length and clustering depend differently on the amount of randomness in the network. *SCL* decreases quickly, while *SCC* drops rather slowly with an increase in *PCLUS*. This can also be seen in table 3, where *SCL* and *PCLUS* exhibit a correlation coefficient of -0.65 (p=0.0091), while *SCC* and *PCLUS* show a correlation coefficient of -0.51 (p=0.053). This leads to a small-world network with a high amount of clustering and a rather short characteristic path length. From a social systems perspective this means that it only takes a small number of well-connected people to make a world small (Collins and Chow, 1998).

Watts and Strogatz (1998) give an empirical example regarding these estimators. They analyze the characteristics of the social network of actors via the Internet Movie Database which includes approximately 90 % of the professional actors. Watts and Strogatz define that two actors are connected if they played in the same movie. Their results include 225 actors (vertices of the graph, see figure 6) with on average 61 connected actors (edges of the graph). For comparison they provide information of a randomly connected network with the same number of vertices and average number of edges (table 4).

Table 4: Social network properties of movie actors (from Watts and Strogatz, 1998).

Social Network	Characteristic Path Length (*SCL*)	Clustering Coefficient (*SCC*)
Film actors	3.65	0.79
Random	2.99	0.00027

4.5.2 The Optimized Model

To be able to compare the model results to a model with a more life-like social network structure the “optimized” model is defined.

New model runs were executed, where the parameter proportion of clustering (*PCLUS*) is excluded from being modified by the MCMC sampler. Hence the only social network related parameters to be manipulated by the sampler were the number of consumers (*NCONS*) and the average number of neighbors (*NNBS*). Each time the MCMC sampler changed one of these two parameters, an additionally nonlinear optimizing algorithm was run (based on golden section search and parabolic interpolation, see, for example, Forsythe et al., 1976; Hagan et al., 1996) to find an optimal value for *PCLUS* for the current sample. This means that with a given *NCONS* and *NNBS* the optimizer was meant to manipulate the parameter *PCLUS* until it got as close as possible to the values for *SCL* and *SCC* shown in table 4. With this obtained "optimal" value for *PCLUS* it was possible to set up the new consumer market for the current sample. The value for *PCLUS* was retained for the following samples until the MCMC sampler again changed one of the two parameters *NCONS* or *NNBS*.

Hence the "optimized" model should explicitly exhibit small-world parameters similar to those shown in table 4.

4.5.3 Model Comparison

An overview of the differences between the optimized and the "normal" model is shown in table 5. Each of the values is discussed below:

- Values of the Bass correlation coefficient (ρ_{max}, see equation 14): Despite the fact that the means of both models look very similar in absolute values they are significantly different at the 5% significance level ($p=0.0218$) due to the low standard deviation. Interestingly there is a trend for the optimized model to have lower standard deviation. This is an indicator for the improved performance of a more life-like social network in showing the stylized fact Bass curves. Given more time and model runs the "normal" MCMC sampler would also find the superior solutions of the "optimized" model, which were found by the use of prior knowledge of social network structure.
- Characteristic path length (*SCL*): The "optimized" MCMC sampler with the Bass energy function seems to prefer social networks with a very stable path length of around 1.74. Since *SCL* measures a global property, the typical separation between two agents in the network, it also depends on the parameters *NCONS* and *NNBS* (section 4.5.1). Thus 1.74 was the closest value the optimizer could find under these given

conditions. Interestingly, these values are close to the “normal” Bass validated model which was not forced to generate small-world network properties. This substantiates the fact that the MCMC sampler in the “normal” model already preferred more realistic social networks in order to increase the probability of the occurrence of Bass curves. Figure 13 (upper picture) shows the rather stable development of both parameters over time for the last 50 % of samples.

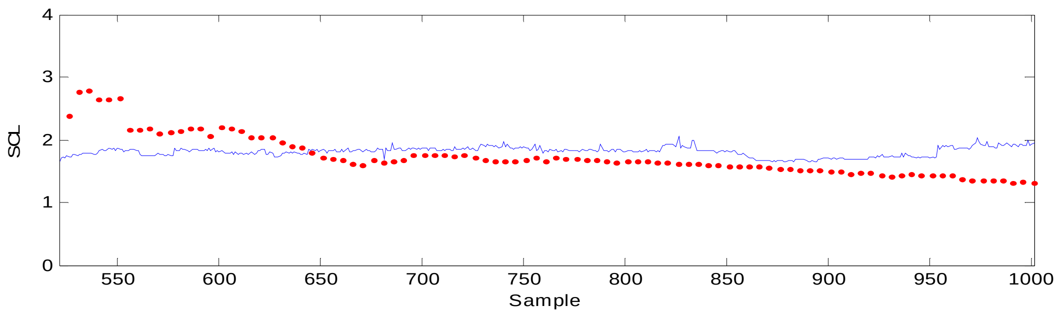
- Clustering coefficient or cliquishness (*SCC*): The mean values for *SCC* of both models are rather close again. Although the optimized model seems to develop its slightly but significantly higher average clustering coefficient (0.595) via a network structure using a higher average number of neighbors (~23) together with a lower proportion of clustering (0.10). The “normal” model seems to reach its neighborhood cliquishness via a higher proportion of clustering (0.26) but with a lower number of neighbors (~17). Since *PCLUS* and *SCC* are negatively correlated (cc=-0.51, p=0.053, see table 3) this is a consistent result. Figure 13 (lower picture) shows the development of the parameters over time. Although the mean values of both models are not that far off, the optimized model has a natural drift in the clustering coefficients towards its predetermined ideal value of 0.79 (from table 4, see figure 14 and 15).
- The social network properties are described by the number of consumers (*NCONS*), the average number of neighbors (*NNBS*), and the proportion of clustering (*PCLUS*). Interestingly both models are very close in their characteristic path length (*SCL*) and in their clustering coefficient or cliquishness (*SCC)*. But they seem to develop these properties in different ways. The optimized model develops a bigger market with more consumers and also more neighbors but exhibiting a lower *PCLUS* on average. The non-optimized model derives its properties of *SCL* and *SCC* by a higher *PCLUS* and lower number of consumers with a smaller neighborhood on average. This suggests that *SCC* and *SCL* may be a better measure for the description of social network structure than *PCLUS* alone and substantiates their usefulness. Given enough time and model runs the MCMC sampler would find all of these equivalent parametrizations.

Our results again support the importance of small-world network properties in consumer markets for the appearance of Bass curves. Interestingly this real-life observed topology emerges without being imposed explicitly in the “normal” Bass validated model.

Table 5: Comparison of the mean parameter values for the "normal" and "optimal" model including all samples (1000 per model). The values are shown with their 95 % confidence interval.[12]

Model comparison	Mean	$Mean_{opt}$	Std.Dev.	$Std.Dev._{opt}$
ρ_{max} (equation 14)*	0.744±0.014	0.747±0.011	0.2198	0.1824
Characteristic Path Length (*SCL*)**	1.811±0.007	1.74 ±0.03	0.0752	0.3298
Clustering Coefficient (*SCC*)**	0.411±0.009	0.595±0.009	0.1288	0.1371
Number of Consumers (*NCONS*)**	73.28 ±0.87	89.24 ±0.76	14.08	12.26
Number of Neighbors (*NNBS*)**	16.47 ±0.37	22.72 ±0.76	5.91	12.28
Proportion of Clustering (*PCLUS*)**	0.261±0.011	0.104±0.007	0.1837	0.112

Mean differences are significant at the *5 % level or at the **1 % level.



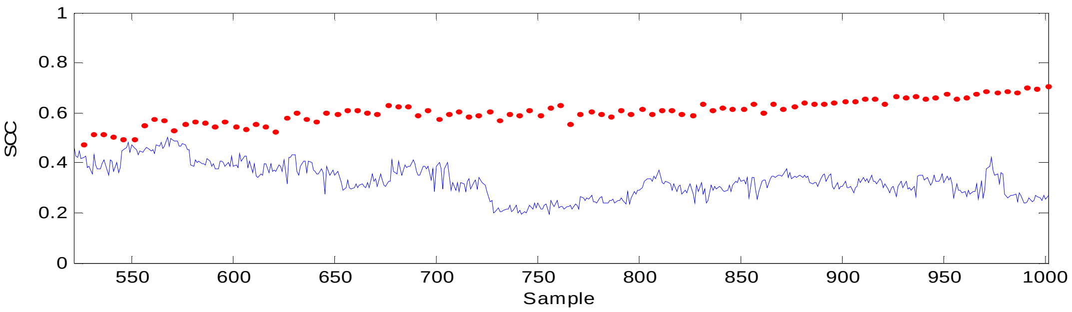


Figure 13: Development of the characteristic path length (*SCL*) and the social clustering coefficient (*SCC*) over time. The straight line represents the non-optimized model, the dotted line the optimized model values. The plot includes the last 50 % of samples of both models.

[12] The mean comparison is based on the nonparametric Wilcoxon ranksum test for independent samples.

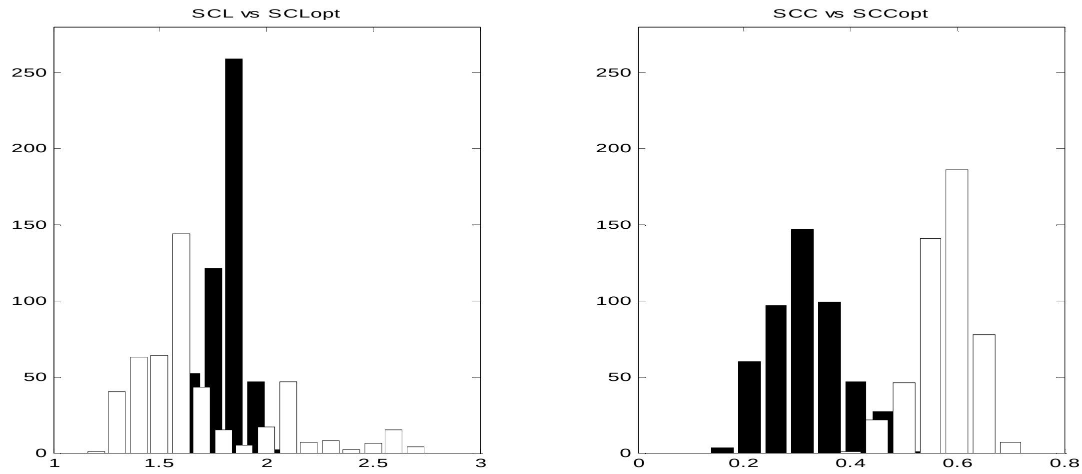


Figure 14: Comparison of the distribution of the characteristic path length (*SCL*, left) and the social clustering coefficient (*SCC*, right). The bars in the background represent the non-optimized model, the bars in the front the optimized model values. The plot includes the last 50 % of samples of both models.

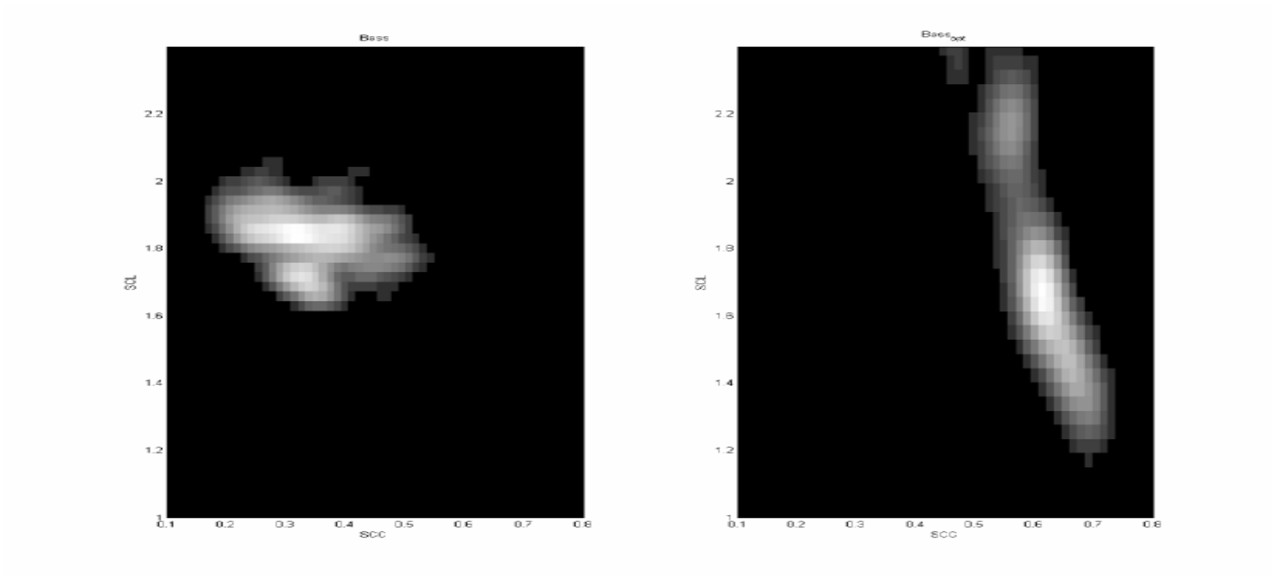


Figure 15: Comparison of the density plots of the characteristic path length (*SCL*) and the social clustering coefficient (*SCC*). The plot includes the last 50 % of samples of both models.

4.6 Market Share Forecasting

Our results so far have revealed the large impact of small-world network properties on the occurrence of Bass curves in the consumer markets (sections 4.4 and 4.5).

Because of this strong evidence multiple linear regressions are run to test the hypothesis that the model parameters can be used as predictors for future market share data. For more detailed comparisons the market share sample data is split into two and four parts. The model parameters were treated as predictors for the emerging future market share potential. In practice, these parameters could be derived, for example, from investigation of the target market's social network topology. Table 6 shows the significant regression coefficients (model parameters) and the goodness of fit of the multiple regression (adjusted R^2) which gives the best measure of the proportion of variance explained by the predictor variables. Table 6 also shows the regression coefficients of the parameters vs. the average of the whole, the average of the halves, and the average of the quarters of the market share time series. The regression results indicate that the parameters proportion of clustering (*PCLUS*, section 4.2.2), weight for social needs (*SNW*, equation 9), and the maximum uncertainty threshold (UNC_{MAX}, section 4.2.3) were able to explain a significant part of the whole, the 1st half, and especially the 1st quarter of the market share. While it should be very easy for marketing practitioners to measure the complexity of the target consumer market's structural complexity (*PCLUS*), this might be rather difficult for social needs (*SNW*) and the uncertainty threshold (UNC_{MAX}). The latter could be discovered by, for example, surveys of consumer needs and consumer satisfaction.

Furthermore, multiple regressions are run where 80 % of the model parameters and market share time series data were used as training set and 20 % as a test set. While table 6 shows the overall results and could be used as forecasting tool applicable to real market share data, table 7 gives the results of the predictions from estimated regression coefficients of the training set to the market share of the test set (out of sample forecast). Table 7 shows the regression coefficients and the goodness of fit of the regression (adjusted R^2) of the training set. σ_{res} gives the standard deviation of the residuals from prediction of the test set data and S.E. the standard error of the residuals (only the significant coefficients were used for market share forecasting).

The multiple regression results again support the relevant role of the complexity parameter (*PCLUS*). In addition it is shown that the weight of social network for the consumer agents' satisfaction (*SNW*), the maximum uncertainty threshold (UNC_{MAX}), and the social clustering coefficient (*SCC*) seem to be consistent and substantial predictors for the emerging market share of the integrated markets model.

Table 6: Multiple linear regression of the integrated markets model parameters against different proportions of average market share. Bold values indicate significant regression coefficients.

	Total average Market Share	1st half	2nd half	1st quarter	2nd quarter	3rd quarter	4th quarter
Adj. R^2	0.11	0.10	0.12	0.14	0.08	0.13	0.10
ν	0.62	0.51	0.73	0.31	0.72	0.63	0.83
Ncons	0.00	0.00	0.00	0.00	0.00	0.00	0.00
γ	-0.23	-0.22	-0.24	-0.35	-0.10	-0.13	-0.34
NN	0.00	0.00	0.00	-0.01	0.00	0.01	0.00
Pclus	**+0.43**	***0.57**	0.29	***0.70**	0.44	0.35	0.22
SNW	****0.82**	****0.88**	****0.77**	****0.95**	****0.80**	****0.76**	****0.78**
Psat	0.01	0.02	-0.01	0.04	0.00	0.00	-0.01
SAT_{min}	-0.10	-0.17	-0.03	***-0.29**	-0.05	-0.01	-0.05
UNC_{max}	****-0.63**	****-0.53**	****-0.73**	***-0.38**	****-0.68**	****-0.74**	****-0.71**
SCC	0.47	0.61	0.34	0.72	0.49	0.37	0.31
SCL	0.15	0.09	0.21	-0.06	0.24	0.31	0.12

regression coefficients significant at the +10 % level, *5 % level, **1 % level

Table 7: Multiple linear regression of the significant integrated markets model parameters against different proportions of average market share. Bold values indicate significant regression coefficients.

	Total average Market Share	1st half	2nd half	1st quarter	2nd quarter	3rd quarter	4th quarter
S.E.	0.12	0.12	0.12	0.12	0.13	0.12	0.10
σ_{res}	0.24	0.26	0.24	0.31	0.26	0.24	0.25
Adj. R^2	0.07	0.06	0.11	0.10	0.06	0.12	0.08
Pclus	****0.69**	****0.80**	***0.58**	***0.86**	***0.75**	****0.66**	**+0.49**
SNW	****0.76**	****0.80**	****0.73**	****0.81**	****0.80**	****0.72**	****0.75**
UNC_{max}	****-0.44**	**+-0.32**	****-0.57**	-0.11	****-0.54**	****-0.59**	****-0.55**
SCC	**+0.97**	**+1.15**	**+0.79**	**+1.20**	**+1.09**	**+0.83**	0.75

regression coefficients significant at the +10 % level, *5 % level, **1 % level

4.7 Discussion

"Why is network anatomy so important to characterize? Because structure always affects function. For instance, the topology of social networks affects the spread of information and disease, and the topology of the power grid affects the robustness and stability of power transmission."

Strogatz (2001)

The simulations in this chapter explore the impact of a cognitive and socially bounded agent based consumer model on the integrated markets model recently introduced by Sallans et al. (2003). First, a new consumer agent model is presented (sections 4.1 and 4.2). The model is embedded in a social structure based on "small-world network" principles (Milgram, 1967; Watts and Strogatz, 1998). Furthermore, the agents follow a rather simple cognitive decision structure, but one which is able to account for valid behavioral dynamics such as habits, imitation and social comparison processes (Janssen and Jager, 2000). In the second part of the chapter the underlying mechanisms of Bass curves are explored by validation of the model generated consumer markets data against empirically estimated time series. Therefore a recently presented Markov chain Monte Carlo (MCMC) method is used (see sections 4.3 and 4.4). The model produces consistent results as suggested by economic theory of network externalities. The results show that the emergence of Bass curves in consumer markets can be explained by the underlying consumer behavior: repetition and imitation behavior, which leads to increasing demand, and deliberation behavior, which refers to positive network externalities and leads to increased price sensitivity (section 4.4.2.3). Furthermore, a good model for Bass curves seems to involve firm agents which are good learners, operating in a market environment of stable complexity and small world network properties (sections 4.4.2.1 and 4.4.2.2).

The most striking fact that is documented is the importance of small-world network properties for the occurrence and prediction of the Bass curves in consumer markets. Interestingly, this real-life observed topology emerges in the integrated markets model by selection of the MCMC sampler without being imposed explicitly (section 4.5). This has implications for marketing practitioners. The results strongly suggest that it is useful to consider the structural properties of the target market, like cliquishness of the consumers' neighborhood or complexity of the market's social structure, and consumers' cognitive parameters, like their (social) needs and consumer satisfaction, to improve the quality of sales forecasts. These results may even have more general applications than just in combination with the Bass model. Further (empirical) research seems to be fruitful and necessary regarding the small-world properties in conjunction with marketing forecasts. A line of research could focus on the grouping of the markets by their social properties and relate these, for example, to innovation and imitation effects (the p and q parameters) of the Bass diffusion model. One could then analyze the impact of the structural properties on future market share development.

5 Synopsis

"Stock markets are psychology."
André Costolani

Modern financial economic theory is based on the assumption that the "representative agent" in the economy is rational in two ways: The representative agent makes decisions according to the axioms of expected utility theory and makes unbiased forecasts about the future. An extreme version of this theory assumes that every agent behaves in accordance with these assumptions. The argument that asset prices are set by rational investors is part of the grand oral tradition in economics and is often attributed to Milton Friedman, one of the greatest economists of the century. But the argument has fundamental problems. First, even if asset prices were set only by rational investors in the aggregate, knowing what individual investors are doing might still be of interest. Secondly, although the argument is intuitively appealing and reassuring, even when the relationship between two prices is easy to calculate and fixed by charter, prices can diverge and arbitrageurs are limited in their ability to restore the prices to parity (see section 2.1 on limits to arbitrage; Thaler, 1999b).

In recent years a body of evidence on security returns has presented a sharp challenge to the traditional view that securities are rationally priced to reflect all publicly available information. Furthermore, over the last decades, prominent researchers in both economics and psychology have criticized the view of neoclassical economics as psychologically unrealistic and proposed alternative assumptions. The underlying idea of this research is far too compelling to consider it temporary: the more realistic the assumptions about economic actors, the better the economics. Thus economists should aim to make assumptions about humans as psychologically realistic as possible.

Behavioral finance argues that empirical financial phenomena, like market anomalies, can plausibly be understood using models in which agents are not fully rational. In particular, behavioral finance has two main building blocks: limits to arbitrage and cognitive psychology or psychology of decision making. Limits to arbitrage refers to the effectiveness of arbitrage forces under different conditions. Cognitive biases refers to the huge psychological evidence documenting that people make systematic errors in the way they come to decisions under the condition of uncertainty. For example, they can be overconfident, they may put too much weight on recent experience, etc.

Perhaps the most important contribution of behavioral finance on the theory side is the careful investigation of the role of markets in aggregating a variety of behaviors. In particular, the publications show that in an economy which includes interacting rational and irrational traders, irrationality can have substantial and long-living impact on prices. One reason is that there are some psychological biases which virtually no one can escape. A second reason is that when traders are risk averse, prices reflect a weighted average of beliefs. Just as rational investors trade to arbitrage away mispricing, irrational investors trade to arbitrage away rational pricing. The presumption that rational beliefs will be victorious is based on the premise that wealth must flow from foolish to wise investors. But if investors are foolishly aggressive in their trading, they may earn higher rewards for bearing more risk (see, for example, DeLong et al., 1990b and 1991) or for exploiting information signals more aggressively (Hirshleifer and Luo, 2001). Thus irrational traders may gain from intimidating competing informed traders (Kyle and Wang, 1997). Indeed, one would expect wealth to flow from smart to dumb traders exactly when mispricing becomes more severe (Shleifer and Vishny, 1997; Xiong, 2000), which could contribute to self-feeding bubbles.

The thesis is organized as follows. In the first part of the thesis seminal theoretical and experimental work on behavioral finance and market anomalies is reviewed. Furthermore the underlying psychological mechanisms and empirical evidence of robust and systematic effects observed in experiments and over a wide area of financial markets data are emphasized (chapter 2).

In the second part of the thesis the novel methodology of agent-based computational economics is presented. This technique provides a framework to study an economic system in a controlled computational environment and is well suited for testing behavioral theories (see chapter 3). Moreover a significant feature of agent-based models is the ability to explicitly model “boundedly rational” agents (Simon, 1982). These agents have explicit limitations on their memory, knowledge or computational abilities. To simulate selected behavioral effects in an artificial economy an agent-based integrated markets model is developed (see chapter 4). The integrated markets model will serve as a testbed, which allows the investigation of market dynamics under conditions, which are too complex to be addressed analytically. The underlying behavioral, cognitive and social mechanisms are explored.

The integrated markets model's environment consists of a financial market with trading agents and a consumer market with cognitive and socially bounded consumer agents is developed. The markets are coupled via learning production firm agents offering their products and shares for sale. The consumer agents are embedded in a social structure based on "small-world network" principles. The cognitive model of the consumer agents enables them to make their decisions according to the behavior of the adjacent social neighborhood and based on the degree of satisfaction and uncertainty they are facing. The potential and limitations of the consumer agent model are explored by applying a recently introduced Markov chain Monte Carlo method. Therefore certain empirical phenomena or "stylized facts" are selected for reproduction within the simulation and the conditions of their occurrence are analyzed. It is shown that the properties of the social network structure and the sensitivity of the agents' cognitive decision making process (heuristics) contribute significantly or are, in fact, enabling the complex phenomena of Bass curves observed in consumer market scenarios. Furthermore, the results indicate that the structural properties of the emerged social networks are stable and match real-life social networks. Moreover it is shown that the network structure has a strong impact on the development of market share. Thus it is suggested to use the social network descriptive parameters, which could be discovered empirically, as predictive factors for market dynamic forecasts.

Nevertheless, anomalies in "real" financial markets must be viewed with caution and scepticism, as spurious mispricings can surface for a variety of reasons, such as errors in defining normal return, data mining, survivorship bias, small sample bias, selection bias, nonsynchronous trading, and misestimation of risk. Although anomalies should disappear in a close to efficient market, they may persist because they are not well understood, arbitrage is too costly, the profit potential is insufficient, trading restrictions exist, and behavioral biases exist (see also chapter 2.1 on limits to arbitrage). Documented and valid anomalies may still be unprofitable because the evidence is based on averages and may therefore include a large fraction of losers. Furthermore, the conditions responsible for the anomaly may change, and trading by informed investors may cause the anomaly to disappear (Singal, 2004).

According to Thaler (1999b) behavioral finance is no longer as controversial as it was. As financial economists become accustomed to think about the role of human behavior in driving stock prices, people will look back at the articles published in the past 15 years and wonder what the excitement was about. Moreover, Thaler predicts that in the not-too-distant future, the term “behavioral finance” will be correctly viewed as a redundant phrase. What other kind of finance is there? Economists will routinely incorporate as much “behavior” into their models as they observe in the real world. After all, to do otherwise would be irrational.

Acknowledgements

This work was partly funded by the Austrian Science Fund (FWF) under grant SFB #010: “Adaptive Information Systems and Modelling in Economics and Management Science". The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry for Education, Science and Culture and by the Austrian Federal Ministry for Transport, Innovation and Technology.

References

Abreu, D., and M. Brunnermeier (2002). Synchronization risk and delayed arbitrage. *Journal of Financial Economics 66, 341-360.*

Alicke, M. D., M. L. Klotz, D. L. Breitenbecher, T. J. Yurak, and D. S. Vredenburg (1995). Personal Contact, Individuation, and the Better-Than-Average Effect. *Journal of Personality and Social Psychology 68, 804-825.*

Alpert, M., and H. Raiffa (1982). A progress report on the training of probability assessors. In: Kahneman, D., P. Slovic and A. Tversky (eds.) *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge, Mass.: Cambridge University Press.

Arrow, K. (1986). Rationality of self and others. In: Hogarth, R. and M. Reder (eds.) *Rational Choice.* University of Chicago Press.

Arthur, W. B., J. H. Holland, B. LeBarron, R. Palmer, and P. Taylor (1997). Asset pricing under endogenous expectations in an artificial stock market. In W. B. Arthur, S. N. Durlauf, and Lane, D. A. (eds.). *The Economy as an Evolving Complex System II.* Reading, MA: Addison-Wesley.

Ayton, P., and D. Önkal (1997). Forecasting football fixtures: Confidence and judged proportion correct. Unpublished manuscript. Department of Psychology, The City University of London.

Barber, B., and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics 116(1), 261-292.*

Barberis, N., M. Huang and T. Santos (2001). Prospect theory and asset prices. *Quarterly Journal of Economics 116, 1-53.*

Barberis, N., A. Shleifer and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49, 307-345.*

Barberis, N., and R. Thaler (2003). A Survey of Behavioural Finance. In: Constantinides, G., M. Harris, and R. Stulz (eds.) *Handbook of the Economics of Finance.* North-Holland: Elsevier.

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science 15, 215-227.*

Baumann, A., R. Deber, and G. Thompson (1991). Overconfidence among Physicians and Nurses: The "Micro-Certainty, Macro-Uncertainty" Phenomenon. *Social Science and Medicine 32, 167-174.*

Bayes, T. (1763/1958). Studies in the History of Probability and Statistics: IX. Thomas Bayes's Essay Towards Solving a Problem in the Doctrine of Chances. *Biometrika 45, 296-315.*

Beardah, C. and M. Baxter (1996). Matlab routines for kernel density estimation and the graphical presentation of archaeological data. In H. Kammermans and K. Fennema (eds.), *Interfacing the Past: Computer Applications and Quantitative Methods in Archaeology 1995, Analecta Prehistorica Leidensia 28(1).* Leiden.

Beck-Bornholdt, H. P., and H. H. Dubben (2001). *Der Hund, der Eier legt: Erkennen von Fehlinformation durch Querdenken.* Hamburg: Rowohlt.

Beck-Bornholdt, H. P., and H. H. Dubben (2003). *Der Schein der Weisen: Irrtümer und Fehlurteile im täglichen Denken*. Hamburg: Rowohlt.

Bell, D. (1982). Regret in decision making under uncertainty. *Operations Research 30, 961-981.*

Benartzi, S. (2001). Excessive extrapolation and the allocation of 401(k) accounts to company stock. *Journal of Finance 56, 1747-1764.*

Benartzi, S., and R. H. Thaler (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics 110, 75-92.*

Benartzi, S., and R. H. Thaler (2001). Naive diversification strategies in defined contribution savings plans. *American Economic Review 91, 79-98.*

Bernard, V., and J. Thomas (1989). Post-earnings announcement drift: delayed price response or risk premium? *Journal of Accounting Research (Supplement), 1-36.*

Bernheim, B. D. (1996). Financial Illiteracy, Education, and Retirement Saving. University of Pennsylvania (Pension Research Council). Working Paper 96-7.

Bertsekas, D. P. and J. N. Tsitsiklis (1996). *Neuro-Dynamic Programming.* Belmont, MA: Athena Scientific.

Birnbaum, M. H. (1983). Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology 96, 85-94.*

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition.* Oxford University Press.

Black, F. (1986). Noise. *Journal of Finance 41, 529-543.*

Blume, M., and I. Friend (1975). The asset structure of individual portfolios with some implications for utility functions. *Journal of Finance 30, 585-604.*

Bodie, Z., A. Kane, A. J. Marcus (2002). *Investments.* (5th ed.). New York: McGraw-Hill.

Borges, B., D. G. Goldstein, A. Ortmann, and G. Gigerenzer (1999). Can Ignorance Beat the Stock Market? In Gigerenzer, G., P. M. Todd, and the ABC Research Group (1999). Simple heuristics that make us smart. (pp. 59-72) New York: Oxford University Press.

Boyd, M. (2001). On Ignorance, Intuition, and Investing: A Bear Market Test of the Recognition Heuristic. *The Journal of Behavioural Finance 2(3), 150-156.*

Brennan, M. J., and Y. Xia (1998). Resolution of a Financial Puzzle. *Finance.* Paper 28-98. URL: http://repositories.cdlib.org/anderson/fin/28-98

Brennan, M. J., and W. N. Torous (1999). Individual Decision-Making and Investor Welfare. Working paper. University of California, Los Angeles.

Brock, W. and C. Hommes (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control 22, 1235-1274.*

Brown, S., W. Goetzmann and S. Ross (1995). Survival. *Journal of Finance 50, 853-873.*

Buehler, R., D. Griffin and M. Ross (1994). Exploring the planning fallacy: why people underestimate their task completion times. *Journality of Personality and Social Psychology 67, 366-381.*

Camerer, C. (1995). Individual decision making. In: Kagel, J. H., and A. E. Roth (eds.) *The Handbook of Experimental Economics.* Princeton University Press.

Camerer, C., and R. Hogarth (1999). The effects of financial incentives in experiments: a review and capital-labor production framework. *Journal of Risk and Uncertainty 19, 7-42.*

Camerer, C., and D. Lovallo (1999). Overconfidence and excess entry: an experimental approach. *American Economic Review 89, 306-318.*

Campbell, J. Y. (1999). Asset prices, consumption and the business cycle. In: Taylor, J. and M. Woodford (eds.) *Handbook of Macroeconomics.* Amsterdam: Elsevier.

Campbell, J. Y., and J. Cochrane (1999). By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107, 205-251.*

Canner, N., G. N. Mankiw, and D. N. Weil (1997). An Asset Allocation Puzzle. *American Economic Review 87, 181-91.*

Charness, G., and M. Rabin (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics 117(3), 817-869.*

Chen, Z., W. Stanzl, and M. Watanabe (2001). Price Impact Costs and the Limit of Arbitrage. ICF Working Paper No. 00-66.

Chew, S. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the allais paradox. *Econometrica 51, 1065-1092.*

Chew, S. (1989). Axiomatic utility theories with the betweenness property. *Annals of Operations Research 19, 273-98.*

Chew, S., and K. MacCrimmon (1979). *Alpha-nu choice theory: an axiomatization of expected utility.* Working Paper. University of British Columbia, Vancouver, BC.

Christensen-Szalanski, J., and J. Bushyhead (1981). Physicians' Use of Probabilistic Information in a Real Clinical Setting. *Journal of Experimental Psychology: Human Perception and Performance 7, 928-935.*

Christensen-Szalanski, J., and C. Willham (1991). The Hindsight Bias: A Meta-Analysis. *Organizational Behaviour and Human Decision Processes 48(1), 147-68.*

Collins, J. J. and C. C. Chow (1998). It's a small world (News and Views). *Nature 393, 409-410*.

Combs, B., and P. Slovic (1979). Causes of death: Biased newspaper coverage and biased judgments. *Journalism Quarterly 56, 837-843.*

Cooper, A., C. Woo, and W. Dunkelberg (1988). Entrepreneurs' Perceived Chances for Success. *Journal of Business Venturing 3, 97-108.*

Cooper, I., and E, Kaplanis (1994). Home bias in equity portfolios, inflation hedging, and international capital market equilibrium. *Review of Financial Studies 7, 45-60.*

Cooper, J., and R. H. Fazio, (1984). A new look at dissonance theory. In: Berkowitz, L. (ed.) *Advances in Experimental Social Psychology, Vol. 17.* Orlando, Florida: Academic Press.

Cooper, M. J., R. C. Gutierrez, and H. Allaudeen (2004). Market States and Momentum. *The Journal of Finance 59(3),1345-1365.*

Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein. (2001). *Introduction to Algorithms.* MIT Press, 2nd edition.

Coval, J. D., and T. J. Moskowitz (2001). The geography of investment: informed trading and asset prices. *Journal of Political Economy 109, 811-841.*

Cox, J. C., and D. M. Grether (1996). The preference reversal phenomenon: response mode, markets and incentives. *Economic Theory 7, 381-405.*

Craik, F. I. M., and M. McDowd (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition 14, 474-479.*

D'Avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics 66, 271-306.*

Damasio, A. R. (1994). *Descartes' Error*. New York: Grosset Putnam.

Dangl, T., E. Dockner, A. Gaunersdorfer, A. Pfister, A. Soegner, and G. Strobl (2001). Adaptive erwartungsbildung und Finanzmarktdynamik. *Zeitschrift für betriebswirtschaftliche Forschung 53, 339-365.*

Daniel, K. D., and S. Titman (1999). Market efficiency in an irrational world. *Financial Analysts Journal 55, 28-40.*

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998). Investor Psychology and Security Market Under-and Overreactions. *Journal of Finance 53, 1839-1885.*

Daniel, K., D. Hirshleifer, and S. H. Teoh (2002). Investor psychology in capital markets: evidence and policy implications. *Journal of Monetary Economics 49(1), 139-209.*

Darley, J. M. and P. H. Gross (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology 44(1), 20-33.*

Daughety, A. F. (ed.) (1988). *Cournot Oligopoly: Characterization and Applications.* Cambridge: Cambridge University Press.

Dawes, R., and R. H. Thaler (1988). Cooperation. *Journal of Economic Perspectives 2(3), 187-197*

De Bondt, W. F. M., and R. H. Thaler (1985). Does the stock market overreact? *Journal of Finance 40(3), 793-805.*

De Bondt, W., and R. H. Thaler (1987). Further evidence on investor overreaction and stock market seasonality. *Journal of Finance 42, 557-581.*

De Bondt, W. F. M., and R. H. Thaler (1995). Financial Decision Making in Markets and Firms. In: Jarrow, R., and V. Maksimovic, and W. T. Ziemba (eds.) Finance. *Series of Handbooks in Operations Research and Management Science.* Elsevier: North Holland.

De Long, J. B., A. Shleifer, L. Summers and R. Waldmann (1990a). Noise trader risk in financial markets. *Journal of Political Economy 98, 703-738.*

De Long, J. B., A. Shleifer, L. Summers and R. Waldmann (1990b). Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance 45, 375-395.*

De Long, J. B., A. Shleifer, L. Summers, and R. J. Waldmann (1991). The survival of noise traders in financial markets. *Journal of Business 64, 1-20.*

De Sousa, R. (1987). *The Rationality of Emotion.* Cambridge, Mass.: MIT Press

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: weakening the independence axiom. *Journal of Economic Theory 40, 304-18.*

Delbecq-Derousné, J., J. F. Beauvois, and T. Shallice (1990). Preserved recall versus impaired recognition. *Brain 113, 1045-1074.*

Devenow, A., and I. Welch (1996). Rational herding in financial economics. *European Economic Review 40, 603-616.*

Dörner, D. (1997). *Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen.* Rowohlt.

Dowie, J. (1976) On the Efficiency and Equity of Betting Markets. *Economica, 43 139-150.*

Duncker, K. (1945). On problem solving (L. S. Lees, Trans.). *Psychological Monographs, 58 (5, Whole No. 270).*

Economides, N. (1996). The Economics of Networks. *International Journal of Industrial Organization 14(6), 673-699.*

Edwards, W. (1968). Conservatism in human information processing. In: Kleinmuntz, B. (ed.) *Formal representation of human judgment.* New York: Wiley.

Ekman, P. (1992). *Telling Lies.* New York: Norton.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics 75, 643-69.*

Elster, J. (1989). *Solomonic Judgements.* Cambridge, Mass.: Cambridge University Press.

Elster, J. (1996). Rationality and the Emotions. *The Economic Journal 106(438), 1386-1397.* http://www.geocities.com/hmelberg/elster/AR96RATE.HTM

Elster, J. (1998). Emotions and Economic Theory. *Journal of Economic Literature 36(1), 47-74.*

Employee Benefit Research Institute. (1997). *EBRI databook on employee benefits.* Washington, DC: EBRI.

Fama, E. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance 25, 383-417.*

Fama, E. (1998). Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics 49, 283-307.*

Fama, E., and K. French (1988). Dividend yields and expected stock returns. *Journal of Financial Economics 22, 3-25.*

Farmer, J. D., and A. W. Lo (1999). Frontiers of finance: Evolution and efficient markets. *Proceedings of the National Academy of Sciences 96, 9991-9992.*

Fehr, E., and S. Gächter (1998). The Economic Implications of Homo Reciprocans. *European Economic Review 42, 845-859.*

Fehr, E., and S. Gächter (2002). Altruistic punishment in humans. *Nature 415, 137-140.*

Fischhoff, B. (1975). Hindsight is not equal to Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty. *Journal of Experimental Psychology: Human Perception and Performance 1(3), 288-299.*

Fischhoff, B. (1982). For Those Condemned to Study the Past: Heuristics and Biases in Hindsight. In: Kahneman, D., P. Slovic, and A. Tversky (eds.) *Judgment under Uncertainty: Heuristics and Biases.* Cambridge and New York: Cambridge University Press.

Fischhoff, B., P. Slovic and S. Lichtenstein (1977). Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Pyschology: Human Perception and Performance 3, 552-564.*

Fishburn, P. C. (1970). *Utility Theory for Decision Making.* New York: John Wiley and Sons.

Forsythe, G. E., M. A. Malcolm, and C. B., Moler (1976). *Computer Methods for Mathematical Computations.* Prentice-Hall.

Fox, C., and A. Tversky (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics 110, 585-603.*

Frank, J. (1935). Some Psychological Determinants of the Level of Aspiration. *American Journal of Psychology 47, 285-293.*

Frank, R. (1988). *Passions within Reason.* New York: Norton.

Friedman, M. (1953). *The case for flexible exchange rates.* In: Essays in Positive Economics. University of Chicago Press.

Friend, I., and M. E. Blume (1975). The Asset Structure of Individual Portfolios and Some Implications for Utility Functions. *Journal of Finance, 30(2), 585-603.*

Froot, K.A., and E.A. Dabora (1999). How are stock prices affected by the location of trade? *Journal of Financial Economics 53, 189-216.*

Fryer, R. G., and M. O. Jackson (2003). *Categorical Cognition: A Psychological Model of Categories and Identification in Decision Making.* National Bureau of Economic Research (NBER) Working Paper No. 9579.

Galef, B. G. (1987). Social influences on the identification of toxic foods by Norway rats. *Animal Learning and Behaviour 15, 327-332.*

Galef, B. G., McQuoid, L. M., and Whiskin, E. E. (1990). Further evidence that Norway rats do not socially transmit learned aversions to toxic baits. *Animal Learning Behaviour 18, 199-205.*

Genesove, D., and C. Mayer (2001). *Loss aversion and seller behavior: evidence from the housing market.* National Bureau of Economic Research (NBER) Working Paper 8143.

Gervais, S., and T. Odean (2001). Learning to be overconfident. *Review of Financial Studies 14, 1-27.*

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review 98, 254-267.*

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In: K. I. Manktelow and D. E. Over (eds.) *Rationality: Psychological and philosophical perspectives.* London: Routledge.

Gigerenzer, G. (1996). On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky. *Psychological Review 103(3), 592-596.*

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world.* New York: Oxford University Press.

Gigerenzer, G., and D. G. Goldstein (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103(4), 650-669.*

Gigerenzer, G., and U. Hoffrage (1995). How to improve Bayesian reasoning without instruction. Frequency formats. *Psychological Review 102, 684-704.*

Gigerenzer, G., and D. J. Murray (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Erlbaum.

Gigerenzer, G., P. M. Todd, and the ABC Research Group (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Gilovich, T., R. Vallone and A. Tversky (1985). The hot hand in basketball: on the misperception of random sequences. *Cognitive Psychology 17, 295-314.*

Gintis, H. (2000). *Game theory evolving.* Princeton NJ: Princeton University Press.

Grajek, M. (2002). Identification of Network Externalities in Markets for Non-Durables. Discussion Paper FS IV 02-32, Wissenschaftszentrum Berlin.

Groner, M., R. Groner, and W. F. Bischof (1983). Approaches to heuristics: A historical review. In: R. Groner, M. Groner, and W. F. Bischof (eds.) *Methods of heuristics.* Hillsdale, NJ: Erlbaum.

Goldstein, D. G., and G. Gigerenzer (1999). The recognition heuristic: How ignorance makes us smart. In Gigerenzer, G., P. M. Todd, and the ABC Research Group (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Goldstein, D. G., and G. Gigerenzer (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109, 75-90.*

Goleman, D. (1995). *Emotional Intelligence.* New York: Bantam Books.

Graham, J. R. (1999). Herding among investment newsletters: theory and evidence. *Journal of Finance 54, 237-268.*

Greenwald, A. (1980). The Totalitarian Ego: Fabrication and Revision of Personal History. *American Psychologist 35, 603-618.*

Griffin, D., and A. Tversky (1992) The Weighing of Evidence and the Determinants of Confidence. *Cognitive Psychology 24, 411-435.*

Griffiths, T. L., and J. B. Tenenbaum (2004). From algorithmic to subjective randomness. *Advances in Neural Information Processing Systems 16.*

Grinblatt, M., and T. Moskowitz (1999). The cross-section of expected returns and its relation to past returns. Working Paper. University of Chicago.

Grinblatt, M., S. Titman, and R. Wermers (1995). Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. *American Economic Review 85, 1088-1105.*

Grundy, B., and J. S. Martin (2001). Understanding the nature of the risks and the source of the rewards to momentum investing. *Review of Financial Studies 14, 29-78.*

Gul, F. (1991). A theory of disappointment in decision making under uncertainty. *Econometrica 59, 667-686.*

Hagan, M. T., H. B. Demuth, and M. H. Beale (1996). *Neural Network Design.* Boston, MA: PWS Publishing.

Hammond, K. R. (1996). *Human judgment and social policy. Irreducible uncertainty, inevitable error, unavoidable injustice.* New York: Oxford University Press.

Harris, L., and E. Gurel (1986). Price and volume effects associated with changes in the S&P 500: new evidence for the existence of price pressure. *Journal of Finance 41, 851-860.*

Hausch, D, W. Ziemba, and M. Rubinstein (1981). Efficiency of the Market for Racetrack Betting. *Management Science 27, 1435-1452.*

Hawkins, S. A., and R. Hastie (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin 107(3), 311-327.*

Heath, C., and A. Tversky (1991). Preference and belief: ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty 4, 5-28.*

Helson, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior.* New York: Harper and Row.

Heuer, R. J., Jr. (1999). *Psychology of Intelligence Analysis.* Washington, DC: Center for the Study of Intelligence, Central Intelligence Agency. URL: http://www.cia.gov/csi/books/19104

Hirshleifer, J. (1987). On the emotions as guarantors of threats and promises. In: Dupre, J. (ed.) *The Last on the Best.* Cambridge, Mass.: MIT Press.

Hirshleifer, D. (2001): Investor psychology and asset pricing. *Journal of Finance 56, 1533-1597.*

Hirshleifer, D., and G. Y. Luo (2001). On the survival of overconfident traders in a competitive security market. *Journal of Financial Markets 4, 73-84.*

Hirshleifer, D., and T. Shumway (2003). Good day sunshine: stock returns and the weather. *The Journal of Finance 58(3), 1009-1032.*

Hirshleifer, D., and S. H. Teoh (2003). Herd Behaviour and Cascading in Capital Markets: A Review and Synthesis. *European Financial Management 9(1), 25-66.*

Hoffrage, U. (1995). The adequacy of subjective confidence judgments: Studies concerning the theory of probabilistic mental models. Unpublished doctoral dissertation, University of Salzburg, Austria.

Hommes, C. H. (2002). Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proceedings of the National Academy of Sciences U.S.A., 99(3), 7221-7228.*

Hong, H., T. Lim and J. Stein (2000). Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance 55, 265-295.*

Hong, H., J. Kubik, and J. C. Stein (2001). Social interactions and stock market participation. Working paper. Stanford Business School.

Huberman, G. (1999). *Familiarity breeds investment.* Working Paper 97-04. Columbia University.

Irwin, F. (1953). Stated Expectations as Functions of Probability and Desirability of Outcomes. *Journal of Personality 21, 329-335.*

Ito, T. (1990). Foreign Exchange Rate Expectations: Micro Survey Data. *American Economic Review 80, 434-449.*

Jacoby, L. J., V. Woloshyn, and C. Kelley (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology 118, 115-125.*

Janssen, M.A., and W. Jager (2000). Fashions, habits and changing preferences: a simulation of psychological factors affecting market dynamics. Research memorandum 2000-31. Free University of Amsterdam. Faculty of Economics and Econometrics.

Jegadeesh, N., and S. Titman (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance 48, 65-91.*

Jensen, M. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics, 6, 95-101.*

John Hancock Financial Services. (1995). *1995 Gallup Survey of Defined Contribution Plan Participants.* November 1995.

Kahneman, D., and S. Frederick (2002). Representativeness revisited: attribute substitution in intuitive judgment. In: T. Gilovich, D. Griffin, and D. Kahneman (eds.) *Heuristics and biases.* New York: Cambridge University Press.

Kahneman, D., and M. Riepe (1998). Aspects of investor psychology. *Journal of Portfolio Management, 24, 52-65.*

Kahneman, D., and A. Tversky. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology 3, 430-454.*

Kahneman, D., and A. Tversky. (1974). Judgement under uncertainty: heuristics and biases. *Science 185, 1124-1131.*

Kahneman, D., and A. Tversky (1979). Prospect Theory: An Analysis of Decisions under Risk. *Econometrica 47, 313-327.*

Kahneman, D., and Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review 103, 582-591.*

Kahneman, D., P. Slovic and A. Tversky (1982). *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge University Press.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy 98(6), 1325-48.*

Kamstra, M., L. Kramer, and M. Levi (2000). Losing Sleep at the Market: The Daylight Saving Anomaly. *American Economic Review 90(4), 1005-1011.*

Kamstra, M., L. Kramer, and M. Levi (2002). Losing Sleep at the Market: The Daylight Saving Anomaly: Reply. *American Economic Review 92(4), 1257-1263.*

Kamstra, M., L. Kramer, and M. Levi (2003). Winter Blues: A SAD Stock Market Cycle. *American Economic Review 93(1), 324-343.*

Katz, M. L. and C. Shapiro (1985). Network Externality, Competition, and Compatibility. *American Economic Review 75(3), 424-440.*

Kaul, A., V. Mehrotra and R. Morck (2000). Demand curves for stocks do slope down: new evidence from an index weights adjustment. *Journal of Finance 55, 893-912.*

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist 28, 107-128.*

Keren, G. (1987). Facing Uncertainty in the Game of Bridge: A Calibration Study. *Organizational Behaviour and Human Decision Processes 39, 98-114.*

Kidd, J. (1970). The Utilization of Subjective Probabilities in Production Planning. *Acta Psychologica 34, 338-347.*

Knight, F. (1921). *Risk, Uncertainty and Profit.* Boston, New York: Houghton Mifflin.

Kreps, D. (1988). *Notes on the Theory of Choice.* Westview Press.

Kunda, Z. (1987). Motivated Inference: Self-serving Generation and Evaluation of Causal Theories. *Journal of Personality and Social Psychology 53, 636-647.*

Kyle, A., and F. A. Wang (1997). Speculation duopoly with agreement to disagree: can overconfidence survive the market test? *Journal of Finance 52, 2073-2090.*

Lakonishok, J., A. Shleifer and R. Vishny (1994). Contrarian investment, extrapolation and risk. *Journal of Finance 49, 1541-1578.*

Lamont, O., and R. Thaler (2003). Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy 111, 227-268.*

Langer, E., and J. Roth (1975). Heads I Win, Tails It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task. *Journal of Personality and Social Psychology 32, 951-955.*

Laplace, P. S. (1774/1986). Memoir on the Probability of the Causes of Events. *Statistical Science, 1(3), 364-378.*

Larwood, L., and W. Whittaker (1977). Managerial myopia: self-serving biases in organizational planning. *Journal of Applied Psychology 62, 94-198.*

LeBaron, B. (2000). Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control 24, 679-702.*

LeBaron, B. (2004). Agent-based Computational Finance. Working Paper. Brandeis University.

LeDoux, J. (1996). *The Emotional Brain.* New York: Simon and Schuster.

Lee, C., A. Shleifer and R. Thaler (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance 46, 75-110.*

Leibenstein, H. (1950). Bandwagon, snob, and Veblen Effects in the theory of consumers' demand. *Quarterly Journal of Economics 64, 183-207.*

LeRoy, S., and R. Porter (1981). The present-value relation: tests based on implied variance bounds. *Econometrica 49, 97-113.*

Lewis, K. K. (1999). Trying to explain home bias in equities and consumption. *Journal of Economic Literature 37, 571-608.*

Lichtenstein, S., B. Fischhoff, and L. Phillips (1982). Calibration of Probabilities: The State of the Art to 1980. In: Kahneman, D., P. Slovic, and A. Tversky (eds.) *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge and New York: Cambridge University Press.

Lo, A. W. and A. C. MacKinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies 3, 175-206.*

Loomes, G., and R. Sugden (1982). Regret theory: an alternative theory of rational choice under uncertainty. *The Economic Journal 92, 805-824.*

Loomes, G., C. Starmer, and R. Sugden (2003). Do Anomalies Disappear in Repeated Markets? *The Economic Journal 113, C153-C166.*

Lord, C., L. Ross and M. Lepper (1979). Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology 37, 2098-2109.*

Lux, T. (1998). The socio-economic dynamics of speculative markets: Interacting agents, chaos, and the fat tails of return distributions. *Journal of Economic Behavior and Organization 33, 143-165.*

Lux, T., and M. Marchesi (2000). Volatility clustering in financial markets: A micro-simulation of interacting agents. *International Journal of Theoretical and Applied Finance 3, 675-702.*

MacKinley, A. C. (1995). Multifactor Models do not Explain Deviations from the CAPM. *Journal of Financial Economics 38, 3-28.*

Macrae, N., and G. Bodenhausen (2000). Social Cognition: Thinking Categorically About Others. *Annual Review of Psychology 51, 93-120.*

Mahajan, V., E. Muller, and F. Bass (1990). New Product Diffusion Models in Marketing: A Review and Directions for Research. *Journal of Marketing 54:1, 1-26.*

Markman, A. B. and D. Gentner, D. (2001). Thinking. *Annual Review of Psychology 52, 223-247.*

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy 60, 151-158.*

Marks, R. (1951). The Effect of Probability, Desirability, and "Privilege" on the Stated Expectations of Children. *Journal of Personality 19, 332-351.*

McKenna, P., and E. K. Warrington (1980). Testing for nominal dysphasia. *Journal of Neurology, Neurosurgery and Psychiatry 43, 781-788.*

Mehra, R., and E. Prescott (1985). The equity premium: a puzzle. *Journal of Monetary Economics 15, 145-161.*

Mellers, B. A., A. Schwartz and A. D. J. Cooke (1998). Judgment and decision making. *Annual Review of Psychology 49, 447-477.*

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica, 41, 867-887.*

Merton, R. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance 42, 483-510.*

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics 21, 1087-1092.*

Milgram, S. (1967). The small world problem. *Psychology Today, 61(1).*

Miller, D., and M. Ross (1975). Self-serving Biases in Attribution of Causality: Fact or Fiction? *Psychological Bulletin 82, 213-225.*

Mitchell, M., T. Pulvino and E. Stafford (2002). Limited arbitrage in equity markets. *Journal of Finance 57, 551-584.*

Mitchell, O. S., and S. P. Zeldes (1996). Social Security Privatization: A Structure for Analysis. *American Economic Review, 86(2), 363-67.*

Modigliani, F., and R. Cohn (1979). Inflation, rational valuation and the market. *Financial Analysts Journal 35, 24-44.*

Morris, S. A. and D. Pratt (2003). Analysis of the Lotka-Volterra competition equations as a technological substitution model. *Technological Forecasting and Social Change 70(2), 103-133.*

Mullainathan, S. (2001). Thinking through Categories. Working Paper. Cambridge, MA: MIT.

Mullainathan, S., and R. Thaler (2000). Behavioural Economics. National Bureau of Economic Research (NBER) Working Paper 7948.

Murphy, A., and R. Winkler (1984). Probability Forecasting in Meteorology. *Journal of the American Statistical Association 79, 489-500.*

Neale, M., and M. Bazerman (1990). *Cognition and Rationality in Negotiation.* New York: The Free Press.

von Neumann, J., and O. Morgenstern (1944). *Theory of Games and Economic Behaviour.* Princeton, NJ: Princeton University Press.

Newell, B. R., and D. R. Shanks (2004). On the Role of Recognition in Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition 30(4), 923-935.*

Newell, B. R., Weston, N. J., and D. R. Shanks (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone "takes-the-best". *Organizational Behaviour and Human Decision Processes 91(1), 82-96.*

Newman, M. E. J., D. J. Watts, and S. H. Strogatz (2002). Random graph models of social networks. *Proc. Natl. Acad. Sci. USA 99, 2566-2572.*

Odean, T. (1998a). Are investors reluctant to realize their losses? *Journal of Finance 53(5), 1775-1798.*

Odean, T. (1998b). Volume, Volatility, Price, and Profit When All Traders Are Above Average. *The Journal of Finance 53(6), 1887-1934.*

Odean, T. (1999): Do investors trade too much? *American Economic Review 89, 1279-1298.*

Ofek, E., and M. Richardson (2003). Dot-com mania: market inefficiency in the internet sector. *Journal of Finance 58, 1113-1137.*

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): rethinking the recognition heuristic. *Cognition, 90, B1-B9.*

Oskamp, S. (1965) Overconfidence in Case-Study Judgments. *Journal of Consulting Psychology 29, 261-265.*

Perkins, D. N. (1981). *The mind's best work.* Cambridge: Harvard University Press.

Peters, E., and P. Slovic (1996). The role of affect and worldviews as orienting dispositions in the perception and acceptance of nuclear power. *Journal of Applied Social Psychology 26, 1427-1453.*

Piattelli-Palmarini, M. (1994). *Inevitable illusions. How mistakes of reason rule our minds.* New York: Wiley.

Plous, S. (1991). Biases in the assimilation of technological breakdowns: Do accidents make us safer? *Journal of Applied Social Psychology 21, 1058-1082.*

Polya, G. (1945). *How To Solve It: A New Aspect of Mathematical Method.* Princeton University Press.

Polya, G. (1954). *Mathematics and plausible reasoning. Vol. 1: Induction and analogy in mathematics.* Princeton, NJ: Princeton University Press.

Prechter, R. R. (2001). Unconscious Herding Behaviour as the Psychological Basis of Financial Market Trends and Patterns. *The Journal of Behavioural Finance 2(3), 120-125.*

de Quervain, D. J. F., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr (2004). The Neural Basis of Altruistic Punishment. Science 305, 1254-1258.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour and Organization 3, 323-343.*

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature 36, 11-46.*

Rabin, M. (2000). Risk aversion and expected utility theory: a calibration theorem. *Econometrica 68, 1281-1292.*

Rabin, M. (2002a). A perspective on psychology and economics. *European Economic Review 46, 657-685.*

Rabin, M. (2002b). Inference by believers in the law of small numbers. *Quarterly Journal of Economics 117, 775-816.*

Rabin, M., and J. Schrag (1999). First impressions matter: A model of confirmatory bias. *Quarterly Journal of Economics 114, 37-82.*

Ritter, J. R. (2003). Behavioural finance. *Pacific-Basin Finance Journal 11(4), 429-437.*

Ritter, J. R., and R. Warr (2002). The decline of inflation and the bull market of 1982-1999. *Journal of Financial and Quantitative Analysis 37, 29-61.*

Rogers, E. M. (1983). *Diffusions of Innovations.* (3rd ed.) New York: The Free Press.

Roll, R. (1983). Vas ist das? The turn-of-the-year effect and the return premia of small firms. *Journal of Portfolio Management 9, 18-28.*

Ross, S. (2001). *Lectures Notes on Market Efficiency*. MIT, Cambridge, MA.

Rubinstein, M. (2001). Rational markets: yes or no? The affirmative case. *Financial Analysts Journal 57(3), 15-29.*

Russo, J. E., and P. J. H. Schoemaker (1992). Managing Overconfidence. *Sloan Management Review 33, 7-17.*

Sallans B., G. Dorffner, and A. Karatzoglou (2002). The Dynamics of Interacting Markets: First Results. Technical Report, Österreichisches Forschungsinstitut für Artificial Intelligence, Wien, TR-2002-01.

Sallans B., A. Pfister, A. Karatzoglou, and G. Dorffner (2003). Simulation and Validation of an Integrated Markets Model. *The Journal of Artificial Societies and Social Simulation 6(4).*

Samuelson, W., and R. J. Zeckhauser (1988). Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty, 1(1), 7-59.*

Sargent, T. (1993). *Bounded Rationality in Macroeconomics.* Oxford University Press.

Sarin, R. K., and M. Weber (1993). Effects of ambiguity in market experiments. *Management Science 39, 602-615.*

Saunders, E. M. J. (1993). Stockprices and Wall Street weather. *American Economic Review 83, 1337-1345.*

Savage, L. J. (1954). *The Foundations of Statistics.* New York: Wiley.

Schacter, D. L., and E. Tulving (1994). What are the memory systems of 1994? In: Schacter, D. L. and E. Tulving (eds.) *Memory systems.* Cambridge, MA: MIT Press.

Sherali, H. (1984). A multiple leader Stackelberg model and analysis. *Operations Research, 32, 390-404.*

Shirreff, D. (1999). Lessons From The Collapse Of Hedge Fund, Long-Term Capital Management. Case Studies from the International Financial Risk Institute. URL: http://riskinstitute.ch/146490.htm.

Schonfield, D., and B. Robertson (1966). Memory storage and aging. *Canadian Journal of Psychology 20, 228-236.*

Schwert, G. W. (2003). Anomalies and Market Efficiency. In: Constantinides, G., M. Harris, and R. Stulz (eds.) *Handbook of the Economics of Finance.* North-Holland: Elsevier.

Segal, U. (1987). Some remarks on Quiggin's anticipated utility. *Journal of Economic Behaviour and Organization 8, 145-154.*

Segal, U. (1989). Anticipated utility: a measure representation approach. *Annals of Operations Research 19, 359-373.*

Semenza, C., and T. M. Sgaramella (1993). Production of proper names: A clinical case study of the effects of phonemic cueing. In: G. Cohen and D. M. Burke (eds.) *Memory for proper names*. Hove, United Kingdom: Erlbaum.

Semenza, C., and M. Zettin (1989). Evidence from aphasia for the role of proper names as pure referring expressions. *Nature 342, 678-679.*

Shefrin, H. and M. Statman (1985). The disposition effect to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance 40(3), 777-790.*

Shefrin, H. and M. Statman (2000). Behavioural Portfolio Theory. *Journal of Financial and Quantitative Analysis 35, 127-151.*

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review 71, 421-436.*

Shiller, R. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity 2, 457-498.*

Shleifer, A. (1986). Do demand curves for stocks slope down? *Journal of Finance 41, 579-90.*

Shleifer, A., and L. Summers (1990). The noise trader approach to finance. *Journal of Economic Perspectives 4, 19-33.*

Shleifer, A., and R. Vishny (1997). The limits of arbitrage. *Journal of Finance 52, 35-55.*

Shleifer, A. (2000). *Inefficient Markets.* New York: Oxford University Press.

Shogren, J., S. Cho, C. Koo, J. List, C. Park, P. Polo, and R. Wilhelmi (2001). Auction mechanisms and the measurement of WTP and WTA. *Resource and Energy Economics 23, 97-109.*

Siegel, J. J. (1998). *Stocks for the Long Run.* (2nd ed.). New York: McGraw-Hill.

Simon, H. A. (1955). A Behavioural Model of Rational Choice. *Quarterly Journal of Economics 69, 99-118.*

Simon, H. A. (1982). *Models of Bounded Rationality, Vol 2: Behavioural Economics and Business Organization.* Cambridge, MA: The MIT Press.

Singal, V. (2004). *Beyond the Random Walk: A Guide to Stock Market Anomalies and Low Risk Investing.* New York: Oxford University Press

Slovic, P., and S. Lichtenstein (1971). Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment. *Organizational Behaviour and Human Performance 6(6), 649-744.*

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations.* Reprinted: Campbell, R. H., and Skinner, A. S. (eds). Oxford: Clarendon Press. 1976.

Squire, L. R., B. Knowlton, and G. Musen (1993). The structure and organization of memory. *Annual Review of Psychology 44, 453-495.*

von Stackelberg, H. (1934). *Marktform und Gleichgewicht.* Vienna: Julius Springer.

Staël von Holstein, C. (1972). Probabilistic Forecasting: An Experiment Related to the Stock Market. *Organizational Behaviour and Human Performance 8, 139-158.*

Statman, M. (1997). *Behavioural Finance.* Contemporary Finance Digest.

Strogatz, S. H. (2001). Exploring complex networks. *Nature 410, 268-276.*

Summers, L. (1986). Does the stock market rationally reflect fundamental values? *Journal of Finance 41, 591-601.*

Sutton, R. S. and A. G. Barto (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: The MIT Press.

Svenson, O. (1981). Are We All Less Risky and More Skillful Than Our Fellow Drivers? *Acta Psychologica 47, 143-148.*

Swinkels, L. (2004). Momentum investing: A survey. *Journal of Asset Management 5(2), 120-143.*

Taylor, S., and J. Brown (1988). Illusion and Well-Being: A Social Psychological Perspective on Mental Health. *Psychological Bulletin 103, 193-210.*

Tesfatsion, L. (2002). Agent-based computational economics: Growing economies from the bottom up. *Artificial Life 8(1), 55-82.*

Tesfatsion, L. (2003). Agent-based computational economics. ISU Economics Working Paper 1. Department of Economics, Iowa State University.

TIAA-CREF. (1997). Premium Allocations and Accumulations in TIAA-CREF - Trends in Participant Choices Among Asset Classes and Investment Accounts. *Research Dialogues 51.*

Thaler, R. H. (1980). Toward a Positive Theory of Consumer Choice. *Journal of Economic Behaviour and Organization, 1(1), 39-60.*

Thaler, R. H. (1985). Mental accounting and consumer choice. *Marketing Science 4(3), 199-214.*

Thaler, R. H. (1990). Anomalies: Saving, fungibility, and mental accounts. *Journal of Economic Perspectives, 4(1), 193-205.*

Thaler, R. H. (1999a). Mental Accounting Matters. *Journal of Behavioural Decision Making 12, 183-206.*

Thaler, R. H. (1999b). The End of Behavioral Finance. Financial Analysts Journal 12-17.

Thaler, R. H. (2000). Mental accounting matters. In: Kahneman, D., and A. Tversky (eds.) *Choice, Values and Frames.* Cambridge, UK: Cambridge University Press.

Thaler, R. H., and E. Johnson (1990). Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Management Science 36, 643-660.*

Tobin, J. (1958). Liquidity Preference as Behaviour towards Risk. *Review of Economic Studies, 25, 68-85.*

Tversky, A. and D. Griffin (1991). Endowment and contrast in judgments of well-being. In: Zeckhauser, R. (ed.) *Strategy and Choice.* Cambridge, Mass.: MIT Press.

Tversky, A., and D. Kahneman (1971). Belief in the law of small numbers. *Psychological Bulletin 76, 105-110.*

Tversky, A., and D. Kahneman (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology 4, 207-232.*

Tversky, A., and D. Kahneman (1986). Rational choice and the framing of decisions. *Journal of Business 59, 251-278.*

Tversky, A., and D. Kahneman (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Quarterly Journal of Economics 106(4), 1039-61.*

Tversky, A., and D. Kahneman (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5, 297-323.*

Van den Bulte, C. (2002). Want to Know How Diffusion Speed Varies Across Countries and Products? Try Using a Bass Model. *Visions, 26 (4), 12-15.*

Villejoubert, G. (2004). Could they have known better? Review of the special issue of Memory Hindsight Bias. *Applied Cognitive Psychology (forthcoming).*

Wagenaar, W., and G. Keren (1986). Does the Expert Know? The Reliability of Predictions and Confidence Ratings of Experts. In: Hollnagel, E., G. Maneini, and D. Woods (eds.) *Intelligent Decision Support in Process Environments.* Berlin: Springer.

Warrington, E. K., and R. A. McCarthy (1988). The fractionation of retrograde amnesia. *Brain and Cognition 7, 184-200.*

Wasserman, S. and K. Faust (1994). *Social Network Analysis: Methods and Applications.* Cambridge: Cambridge University Press.

Watt, R. (2002). A Generalized Oligopoly Model. *Metroeconomica 53 (1), 46-55.*

Watts, D. J., and S. H. Strogatz (1998). Collective dynamics of 'small-world' networks. *Nature 393, 440-442.*

Weinstein, N. (1980). Unrealistic Optimism About Future Life Events. *Journal of Personality and Social Psychology 39, 806-820.*

Welch, I. (2000). Herding among security analysts. *Journal of Financial Economics 58, 369-396.*

Wilson, A. (2004). Bounded Memory and Biases in Information Processing. Job Market Paper. Princeton University.

Winter, S. G., Jr. (1964). Economic 'Natural Selection' and the Theory of the Firm. *Yale Econ. Essays, 4(1), 225-72.*

Wurgler, J., and K. Zhuravskaya (2002). Does arbitrage flatten demand curves for stocks? *Journal of Business 75, 583-608.*

Xiong, W. (2000). Convergence trading with wealth effects: an amplification mechanism in financial markets. Mimeo, Duke University.

Yaari, M. (1987). The dual theory of choice under risk. *Econometrica 55, 95-115.*

Yates, J. F. (1990). *Judgment and Decision Making.* Englewood Cliffs, New Jersey: Prentice Hall.

Zweig, J. (1998). Five Investing Lessons from America's Top Pension Fund. *Money, January, 115-18.*

List of Figures

List of Tables