Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors

BRAD M. BARBER and TERRANCE ODEAN*

ABSTRACT

Individual investors who hold common stocks directly pay a tremendous performance penalty for active trading. Of 66,465 households with accounts at a large discount broker during 1991 to 1996, those that trade most earn an annual return of 11.4 percent, while the market returns 17.9 percent. The average household earns an annual return of 16.4 percent, tilts its common stock investment toward high-beta, small, value stocks, and turns over 75 percent of its portfolio annually. Overconfidence can explain high trading levels and the resulting poor performance of individual investors. Our central message is that trading is hazardous to your wealth.

In 1996, approximately 47 percent of equity investments in the United States were held directly by households, 23 percent by pension funds, and 14 percent by mutual funds (Securities Industry Fact Book, 1997). Financial economists have extensively analyzed the return performance of equities managed by mutual funds. There is also a fair amount of research on the performance of equities managed by pension funds. Unfortunately, there is little research on the return performance of equities held directly by households, despite their large ownership of equities.

* Graduate School of Management, University of California, Davis. We are grateful to the discount brokerage firm that provided us with the data for this study. We appreciate the comments of Christopher Barry, George Bittlingmayer, Eugene Fama, Ken French, Laurie Krigman, Bing Liang, John Nofsinger, Srinivasan Rangan, Mark Rubinstein, René Stulz (the editor), Avanidhar Subrahmanyam, Kent Womack, Jason Zweig, two anonymous reviewers, seminar participants at the American Finance Association Meetings (New York, 1999), the 9th Annual Conference on Financial Economics and Accountancy at New York University, Notre Dame University, the University of Illinois, and participants in the Compuserve Investor Forum. All errors are our own.
In this paper, we attempt to shed light on the investment performance of common stocks held directly by households. To do so, we analyze a unique data set that consists of position statements and trading activity for 78,000 households at a large discount brokerage firm over a six-year period ending in January 1997.

Our analyses also allow us to test two competing theories of trading activity. Using a rational expectation framework, Grossman and Stiglitz (1980) argue that investors will trade when the marginal benefit of doing so is equal to or exceeds the marginal cost of the trade. In contrast Odean (1998b), Gervais and Odean (1998), and Caballé and Sákovics (1998) develop theoretical models of financial markets where investors suffer from overconfidence. These overconfidence models predict that investors will trade to their detriment.\(^1\)

Our most dramatic empirical evidence supports the view that overconfidence leads to excessive trading (see Figure 1). On one hand, there is very little difference in the gross performance of households that trade frequently (with monthly turnover in excess of 8.8 percent) and those that trade infrequently. In contrast, households that trade frequently earn a net annualized geometric mean return of 11.4 percent, and those that trade infrequently earn 18.5 percent. These results are consistent with models where trading emanates from investor overconfidence, but are inconsistent with models where trading results from rational expectations. Though liquidity, risk-based rebalancing, and taxes can explain some trading activity, we argue that it belies common sense that these motivations for trade, even in combination, can explain average annual turnover of more than 250 percent for those households that trade most.

We also document that, overall, the households we analyze significantly underperform relevant benchmarks, after a reasonable accounting for transaction costs. These households earn gross returns (before accounting for transaction costs) that are close to those earned by an investment in a value-weighted index of NYSE/AMEX/Nasdaq stocks. During our sample period, an investment in a value-weighted market index earns an annualized geometric mean return of 17.9 percent, the average household earns a gross return of 18.7 percent, and in aggregate households earn a gross return of 18.2 percent. In contrast, the net performance (after accounting for the bid-ask spread and commissions) of these households is below par, with the average household earning 16.4 percent and in aggregate households earning 16.7 percent. The empirical tests supporting these conclusions come from abnormal return calculations that allow each household to self-select its own

---

\(^1\) In an exception to this finding, Kyle and Wang (1997) argue that when traders compete for duopoly profits, overconfident traders may reap greater profits. This prediction is based on several assumptions that do not apply to individuals trading common stocks. Benos (1998) has a similar result. Daniel, Hirshleifer, and Subrahmanyam (1998) consider the asset price implications of overconfidence but do not directly address investor welfare.
investment style and from time-series regressions that employ either the Capital Asset Pricing Model (CAPM) or the three-factor model developed by Fama and French (1993) as our benchmark.

Our descriptive analysis provides several additional conclusions that are noteworthy:

1. Households\(^2\) trade common stocks frequently. The average household turns over more than 75 percent of its common stock portfolio annually.
2. Trading costs are high. The average round-trip trade in excess of $1,000 costs three percent in commissions and one percent in bid-ask spread.
3. Households tilt their investments toward small, high-beta stocks. There is a less obvious tilt toward value (high book-to-market) stocks.

\(^2\) Throughout this paper, “households” and “individual investors” refer to households and investors with discount brokerage accounts. Though we believe that our findings generalize to customers at other discount brokerages, we suspect that the trading practices of retail customers differ. Some of our sample households may have both retail and discount accounts. In these cases, our observations are limited to their discount accounts.
It is the cost of trading and the frequency of trading, not portfolio selections, that explain the poor investment performance of households during our sample period. In fact, the tilt of households toward small stocks and, to a lesser extent, value stocks helps their performance during our sample period (during which small stocks outperform large stocks by 15 basis points per month and value outperforms growth by 20 basis points per month).\(^3\)

The remainder of this paper is organized as follows. We discuss related research in Section I and our data and empirical methods in Section II. Our main descriptive results are presented in Section III. We test the models of investor overconfidence in Section IV. We discuss the impact of price momentum on individual investor performance in Section V and liquidity, risk, and taxes as motivations for trading in Section VI. Concluding remarks are made in Section VII.

I. Related Research

To our knowledge, the current investigation is the first comprehensive study of the aggregate common stock performance of individual investors who manage their own equity investments without the advice of a full-service broker. Schlarbaum, Lewellen, and Lease (1978a) analyze the aggregate common stock performance of investors at a full-service brokerage firm. Odean (1999) and Schlarbaum, Lewellen, and Lease (1978b) analyze the profitability of common stock trades (as distinct from positions held) by individual investors.

Schlarbaum et al. (1978a) calculate monthly gross and net portfolio returns for 2,500 accounts at a retail brokerage firm over a seven-year period ending in December 1970. In a separate paper, Schlarbaum et al. (1978b) analyze the gross and net returns of round-trip trades made by the same 2,500 accounts over the same period. Though they emphasize that their results are conjectural, they conclude that their results “portray an overall picture of quite respectable individual investor security selection acumen.” In contrast, we document that individual investors at a discount brokerage firm during the six-year period ending January 1997 perform poorly.

There are at least three reasons why our results might differ from those in Schlarbaum et al. (1978a, 1978b). First, we analyze households that hold their investments at a discount brokerage firm rather than at a retail brokerage firm. A wide variety of investment advice is available to both retail and discount investors from sources such as newsletters, Value Line, and the financial press. Retail brokerage firms also provide stock selection advice to their clients. If this advice is valuable and if investors attend to it, it is

---

\(^3\) These figures are based on the mean return from February 1991 through January 1997 for the size and book-to-market factors constructed by Fama and French (1993). In the remainder of this paper, when we refer to a size or value premium, our inference is based on the returns of these zero-investment portfolios.
plausible that individual investors at these firms earn both better gross returns and net returns. We would welcome the opportunity to test this hypothesis directly by obtaining a data set similar to that employed in our study from a retail brokerage firm. Barber et al. (1998) and Womack (1996) present evidence that the recommendations of brokerage-house analysts have investment value.

Second, the analysis in Schlarbaum et al. (1978b) focuses on the returns from round-trip trades. There is now evidence that investors have a tendency to sell winning investments and hold on to losing investments (Odean (1998a)). Thus, by analyzing trades rather than position statements (as we do in the current study), Schlarbaum et al. may upwardly bias their return estimates. Schlarbaum et al. (1978a) do attempt to reconstruct monthly positions from trading records and partial end-of-period positions. However, as they point out, stocks purchased before 1964 and sold after 1970 may not appear in their study.

Third, although Schlarbaum et al. (1978a, 1978b) evaluate performance using a variety of market indexes, they do not consider the tendency for individual investors to tilt toward small stocks (though of course firm size did not have the same celebrity status in 1978 that it enjoys today). They do not explicitly address whether such a tilt exists among the individual investors they analyze, but we suspect that it does. This small-stock tilt is likely to be extremely important because small stocks outperform large stocks by 67 basis points per month during their sample period.

As do Schlarbaum et al. (1978b), Odean (1999) focuses on the trades of individual investors. He analyzes the timing of trades made by individual investors at a large discount brokerage firm during the seven years ending in December 1993, a sample period that overlaps with ours. (The data sets employed in Odean (1999) and this study are different.) He documents that the stocks individuals sell subsequently outperform the stocks they buy. Thus, the implications of his study and the current investigation are similar: Individual investors trade too much. However, Odean does not analyze the aggregate performance of all stocks held by individuals. Consequently, he is unable to conclude whether individual investors perform well in aggregate, which is the focus of our investigation.

II. Data and Methods

A. Household Account Data

The primary data set for this research is information from a large discount brokerage firm on the investments of 78,000 households from January 1991 through December 1996. Of the sampled households, 42 percent are in

---

4 The month-end position statements for this period allow us to calculate returns for February 1991 through January 1997. Data on trades are from January 1991 through November 1996.
the western part of the United States, 19 percent in the East, 24 percent in
the South, and 15 percent in the Midwest. The data set includes all accounts
opened by each household at this discount brokerage firm. The sample
selection was performed at the household level and was stratified based on
whether the discount brokerage firm labeled the household as a general
(60,000 households), affluent (12,000 households), or active trader house-
hold (6,000 households). The firm labels households that make more than 48
trades in any year as active traders, households with more than $100,000 in
equity at any point in time as affluent, and all other households as general.
If a household qualifies as either active trader or affluent, it is assigned the
active trader label. In 1997, approximately 61 percent of all retail accounts
at this brokerage firm were classified as general, 28 percent as affluent, and
11 percent as active. Sampled households were required to have an open
account with the discount brokerage firm during 1991. Roughly half of the
accounts in our analysis were opened prior to 1987 and half were opened

In this research, we focus on the common stock investments of house-
holds. We exclude from the current analysis investments in mutual funds
(both open-end and closed-end), American Depositary Receipts (ADRs), war-
rants, and options. Of the 78,000 sampled households, 66,465 have posi-
tions in common stocks during at least one month; the remaining accounts
hold either cash or investments in other than individual common stocks.
Households have, on average, two accounts: 48 percent have a single ac-
count, 27 percent have two, 14 percent have three, and the remaining
11 percent have more than three. The most common reason for two ac-
counts is the tax-preferred status of retirement accounts (e.g., IRAs and
Keoghs). Some households also have different accounts for different house-
hold members (e.g., custodial accounts for children). Roughly 60 percent of
the market value in the accounts is held in common stocks. In these house-
holds, more than 3 million trades are made in all securities during the
sample period, with common stocks accounting for slightly more than 60
percent of all trades. On average during our sample period, the mean house-
hold holds 4.3 stocks worth $47,334, though each of these figures is posi-
tively skewed. The median household holds 2.61 stocks worth $16,210. In
December 1996, these households held more than $4.5 billion in common
stock.

In Table I, we present descriptive information on the trading activity for
our sample. Panels A and B show there are slightly more purchases (1,082,107)
than sales (887,594) during our sample period, though the average value of
stocks sold ($13,707) is slightly higher than the value of stocks purchased
($11,205). As a result, the aggregate value of purchases and sales is roughly
equal ($12.1 and $12.2 billion, respectively). The average trade is transacted
at a price of $31 per share. The value of trades and the transaction price of
trades are positively skewed; the medians for both purchases and sales are
substantially less than the mean values.
Table I
Descriptive Statistics on Trade Size, Trade Price, Transaction Costs, and Turnover

The sample is account records for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. Spread is calculated as the transaction price divided by the closing price on the day of the transaction minus one (and then multiplied by minus one for purchases). Commission is calculated as the commission paid divided by the value of the trade. Monthly turnover is the beginning-of-month market value of shares purchased in month $t-1$ (or sold in month $t$) divided by the total beginning-of-month market value of shares held in month $t$. Trade-weighted spread and commission are averages weighted by trade size. Aggregate turnover is the aggregate value of sales (or purchases) divided by the aggregate value of positions held during our sample period.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>25th Percentile</th>
<th>Median</th>
<th>75th Percentile</th>
<th>Standard Deviation</th>
<th>No. of Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Purchases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade size ($)</td>
<td>11,205</td>
<td>2,513</td>
<td>4,988</td>
<td>10,500</td>
<td>32,179</td>
<td>1,082,107</td>
</tr>
<tr>
<td>Price/share</td>
<td>31.06</td>
<td>11.00</td>
<td>23.00</td>
<td>40.00</td>
<td>117.82</td>
<td>1,082,107</td>
</tr>
<tr>
<td>Monthly turnover (%)</td>
<td>6.49</td>
<td>0.54</td>
<td>2.67</td>
<td>7.08</td>
<td>11.89</td>
<td>66,465</td>
</tr>
<tr>
<td>Commission (%)</td>
<td>1.58</td>
<td>0.78</td>
<td>1.29</td>
<td>2.10</td>
<td>1.45</td>
<td>966,492</td>
</tr>
<tr>
<td>Spread (%)</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Sales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade size ($)</td>
<td>13,707</td>
<td>2,688</td>
<td>5,738</td>
<td>13,000</td>
<td>38,275</td>
<td>887,594</td>
</tr>
<tr>
<td>Price/share</td>
<td>31.22</td>
<td>12.00</td>
<td>24.00</td>
<td>41.00</td>
<td>113.03</td>
<td>887,594</td>
</tr>
<tr>
<td>Monthly turnover (%)</td>
<td>6.23</td>
<td>0.39</td>
<td>2.58</td>
<td>6.95</td>
<td>11.36</td>
<td>66,465</td>
</tr>
<tr>
<td>Commission (%)</td>
<td>1.45</td>
<td>0.70</td>
<td>1.16</td>
<td>1.91</td>
<td>1.06</td>
<td>785,206</td>
</tr>
<tr>
<td>Spread (%)</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>845,644</td>
</tr>
<tr>
<td>Panel C: Trade-Weighted and Aggregate Purchases</td>
<td>6.05</td>
<td>Not Applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade-weighted commission (%)</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade-weighted spread (%)</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel D: Trade-Weighted Sales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate monthly turnover (%)</td>
<td>6.06</td>
<td>Not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade-weighted commission (%)</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade-weighted spread (%)</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Commissions are calculated based on trades in excess of $1,000. Including smaller trades results in a mean buy (sale) commission of 2.09 (3.07) percent.
For each trade, we estimate the bid-ask spread component of transaction costs for purchases ($spr_{db}$) and sales ($spr_{ds}$) as

\begin{equation}
\text{spr}_{ds} = \left( \frac{P^c_{ds}}{P^s_{ds}} - 1 \right) \quad \text{and} \quad \text{spr}_{db} = -\left( \frac{P^c_{db}}{P^b_{db}} - 1 \right),
\end{equation}

where $P^c_{ds}$ and $P^c_{db}$ are the reported closing prices from the Center for Research in Security Prices (CRSP) daily stock return files on the day of a sale and purchase, respectively, and $P^s_{ds}$ and $P^b_{db}$ are the actual sale price and purchase price from our account database.\(^5\) Our estimate of the bid-ask spread component of transaction costs includes any market impact that might result from a trade. It also includes an intraday return on the day of the trade. (In Appendix A, we provide a detailed reconciliation of our return calculations.) The commission component of transaction costs is estimated as the dollar value of the commission paid scaled by the total principal value of the transaction, both of which are reported in our account data.

The average purchase costs an investor 0.31 percent, and the average sale costs an investor 0.69 percent in bid-ask spread. Our estimate of the bid-ask spread is very close to the trading cost of 0.21 percent for purchases and 0.63 percent for sales paid by open-end mutual funds from 1966 to 1993 (Carhart (1997)).\(^6\) The average purchase in excess of $1,000 costs 1.58 percent in commissions, and the average sale in excess of $1,000 costs 1.45 percent.\(^7\)

In Panels C and D of Table I, we calculate the trade-weighted (weighted by trade size) spreads and commissions. These figures can be thought of as the total cost of conducting the $24 billion in common stock trades ($12 billion each in purchases and sales). Trade size has little effect on spread costs (0.27 percent for purchases and 0.69 percent for sales) but substantially reduces the commission costs (0.77 percent for purchases and 0.66 percent for sales).

In sum, the average trade incurs a round-trip transaction cost of about one percent for the bid-ask spread and about three percent in commissions. In aggregate, round-trip trades cost about one percent for the bid-ask spread and about 1.4 percent in commissions.

---

\(^5\) Kraus and Stoll (1972), Holthausen, Leftwich, and Mayers (1987), Laplante and Muscarella (1997), and Beebower and Priest (1980) use closing prices either before or following a transaction to estimate effective spreads and market impact. See Keim and Madhavan (1998) for a review of different approaches to calculating transactions costs.

\(^6\) Odean (1999) finds that individual investors are more likely to both buy and sell particular stocks when the prices of those stocks are rising. This tendency can partially explain the asymmetry in buy and sell spreads. Any intraday price rises following transactions subtract from our estimate of the spread for buys and add to our estimate of the spread for sells.

\(^7\) To provide more representative descriptive statistics on percentage commissions, we exclude trades of less than $1,000. The inclusion of these trades results in a round-trip commission cost of five percent on average (2.1 percent for purchases and 3.1 percent for sales).
Finally, we calculate the monthly portfolio turnover for each household. In each month during our sample period, we identify the common stocks held by each household at the beginning of month $t$ from their position statements. To calculate monthly sales turnover, we match these positions to sales during month $t$. The monthly sales turnover is calculated as the shares sold times the beginning-of-month price per share divided by the total beginning-of-month market value of the household’s portfolio. To calculate monthly purchase turnover, we match these positions to purchases during month $t - 1$. The monthly purchase turnover is calculated as the shares purchased times the beginning-of-month price per share divided by the total beginning-of-month market value of the portfolio. In Panels A and B of Table I we report that, on average, households purchase 6.49 percent and sell 6.23 percent of their stock portfolio each month, though the median household trades much less frequently (buying 2.67 percent of their stock portfolio and selling 2.58 percent). In Panels C and D, we calculate aggregate purchase (sales) turnover by summing all purchases (sales) and dividing by the sum of all positions during our sample period. The aggregate purchase turnover is 6.05 percent and the aggregate sales turnover is 6.06 percent.

In sum, these investors trade their common stocks quite frequently. The average household turns over more than 75 percent of its common stock portfolio each year. This result is uncannily close to the average turnover of 77 percent reported by U.S. common stock mutual funds for the period 1966 to 1993 (Carhart (1997)). In aggregate, these investors turn over more than 70 percent of their invested wealth each year.

B. Measuring Return Performance

The focus of our analysis is the return performance of investments in common stocks by households. We analyze both the gross performance and net performance (after a reasonable accounting for commissions, the bid-ask spread, and the market impact of trades).

We estimate the gross monthly return on each common stock investment using the beginning-of-month position statements from our household data and the CRSP monthly returns file. In so doing, we make two simplifying assumptions. First, we assume that all securities are bought or sold on the last day of the month. Thus, we ignore the returns earned on stocks purchased from the purchase date to the end of the month and include the returns earned on stocks sold from the sale date to the end of the month.

---

8 If more shares are sold than were held at the beginning of the month (e.g., because an investor purchases additional shares after the beginning of the month), we assume the entire beginning-of-month position in that security is sold. Similarly, if more shares were purchased in the preceding month than are held in the position statement, we assume the entire position is purchased in the preceding month. Thus, turnover, as we have calculated it, cannot exceed 100 percent in a month.
Second, we ignore intramonth trading (e.g., a purchase on March 6 and a sale of the same security on March 20), though we do include in our analysis short-term trades that yield a position at the end of a calendar month.

In Appendix A, we document that accounting for the exact timing of trades would reduce the performance of individual investors by about two basis points per month. In Appendix B, we document that accounting for intramonth trades would improve the performance of individual investors reported in our main results by less than one basis point per month. More important, a careful accounting for both the exact timing of trades and the profitability of intramonth trades indicates that the results we report in the main text are slightly high for our full sample and for every sample partition that we analyze.

Consider the common stock portfolio for a particular household. The gross monthly return on the household’s portfolio \( R_{ht}^{gr} \) is calculated as

\[
R_{ht}^{gr} = \sum_{i=1}^{s_{ht}} p_{it} R_{it}^{gr},
\]

where \( p_{it} \) is the beginning-of-month market value for the holding of stock \( i \) by household \( h \) in month \( t \) divided by the beginning-of-month market value of all stocks held by household \( h \), \( R_{it}^{gr} \) is the gross monthly return for stock \( i \), and \( s_{ht} \) is the number of stocks held by household \( h \) in month \( t \).

For security \( i \) in month \( t \), we calculate a monthly return net of transaction costs \( R_{it}^{net} \) as

\[
(1 + R_{it}^{net}) = (1 + R_{it}^{gr}) \frac{(1 - c_{it}^s)}{(1 + c_{it,t-1}^b)},
\]

where \( c_{it}^s \) is the cost of sales scaled by the sales price in month \( t \) and \( c_{it,t-1}^b \) is the cost of purchases scaled by the purchase price in month \( t - 1 \). The costs of purchases and sales include the commissions and bid-ask spread components, which are estimated individually for each trade as previously described. Thus, for a security purchased in month \( t - 1 \) and sold in month \( t \), both \( c_{it}^s \) and \( c_{it,t-1}^b \) are positive; for a security neither purchased in month \( t - 1 \) nor sold in month \( t \), both \( c_{it}^s \) and \( c_{it,t-1}^b \) are zero. Because the timing and cost of purchases and sales vary across households, the net return for security \( i \) in month \( t \) varies across households. The net monthly portfolio return for each household is

\[
R_{ht}^{net} = \sum_{i=1}^{s_{ht}} p_{it} R_{it}^{net}.
\]
If only a portion of the beginning-of-month position in stock $i$ is purchased or sold, the transaction cost is applied only to that portion. We estimate the aggregate gross and net monthly return earned by individual investors as

$$RAG_{t}^{gr} = \sum_{h=1}^{n_{ht}} x_{ht} R_{ht}^{gr} \quad \text{and} \quad RAG_{t}^{net} = \sum_{h=1}^{n_{ht}} x_{ht} R_{ht}^{net},$$

where $n_{ht}$ is the number of households with common stock investment in month $t$ and $x_{ht}$ is the beginning-of-month market value of common stocks held by household $h$ divided by the beginning-of-month market value of common stock held by all households. We estimate the gross and net monthly return earned by the average household as

$$RH_{t}^{gr} = \frac{1}{n_{ht}} \sum_{h=1}^{n_{ht}} R_{ht}^{gr} \quad \text{and} \quad RH_{t}^{net} = \frac{1}{n_{ht}} \sum_{h=1}^{n_{ht}} R_{ht}^{net}.$$

### C. Risk-Adjusted Return Performance

We calculate four measures of risk-adjusted performance. First, we calculate an own-benchmark abnormal return for individual investors, which is similar in spirit to that proposed by Grinblatt and Titman (1993) and Lakonishok, Shleifer, and Vishny (1992). In this abnormal return calculation, the benchmark for household $h$ is the month $t$ return of the beginning-of-year portfolio held by household $h$. It represents the return that the household would have earned had it merely held its beginning-of-year portfolio for the entire year. The own-benchmark abnormal return is the return earned by household $h$ less the own-benchmark return; if the household did not trade during the year, the own-benchmark return is zero for all 12 months during the year. In each month, the abnormal returns across households are averaged, yielding a 72-month time-series of mean monthly own-benchmark abnormal returns. Statistical significance is calculated using $t$-statistics based on this time-series. The advantage of the own-benchmark abnormal return

9 A fifth alternative measure of risk-adjusted returns is the Sharpe ratio, the mean excess return divided by its standard deviation. The average Sharpe ratio for the gross (net) return of the average household in our sample is 0.179 (0.134). The Sharpe ratio for the market during our sample period is 0.366 = (1.0578/2.8880). We do not report Sharpe ratios for most partitions of the data because we do not observe the entire portfolios of these households. Unobserved assets such as equities at other brokerage firms and mutual fund holdings are unlikely to greatly change average observed portfolio returns, but they are likely to reduce average observed volatility. Thus we tend to underestimate the total portfolio Sharpe ratios of investors with significant unobserved assets.

10 When calculating this benchmark, we begin the year on February 1. We do so because our first monthly position statements are from the month end of January 1991. If the stocks held by a household at the beginning of the year are missing CRSP returns data during the year, we assume that stock is invested in the remainder of the household’s portfolio.
measure is that it does not adjust returns according to a particular risk model. No model of risk is universally accepted; furthermore, it may be inappropriate to adjust investors’ returns for stock characteristics that they do not associate with risk. The own-benchmark measure allows each household to self-select the investment style and risk profile of its benchmark (i.e., the portfolio it held at the beginning of the year), thus emphasizing the effect trading has on performance.

Second, we calculate the mean monthly market-adjusted abnormal return for individual investors by subtracting the return on a value-weighted index of NYSE/AMEX/Nasdaq stocks from the return earned by individual investors.

Third, we employ the theoretical framework of the capital asset pricing model and estimate Jensen’s alpha by regressing the monthly excess return earned by individual investors on the market excess return. For example, to evaluate the gross monthly return earned by individual investors in aggregate, we estimate the following monthly time-series regression:

$$ (RAG^i_t - R_{ft}) = \alpha_i + \beta_i (R_{mt} - R_{ft}) + \epsilon_{it}, $$

where $R_{ft}$ is the monthly return on T-bills,$^{11}$ $R_{mt}$ is the monthly return on a value-weighted market index, $\alpha_i$ is the CAPM intercept (Jensen’s alpha), $\beta_i$ is the market beta, and $\epsilon_{it}$ is the regression error term. The subscript $i$ denotes parameter estimates and error terms from regression $i$, where we estimate four regressions: one each for the gross and net performance of individual investors in aggregate, and one each for the gross and net performance of the average household.

Fourth, we employ an intercept test using the three-factor model developed by Fama and French ($^{1993}$). For example, to evaluate the performance of individuals in aggregate, we estimate the following monthly time-series regression:

$$ (RAG^i_t - R_{ft}) = \alpha_j + \beta_j (R_{mt} - R_{ft}) + s_j SMB_t + h_j HML_t + \epsilon_{jt}, $$

where $SMB_t$ is the return on a value-weighted portfolio of small stocks minus the return on a value-weighted portfolio of large stocks and $HML_t$ is the return on a value-weighted portfolio of high book-to-market stocks minus the return on a value-weighted portfolio of low book-to-market stocks.$^{12}$ The regression yields parameter estimates of $\alpha_j$, $\beta_j$, $s_j$, and $h_j$. The error term in the regression is denoted by $\epsilon_{jt}$. The subscript $j$ denotes parameter estimates and error terms from regression $j$, where we again estimate four regres-

$^{11}$ The return on Treasury bills is from Stocks, Bonds, Bills, and Inflation, 1997 Yearbook, Ibbotson Associates, Chicago, Ill.

$^{12}$ The construction of these portfolios is discussed in detail in Fama and French ($^{1993}$). We thank Kenneth French for providing us with these data.
sions. We place particular emphasis on the Fama–French intercept tests, since individual investors tilt their portfolios toward small stocks. The three-factor model provides a reasonable adjustment for this small stock tilt.\footnote{Lyon, Barber, and Tsai (1999) document that intercept tests using the three-factor model are well specified in random samples and samples of large or small firms. Thus, the Fama–French intercept tests employed here account well for the small stock tilt of individual investors.}

Fama and French (1993) argue that the risk of common stock investments can be parsimoniously summarized as risk related to the market, firm size, and a firm’s book-to-market ratio. We measure these three risk exposures using the coefficient estimates on the market excess return ($R_{mt} - R_{ft}$), the size zero investment portfolio ($SMB_t$), and the book-to-market zero-investment portfolio ($HML_t$) from the three-factor regressions. Portfolios with above-average market risk have betas greater than one, $\beta_j > 1$. Portfolios with a tilt toward small (value) stocks relative to a value-weighted market index have size (book-to-market) coefficients greater than zero, $s_j > 0$ ($h_j > 0$).

We suspect there is little quibble with interpreting the coefficient on the market excess return ($\beta_j$) as a risk factor. Interpreting the coefficient estimates on the size and the book-to-market zero-investment portfolios is more controversial. For the purposes of this investigation, we are interested in measuring risk as perceived by individual investors. As such, it is our casual observation that investors view common stock investment in small firms as riskier than that in large firms. Thus, we would willingly accept a stronger tilt toward small stocks as evidence that a particular group of investors is pursuing a strategy that it perceives as riskier. It is less clear to us whether a tilt toward high book-to-market stocks (which tend to be ugly, financially distressed, firms) or toward low book-to-market stocks (which tend to be high-growth firms) is perceived as riskier by investors. As such, we interpret the coefficient estimates on the book-to-market zero-investment portfolio with a bit more trepidation.\footnote{Some authors have also identified price momentum effects in stock returns. We discuss momentum in Section V.}

### III. Results

#### A. Full Sample Results

Our main findings for the full sample can be summarized simply. The gross return earned by individual investors in aggregate ($R_{AG}^{gr}$) and the gross return earned by the average household ($R_{HG}^{gr}$) are remarkably close to that earned by an investment in a value-weighted index of NYSE/AMEX/Nasdaq stocks.\footnote{We use the NYSE/AMEX/Nasdaq value-weighted market index constructed by Fama and French (1993). Firms comprising the index must have data for firm size and book-to-market ratio. The correlation between this market index and the NYSE/AMEX/Nasdaq value-weighted index from CRSP is 99.9 percent.} The annualized geometric mean return earned by individ-

\begin{align*}
\text{\footnotesize Footnotes}:
13 \quad & \text{Lyon, Barber, and Tsai (1999) document that intercept tests using the three-factor model are well specified in random samples and samples of large or small firms. Thus, the Fama–French intercept tests employed here account well for the small stock tilt of individual investors.} \\
14 \quad & \text{Some authors have also identified price momentum effects in stock returns. We discuss momentum in Section V.} \\
15 \quad & \text{We use the NYSE/AMEX/Nasdaq value-weighted market index constructed by Fama and French (1993). Firms comprising the index must have data for firm size and book-to-market ratio. The correlation between this market index and the NYSE/AMEX/Nasdaq value-weighted index from CRSP is 99.9 percent.}
\end{align*}
ual investors in aggregate, the average household, and the value-weighted market index are 18.2, 18.7, and 17.9 percent, respectively. In contrast, the net returns earned by individual investors in aggregate \( (R_{AGt}^{\text{net}}) \) and the net return earned by the average household \( (R_{IHt}^{\text{net}}) \) underperform the value-weighted index by more than 100 basis points annually. The net annualized geometric mean return earned by individual investors in aggregate and by the average household are 16.7 and 16.4 percent, respectively.

The results of this analysis are presented in Table II. Panel A presents results for the gross performance of individual investors in aggregate, Panel B presents results for the average household. Three of the four performance measures indicate that the gross performance of individual investors is unremarkable; neither the market-adjusted return, Jensen’s alpha, nor the intercept test from the Fama–French three-factor model is reliably different from zero. The fourth performance measure, the own-benchmark abnormal return, is reliably negative. This result indicates that the investors would have earned higher returns from following a buy-and-hold strategy; they hurt their gross performance by trading.

Also noteworthy in these results are the coefficient estimates on the market, size, and book-to-market factors. Individual investors tilt toward small stocks with high market risk. The market beta for stocks held by individual investors is reliably greater than one and the coefficient estimate on \( SMB_t \) is reliably positive. Though in aggregate individual investors have no tilt toward value or growth, the average household has a slight tilt toward value stocks (those with high book-to-market ratios) and a more pronounced tilt toward small stocks.\(^\text{16}\) These tilts serve individual investors well during our period of analysis; the mean monthly returns on \( SMB_t \) and \( HML_t \) during our 72-month sample period are 0.15 and 0.20 percent, respectively. This observation can account for the fact that the market-adjusted return performance of individual investors is positive (albeit unreliably so), while Jensen’s alpha (CAPM intercept) and the intercept test from the Fama–French three-factor model are negative.

The style preferences of individual investors complement those of institutions. Institutional investors have a clear preference for large stocks. Gompers and Metrick (1998) document this preference for large institutions; Carhart (1997) and Falkenstein (1996) document a similar bias for mutual funds. As is the case for individual investors, the growth or value preference of institutions is less obvious. Gompers and Metrick (1998) document that large institutions prefer value stocks, but Carhart (1997, Table III) documents that mutual fund holdings tilt toward growth stocks.\(^\text{17}\)

\(^\text{16}\) Aggregate measures weight each household by the value of that household's common stocks. Average household measures weight each household equally.

\(^\text{17}\) Kang and Stulz (1997) document that foreign investors in Japanese equity markets prefer large growth stocks. It is likely that these foreign investors are predominantly institutions.
Table II
Summary of the Percentage Monthly Abnormal Return Measures for the Average Household and Aggregate Household

Returns are based on month-end position statements for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. Panel A (Panel C) presents results for the gross (net) return on a portfolio that mimics the aggregate investment of all households. Panel B (Panel D) presents results for the gross (net) return on a portfolio that mimics the investment of the average household. Own-benchmark abnormal return is the return on the household portfolio minus the return on the portfolio the household held at the end of the previous January. Market-adjusted return is the return on the household portfolio less the return on a value-weighted NYSE/AMEX/Nasdaq index. CAPM is the results from a time-series regression of the household excess return on the market excess return ($R_{mt} - R_{ft}$). Fama–French three-factor is the results from time-series regression of household excess return on the market excess return, a zero-investment book-to-market portfolio ($HML_t$), and a zero-investment size portfolio ($SMB_t$). *p*-values are presented in parentheses.

<table>
<thead>
<tr>
<th>Panel A: Gross Percentage Monthly Returns in Aggregate</th>
<th>Coefficient Estimate on:</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own-benchmark abnormal return</td>
<td>-0.049**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.723)</td>
<td></td>
</tr>
<tr>
<td>CAPM</td>
<td>-0.067</td>
<td>1.100***</td>
</tr>
<tr>
<td></td>
<td>(0.543)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Fama–French three-factor</td>
<td>-0.076</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>(0.357)</td>
<td>(0.324)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Gross Percentage Monthly Returns for the Average Household</th>
<th>Coefficient Estimate on:</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own-benchmark abnormal return</td>
<td>-0.048**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.672)</td>
<td></td>
</tr>
<tr>
<td>CAPM</td>
<td>-0.014</td>
<td>1.087</td>
</tr>
<tr>
<td></td>
<td>(0.944)</td>
<td>(0.177)</td>
</tr>
<tr>
<td>Fama–French three-factor</td>
<td>-0.154</td>
<td>1.120***</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.005)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Net Percentage Monthly Returns in Aggregate</th>
<th>Coefficient Estimate on:</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own-benchmark abnormal return</td>
<td>-0.155***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>-0.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.496)</td>
<td></td>
</tr>
<tr>
<td>CAPM</td>
<td>-0.175</td>
<td>1.096***</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Fama–French three-factor</td>
<td>-0.180**</td>
<td>-0.040</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.251)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: Net Percentage Monthly Returns for the Average Household</th>
<th>Coefficient Estimate on:</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own-benchmark abnormal return</td>
<td>-0.194***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>-0.080</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.621)</td>
<td></td>
</tr>
<tr>
<td>CAPM</td>
<td>-0.177</td>
<td>1.082</td>
</tr>
<tr>
<td></td>
<td>(0.360)</td>
<td>(0.194)</td>
</tr>
<tr>
<td>Fama–French three-factor</td>
<td>-0.311**</td>
<td>0.131**</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.008)</td>
</tr>
</tbody>
</table>

***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively (two-tailed). The null hypothesis for beta (the coefficient estimate on the market excess return) is $H_0: \beta = 1$. 

Trading Is Hazardous to Your Wealth 787
Table III
Descriptive Statistics, Gross Returns, and Net Returns for Household Quintiles formed on Beginning Position Value

The sample is account records for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. Households are sorted into quintiles based on the market value of common stocks in the first month that a household appears during our sample period. Quintile 1 contains households with the smallest market value of common stock holdings, quintile 5 contains households with the largest value. Beginning position value is the market value of common stocks held in the first month that the household appears during our sample period. Mean monthly turnover is the average of sales and purchase turnover. Coefficient estimates are those from a time-series regression of the gross average household excess return on the market excess return \( (R_{mt} - R_p) \), a zero-investment book-to-market portfolio \( (HML_t) \), and a zero-investment size portfolio \( (SMB_t) \). Raw return is the average monthly return for the average household. Own-benchmark abnormal return is the return on the household portfolio less the return on a value-weighted NYSE/AMEX/Nasdaq index. CAPM intercept is the estimated intercept from a time-series regression of the household excess return on the market excess return \( (R_{mt} - R_p) \). Fama–French intercept is the estimated intercept from time-series regressions of household excess return on the market excess return, a zero-investment book-to-market portfolio \( (HML_t) \), and a zero-investment size portfolio \( (SMB_t) \). p-values are presented in parentheses.

### Panel A: Descriptive Statistics

<table>
<thead>
<tr>
<th>Quintile</th>
<th>1 (Small)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 (Large)</th>
<th>Difference: Lrg – Sml</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean beginning position value</td>
<td>1,581</td>
<td>4,653</td>
<td>8,599</td>
<td>16,725</td>
<td>149,710</td>
<td>N.A.</td>
</tr>
<tr>
<td>Mean monthly turnover (%)</td>
<td>6.68</td>
<td>6.35</td>
<td>6.31</td>
<td>6.13</td>
<td>6.33</td>
<td>−0.35*** (0.000)</td>
</tr>
</tbody>
</table>

Coefficient estimate on: \( (R_{mt} - R_p) \)

<table>
<thead>
<tr>
<th>Quintile</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Difference: Lrg – Sml</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean monthly turnover (%)</td>
<td>6.68</td>
<td>6.35</td>
<td>6.31</td>
<td>6.13</td>
<td>6.33</td>
<td>−0.35*** (0.000)</td>
</tr>
</tbody>
</table>

### Panel B: Gross Average Household Percentage Monthly Return

| Raw return | 1.722 | 1.511 | 1.473 | 1.424 | 1.400 | −0.322 |
| Own-benchmark | −0.071 | −0.051** | −0.038a | −0.038a | −0.037a | 0.034 |
| (0.101) | (0.022) | (0.070) | (0.061) | (0.077) | (0.487) |
| Market-adjusted return | 0.302 | 0.091 | 0.053 | 0.004 | −0.020 | −0.322 |
| (0.370) | (0.648) | (0.755) | (0.980) | (0.857) | (0.185) |
| CAPM intercept | 0.182 | −0.015 | −0.043 | −0.089 | −0.072 | −0.253 |
| (0.612) | (0.942) | (0.811) | (0.570) | (0.541) | (0.328) |
| Fama–French intercept | −0.137 | −0.152 | −0.149 | −0.186a | −0.140 | 0.003 |
| (0.510) | (0.227) | (0.206) | (0.082) | (0.101) | (0.983) |

### Panel C: Net Average Household Percentage Monthly Return

| Raw return | 1.478 | 1.328 | 1.313 | 1.280 | 1.279 | −0.199 |
| Own-benchmark | −0.270*** | −0.206*** | −0.178*** | −0.169*** | −0.150*** | 0.120*** |
| (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.023) |
| Market-adjusted return | 0.059 | −0.092 | −0.107 | −0.140 | −0.141 | −0.199 |
| (0.860) | (0.635) | (0.521) | (0.339) | (0.200) | (0.404) |
| CAPM intercept | −0.056 | −0.193 | −0.198 | −0.229 | −0.139 | −0.133 |
| (0.875) | (0.350) | (0.264) | (0.140) | (0.105) | (0.602) |
| Fama–French intercept | −0.366a | −0.323** | −0.298*** | −0.319*** | −0.254*** | 0.112 |
| (0.079) | (0.011) | (0.013) | (0.003) | (0.004) | (0.450) |

***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively (two-tailed). The null hypothesis for beta (the coefficient estimate on the market excess return) is \( H_0: \beta = 1 \) except in the difference column, where the null hypothesis is \( H_0: \beta = 0 \).
The more interesting findings of our analysis are contained in Panels C and D of Table II. Net of transaction costs, individual investors perform poorly. Both the market-adjusted return and the CAPM intercepts are negative, though unreliably so. The own-benchmark abnormal return and the Fama–French intercept provide the most compelling evidence of underperformance. These performance measures indicate significant underperformance of 15 to 31 basis points per month (1.8 percent to 3.7 percent per year, with $t$-statistics ranging from $-2.20$ to $-10.21$). These two performance measures are most appropriate in our setting because they control for the style preference of individual investors: small stocks with above-average market risk. In particular, the own-benchmark abnormal returns indicate individual investors would have increased their annual return by about two percent had they merely held their beginning-of-year portfolio. In combination, these results indicate that the net return performance of individual investors is reliably negative.

One might wonder whether our results are driven by a short sample period coinciding with an unusual stock market. Though the market returned about 18 percent per year during our sample period, the market return was negative in 20 of the 72 months. When we compare the performance of individual investors during the 20 months when the market was down to the 52 months in which the market was up, the performance measures presented in Table II are virtually identical.

**B. Sorting on Portfolio Size**

We test the robustness of our results across different position sizes by partitioning the households into quintiles on the basis of portfolio size. We define portfolio size as the market value of common stocks held in the first month for which there is a position statement.\(^{18}\) Each quintile represents the common stock investments of more than 12,000 households.

Descriptive statistics on the partition by portfolio size are presented in Table III, Panel A. The largest portfolios have a mean beginning position market value of $149,750, the smallest portfolios average $1,581. Small portfolios have slightly higher monthly turnover (6.68 percent) than large portfolios (6.33 percent). As before, we estimate the parameters of the Fama–French three-factor model, where the dependent variable is the monthly mean gross household excess return for each quintile.\(^{19}\) The coefficient estimates on the market, size, and book-to-market factors reveal that small portfolios tilt more heavily toward high-beta, small, value stocks than do large portfolios.

---

\(^{18}\) If the first position statement appears after January 1991, we do not discount the market value of the common stocks to January 1991 in our rankings. Our results are virtually identical if we discount the market value of these common stocks using the return on the value-weighted market index.

\(^{19}\) In the interest of parsimony, here and in the remainder of the paper we do not report results for the aggregate performance of each partition. We note when conclusions are different using the aggregate performance.
The gross and net returns for each quintile are presented in Table III, Panels B and C. Focusing first on the gross performance (Panel B), we find that small portfolios (quintile 1) earn higher average returns than large portfolios (quintile 5), though the difference is not reliably different from zero. This difference is likely attributable to the fact that small portfolios tilt more heavily toward small value stocks, which performed well during our sample period. The net performance results are presented in Panel C. The market-adjusted return and Jensen’s alpha are similar to those reported for the full sample for each quintile. Though the point estimates are consistently negative, they are not reliably so. Of course, these risk-adjustments ignore the fact that investors are tilting toward small value stocks. In contrast, the own-benchmark abnormal returns and the intercept tests from the Fama–French three-factor model indicate significant underperformance, ranging from 15 to 37 basis points per month, in each of the quintiles. In sum, after a reasonable accounting for the size and value tilts of small investors, we document that both small and large portfolios underperform.

C. Cross-Sectional Variation in Performance

We should emphasize that the aggregate performance and average household performance, though germane and interesting, mask considerable cross-sectional variation in the performance across households. For each household, we calculate the mean monthly market-adjusted abnormal return. We present the distribution of these means in Table IV. Consistent with the results presented in Table II, the median household earns a gross monthly market-adjusted return of −0.01 percent and a net return of −0.14 percent. Though 49.3 percent of households outperform a value-weighted market index before transaction costs, only 43.4 percent outperform the index after costs. Nonetheless, many households perform very well: 25 percent of all households beat the market, after accounting for transaction costs, by more than 0.50 percent per month (more than six percent annually). Conversely, many households perform very poorly: 25 percent of all households underperform the market, after accounting for transaction costs, by more than 0.73 percent per month (more than eight percent annually).

IV. Overconfidence and Performance

It is well documented that people tend to be overconfident (e.g., Alpert and Raiffa (1982), Griffin and Tversky (1992); see Odean (1998b) for a more detailed review). Odean (1998b), Gervais and Odean (1998), and Caballé and Sákovics (1998) develop theoretical models of financial markets where in-

\[\text{We omit from this analysis accounts that held common stocks for fewer than 12 months during our 72-month sample period.}\]
Investors suffering from overconfidence trade too much (i.e., trading, at the margin, reduces their expected utility). In contrast, in a rational expectation framework, Grossman and Stiglitz (1980) argue that investors will trade when the marginal benefit of doing so is equal to or exceeds the marginal cost of the trade (including the cost of acquiring information). Odean (1998b) analyzes a variation of Grossman and Stiglitz’s model in which investors are overconfident. The two models yield different predictions about the gains of trading. The rational expectations model predicts that investors who trade more (i.e., those whose expected trading is greater) will have the same expected utility as those who trade less. The overconfidence model predicts that investors who trade more will have lower expected utility.

Consider the implications of these two models in our empirical setting. The overconfidence model predicts that the net return performance of households with high turnover will be lower than that of households with low turnover, while making no prediction about the differences in gross returns. In Grossman–Stiglitz, active and passive investors have equivalent expected utilities. Active traders must earn higher expected gross returns in order to

<table>
<thead>
<tr>
<th>Gross Monthly Market-Adjusted Return (%)</th>
<th>Net Monthly Market-Adjusted Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>−19.46</td>
</tr>
<tr>
<td>1st percentile</td>
<td>−4.32</td>
</tr>
<tr>
<td>5th percentile</td>
<td>−2.12</td>
</tr>
<tr>
<td>10th percentile</td>
<td>−1.34</td>
</tr>
<tr>
<td>25th percentile</td>
<td>−0.57</td>
</tr>
<tr>
<td>Median</td>
<td>−0.01</td>
</tr>
<tr>
<td>75th percentile</td>
<td>0.66</td>
</tr>
<tr>
<td>90th percentile</td>
<td>1.62</td>
</tr>
<tr>
<td>95th percentile</td>
<td>2.41</td>
</tr>
<tr>
<td>99th percentile</td>
<td>4.86</td>
</tr>
<tr>
<td>Maximum</td>
<td>48.53</td>
</tr>
<tr>
<td>Total households</td>
<td>62,439</td>
</tr>
<tr>
<td>Percentage &gt; 0</td>
<td>49.3%***</td>
</tr>
<tr>
<td>Binomial Z-statistic</td>
<td>−3.38</td>
</tr>
</tbody>
</table>

*** indicates significant difference from 50 percent at the 1% level.
offset their greater trading costs.\textsuperscript{21} The Grossman–Stiglitz model therefore predicts that the gross risk-adjusted return performance of households with high turnover will be higher than that of households with low turnover, but there will be little difference in the net risk-adjusted returns.

To test these competing models, we partition our sample of households into quintiles on the basis of mean monthly turnover (defined as the average of purchase and sale turnover). Each quintile represents the common stock investments of more than 12,000 households. Descriptive statistics for each of the quintiles are presented in Table V, Panel A. The households with low turnover average 0.19 percent turnover per month, those with high turnover average 21.49 percent. To qualify as a high turnover portfolio, a household would need to turn over at least 8.7 percent of its portfolio in an average month. Households with low turnover also tend to have larger accounts.

As before, we estimate the parameters of the Fama–French three-factor model, where the dependent variable is the monthly mean gross household excess return for each turnover quintile. The coefficient estimates on the market, size, and book-to-market factors reveal that the high turnover households tilt more heavily toward high-beta, small, growth stocks than do the low turnover households.

The gross and net returns for each turnover quintile are presented in Table V, Panels B and C. Focusing first on the gross performance (Panel B), we find that high turnover households (quintile 5) do not significantly outperform low turnover households (quintile 1). In fact, the intercept test based on the Fama–French three-factor model, which accounts for the tendency of the high turnover portfolio to tilt more heavily toward high-beta, small, growth stocks, indicates that the two high turnover quintiles (quintiles 4 and 5) underperform by 24 and 36 basis points per month. Though marginally statistically significant ($p$-values of 0.143 and 0.104, respectively), we believe these figures to be economically large (approximately three to four percent annually). Regardless of whether one accepts these results as statistically significant, the prediction of the Grossman and Stiglitz model is not supported; those who trade most do not earn higher gross returns.

The analysis of net returns (Panel C) is quite interesting. Regardless of the method used to measure performance, the high turnover households (quintile 5) underperform the low turnover households (quintile 1). The underperformance ranges from 46 basis points per month (5.5 percent per year, $t = -1.56$) using market-adjusted returns to an astoundingly high 80 basis points per month (9.6 percent per year, $t = -4.59$) based on the Fama–French intercept. The own-benchmark abnormal returns indicate that the

\begin{footnote}{Rather than increasing their gross returns, active traders could alternatively achieve the same expected utility as less active traders by lowering their volatility through trading. We find no evidence of this however. For example, the average (net) Sharpe ratio of the quintile that trades most actively (0.092) is one-half that of the quintile that trades least actively (0.180). Though these Sharpe ratios do not consider investors’ total portfolios of assets (see footnote 9), they indicate that active traders do not have higher volatility adjusted returns within the observed equity portfolios.}

\end{footnote}
Table V
Descriptive Statistics, Gross Returns, and Net Returns for Household Quintiles Formed on Mean Turnover

The sample is account records for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. Households are sorted into quintiles based on monthly turnover (the average of sales and purchase turnover) during our sample period. Quintile 1 contains households with the lowest turnover, quintile 5 contains households with the highest. Beginning position value is the market value of common stocks held in the first month that the household appears during our sample period. Mean monthly turnover is the average of sales and purchase turnover. Coefficient estimates are those from a time-series regression of the gross average household excess return on the market excess return ($R_{mt} - R_p$), ($HML_t$), and a zero-investment size portfolio ($SMB_t$). Raw return is the average monthly return for the average household. Own-benchmark abnormal return is the return on the household portfolio minus the return on the portfolio the household held at the end of the previous January. Market-adjusted return is the return on the household portfolio less the return on a value-weighted NYSE/AMEX/Nasdaq index. CAPM intercept is the estimated intercept from a time-series regression of the household excess return on the market excess return ($R_{mt} - R_p$). Fama–French intercept is the estimated intercept from time-series regressions of household excess return on the market excess return, a zero-investment book-to-market portfolio ($HML_t$), and a zero-investment size portfolio ($SMB_t$). $p$-values are presented in parentheses.

<table>
<thead>
<tr>
<th>Quintile</th>
<th>1 (Low)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 (High)</th>
<th>Difference: High – Low</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Descriptive Statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean monthly turnover (%)</td>
<td>0.19</td>
<td>1.24</td>
<td>2.89</td>
<td>5.98</td>
<td>21.49</td>
<td>N.A.</td>
</tr>
<tr>
<td>Mean beginning position value</td>
<td>34,169</td>
<td>26,046</td>
<td>22,945</td>
<td>19,102</td>
<td>21,560</td>
<td></td>
</tr>
<tr>
<td>Coefficient estimate on $(R_{mt} - R_p)$</td>
<td>1.03</td>
<td>1.06*</td>
<td>1.11**</td>
<td>1.18***</td>
<td>1.29***</td>
<td>0.26***</td>
</tr>
<tr>
<td>($HML_t$)</td>
<td>0.20***</td>
<td>0.10***</td>
<td>0.13**</td>
<td>0.13*</td>
<td>0.12</td>
<td>-0.08</td>
</tr>
<tr>
<td>($SMB_t$)</td>
<td>0.24***</td>
<td>0.29***</td>
<td>0.51***</td>
<td>0.72***</td>
<td>1.02***</td>
<td>0.78***</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>96.1</td>
<td>94.7</td>
<td>92.2</td>
<td>90.4</td>
<td>87.6</td>
<td>71.8</td>
</tr>
<tr>
<td><strong>Panel B: Gross Average Household Percentage Monthly Return</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw return</td>
<td>1.483</td>
<td>1.472</td>
<td>1.489</td>
<td>1.511</td>
<td>1.548</td>
<td>0.065</td>
</tr>
<tr>
<td>Own-benchmark</td>
<td>-0.009</td>
<td>-0.026*</td>
<td>-0.052**</td>
<td>-0.079***</td>
<td>-0.096*</td>
<td>-0.087</td>
</tr>
<tr>
<td>abnormal return</td>
<td>(0.156)</td>
<td>(0.064)</td>
<td>(0.014)</td>
<td>(0.007)</td>
<td>(0.093)</td>
<td>(0.116)</td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>0.063</td>
<td>0.052</td>
<td>0.069</td>
<td>0.091</td>
<td>0.125</td>
<td>0.065</td>
</tr>
<tr>
<td>CAPM intercept</td>
<td>(0.534)</td>
<td>(0.660)</td>
<td>(0.710)</td>
<td>(0.726)</td>
<td>(0.729)</td>
<td>(0.832)</td>
</tr>
<tr>
<td>Fama–French intercept</td>
<td>0.090</td>
<td>0.022</td>
<td>-0.015</td>
<td>-0.078</td>
<td>-0.167</td>
<td>-0.257</td>
</tr>
<tr>
<td></td>
<td>(0.409)</td>
<td>(0.865)</td>
<td>(0.936)</td>
<td>(0.774)</td>
<td>(0.663)</td>
<td>(0.407)</td>
</tr>
<tr>
<td><strong>Panel C: Net Average Household Percentage Monthly Return</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw return</td>
<td>1.470</td>
<td>1.411</td>
<td>1.361</td>
<td>1.267</td>
<td>1.099</td>
<td>-0.460</td>
</tr>
<tr>
<td>Own-benchmark</td>
<td>-0.021***</td>
<td>-0.079***</td>
<td>-0.167***</td>
<td>-0.300***</td>
<td>-0.587***</td>
<td>-0.566***</td>
</tr>
<tr>
<td>abnormal return</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>0.050</td>
<td>-0.009</td>
<td>-0.059</td>
<td>-0.153</td>
<td>-0.411</td>
<td>-0.460</td>
</tr>
<tr>
<td>CAPM intercept</td>
<td>(0.625)</td>
<td>(0.937)</td>
<td>(0.749)</td>
<td>(0.547)</td>
<td>(0.253)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>Fama–French intercept</td>
<td>0.077</td>
<td>-0.038</td>
<td>-0.140</td>
<td>-0.314</td>
<td>-0.692*</td>
<td>-0.768**</td>
</tr>
<tr>
<td></td>
<td>(0.480)</td>
<td>(0.764)</td>
<td>(0.474)</td>
<td>(0.242)</td>
<td>(0.066)</td>
<td>(0.012)</td>
</tr>
<tr>
<td></td>
<td>-0.061</td>
<td>-0.130</td>
<td>-0.269**</td>
<td>-0.464***</td>
<td>-0.864***</td>
<td>-0.802***</td>
</tr>
<tr>
<td></td>
<td>(0.422)</td>
<td>(0.172)</td>
<td>(0.037)</td>
<td>(0.005)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively (two-tailed). The null hypothesis for beta (the coefficient estimate on the market excess return) is $H_0: \beta = 1$ except in the difference column, where the null hypothesis is $H_0: \beta = 0$. 
trading of high turnover households costs them 57 basis points per month (6.8 percent per year) relative to the returns earned by low turnover households. Again, these differences are not consistent with the Grossman and Stiglitz model, but are consistent with the predictions of the overconfidence models.

In sum, differences in gross returns across the turnover quintiles are small. An investment mimicking that of the average household in each quintile would have earned a gross annualized mean geometric return that ranged from 18.5 percent (for quintile 2) to 18.7 percent (for quintile 1). However, there are dramatic differences in the net returns across the turnover quintiles. An investment mimicking the average household of the high turnover quintile would have earned a net annualized mean geometric return of 11.4 percent, while an investment that mimicked the low turnover quintile would have earned 18.5 percent. These returns are graphed in Figure 1.

V. Price Momentum

Some authors have identified price momentum effects in stock returns—that is, stocks that have performed well recently tend to earn higher returns than those that have not (Jegadeesh and Titman (1993)). It is unlikely, however, that individual investors view momentum as a risk factor. Thus, we do not include momentum when calculating risk-adjusted returns.

Nonetheless, it is interesting to consider how momentum affects the performance of individual investors. In general, the sampled investors are anti-momentum investors; that is, on average they tend to hold stocks that have recently underperformed the market. This is consistent with the evidence that individual investors tend to hold their losers and sell their winning investments (Odean (1998a)).

To investigate the effect of price momentum on the performance of individual investors, we add a zero-investment price-momentum portfolio to the Fama–French three-factor regressions described in Section II.C.22 This portfolio is long stocks that have performed well recently and short those that have performed poorly. We then estimate time-series regressions for each of the sample partitions described in the main text. In all sample partitions, the estimated coefficient estimate on the zero-investment price-momentum portfolio is negative; individuals tend to tilt their investments toward stocks that have performed poorly recently.

The net performance of individual investors in aggregate (on average) is $-0.053 (-0.041)$ percent per month when price momentum is included as an additional characteristic. Though still negative, these intercepts are smaller in magnitude than those from the Fama–French three-factor regressions and are not statistically significant.

22 The construction of the zero-investment price-momentum portfolio is described in Carhart (1997). We thank Mark Carhart for providing us with the returns data.
Our principal finding—that those investors who trade most actively realize, on average, the lowest net returns—is unaffected by the inclusion of a momentum characteristic in the regressions. These time-series regressions result in an intercept of \(-0.398\) percent per month for those who trade most actively (quintile 5) and \(0.070\) percent per month for those who trade least (quintile 1). Thus, when one controls for their tendency to hold poorly performing stocks, those investors who trade least actively achieve reasonable performance. More important, however, is the finding that active investors continue to underperform less active investors. The differences in the intercepts remains large and statistically significant: \(-0.468\) percent per month.

VI. Liquidity, Rebalancing, and Tax-Motivated Trading

To this point, we have focused on information-motivated versus overconfidence-motivated trading. The empirical evidence we have presented solidly favors overconfidence as the major motivation for trading, since trading unambiguously hurts investor performance; however, there are other motivations for trading, which we consider in this section.

A. Liquidity

Investors who face liquidity shocks over time will trade as a rational response to those shocks. Thus, liquidity shocks can explain some trading activity. But, they seem implausible as an explanation of the 75 percent annual turnover that we document for the average individual investor and belie common sense as an explanation of the more than 250 percent annual turnover of the households who trade most. Investors facing rapidly fluctuating liquidity needs can, in most cases, find less expensive means to finance these than rapid trading in and out of stocks.

Moreover, the trading that results from liquidity shocks can be accomplished at a much lower cost by investing in mutual funds than by investing in individual common stocks. To illustrate this point, we analyze the returns on the Vanguard Index 500 mutual fund, a large passive mutual fund that claims to match the performance of the Standard and Poor’s 500. Investors can move in and out of this fund at no cost. In contrast to the performance of the average or aggregate household, this index fund does not underperform when compared to any of the standard performance benchmarks. During our sample period, this fund earned an annualized geometric mean return of 17.8 percent while the value-weighted market index earned 17.9 percent. The market-adjusted return, the CAPM intercept, and the Fama–French intercept for the Vanguard Index 500 were \(-0.002\), \(-0.004\), and \(0.009\) percent, respectively. A passively managed mutual fund clearly provides a lower cost means of managing liquidity shocks than does investment in individual common stocks.
B. Rebalancing

Investors who desire a portfolio with certain risk characteristics will rationally rebalance their portfolio to maintain this risk profile. With an average holding of four common stocks, we believe that risk-based rebalancing is not a significant motivation for trading in the households that we study. Risk-based rebalancing as an explanation of the 75 percent annual turnover that we document for the average household belies common sense. Investors can manage the risk composition of their portfolio at much lower cost by carefully selecting a portfolio of mutual funds.

C. Taxes

The single most compelling reason for investors to hold individual common stocks in lieu of mutual funds is taxes. Investors who hold stocks that have lost value since their purchase can realize those losses. These losses can be used to shelter gains and thereby reduce the investor’s tax liability.°

Tax-loss selling cannot completely explain the results that we document here for three reasons. First, it is implausible that tax-motivated trading would yield an annual turnover rate of 75 percent. A simple example illustrates this point: Consider an investor who buys the value-weighted market index on January 1 of each year 1991 to 1996. In December of the average year, this investor would be able to sell 24 percent of her portfolio for a loss. Of course, this example assumes a holding period of 12 months. The turnover resulting from tax-loss selling will decline as this holding period increases.

Second, we find high turnover and significant underperformance in both taxable and tax-deferred accounts. If tax-loss selling is the major motivation for trading we would expect to find little trading in tax-deferred accounts. On the other hand, if overconfidence is the major motivation for trading, we would expect to find, as we do, active trading and significant underperformance in both taxable and tax-deferred accounts. We partition the accounts in our sample into taxable and tax-deferred accounts (i.e., Individual Retirement Accounts and Keogh Accounts). In Table VI, Panel A, we present descriptive statistics for the taxable and tax-deferred accounts. Turnover in tax-deferred accounts is high: 67.6 percent annually (monthly turnover of 5.63 percent times 12), though not as high as in taxable accounts: 89.4 percent annually (monthly turnover of 7.45 percent times 12). The difference in turnover may result from tax-motivated trading or it may be that investors associate their retirement accounts with future safety and therefore trade less speculatively in these accounts.

In Table VI, Panels B and C, we present the gross and net return performances of taxable and tax-deferred accounts. The gross returns earned by taxable and tax-deferred accounts are quite similar (see Panel B). The net

° Though losses on mutual funds can also be used to reduce an investor’s tax liability, the probability of having a loss on a mutual fund is less than the probability of observing at least one losing investment in a well-diversified portfolio of common stocks.
Descriptive Statistics, Gross Return, and Net Return for Taxable and Tax-Deferred Accounts

The sample is account records for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. Accounts are partitioned as either taxable or tax deferred (IRA, Keogh, SEP-IRA). Beginning position value is the market value of common stocks held in the first month that the household appears during our sample period. Mean monthly turnover is the average of sales and purchase turnover. Coefficient estimates are those from a time-series regression of the gross average household excess return on the market excess return ($\Delta R_{mt}$ - $\Delta R_{ft}$), a zero-investment book-to-market portfolio ($HML_t$), and a zero-investment size portfolio ($SMB_t$). Raw return is the average monthly return for the average household. Own-benchmark abnormal return is the return on the household portfolio minus the return on the portfolio the household held at the end of the previous January. Market-adjusted return is the return on the household portfolio less the return on a value-weighted NYSE/AMEX/Nasdaq index. CAPM intercept is the estimated intercept from a time-series regression of the household excess return on the market excess return ($\Delta R_{mt}$ - $\Delta R_{ft}$). Fama–French intercept is the estimated intercept from time-series regressions of household excess return on the market excess return, a zero-investment book-to-market portfolio ($HML_t$), and a zero-investment size portfolio ($SMB_t$). *p*-values are presented in parentheses.

### Panel A: Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Taxable</th>
<th>Tax Deferred</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of households</td>
<td>54,434</td>
<td>30,554</td>
<td>N/A</td>
</tr>
<tr>
<td>Mean beginning position value</td>
<td>26,303</td>
<td>14,042</td>
<td>12,261***</td>
</tr>
<tr>
<td>Mean monthly turnover (%)</td>
<td>7.45</td>
<td>5.63</td>
<td>1.82*** (0.000)</td>
</tr>
<tr>
<td>Coefficient estimate on:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>($\Delta R_{mt}$ - $\Delta R_{ft}$)</td>
<td>1.13*** (0.004)</td>
<td>1.12*** (0.007)</td>
<td>0.01 (0.346)</td>
</tr>
<tr>
<td>$HML_t$</td>
<td>0.14*** (0.010)</td>
<td>0.18*** (0.001)</td>
<td>-0.04*** (0.000)</td>
</tr>
<tr>
<td>$SMB_t$</td>
<td>0.56*** (0.000)</td>
<td>0.52*** (0.000)</td>
<td>0.04*** (0.000)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>92.6</td>
<td>92.2</td>
<td>46.7</td>
</tr>
</tbody>
</table>

### Panel B: Gross Average Household Percentage Monthly Return

<table>
<thead>
<tr>
<th></th>
<th>Taxable</th>
<th>Tax Deferred</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw return</td>
<td>1.496</td>
<td>1.532</td>
<td>-0.036</td>
</tr>
<tr>
<td>Own-benchmark abnormal return</td>
<td>-0.048*** (0.009)</td>
<td>-0.037* (0.055)</td>
<td>-0.010 (0.107)</td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>0.076</td>
<td>0.112</td>
<td>-0.036</td>
</tr>
<tr>
<td>CAPM intercept</td>
<td>-0.027</td>
<td>0.031</td>
<td>-0.058** (0.899)</td>
</tr>
<tr>
<td>Fama–French intercept</td>
<td>-0.174</td>
<td>-0.133</td>
<td>-0.041* (0.174)</td>
</tr>
</tbody>
</table>

### Panel C: Net Average Household Percentage Monthly Return

<table>
<thead>
<tr>
<th></th>
<th>Taxable</th>
<th>Tax Deferred</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw return</td>
<td>1.313</td>
<td>1.379</td>
<td>-0.066**</td>
</tr>
<tr>
<td>Own-benchmark abnormal return</td>
<td>-0.203*** (0.583)</td>
<td>-0.166*** (0.823)</td>
<td>-0.036*** (0.012)</td>
</tr>
<tr>
<td>Market-adjusted return</td>
<td>-0.107</td>
<td>-0.042</td>
<td>-0.066**</td>
</tr>
<tr>
<td>CAPM intercept</td>
<td>-0.204</td>
<td>-0.119</td>
<td>-0.085***</td>
</tr>
<tr>
<td>Fama–French intercept</td>
<td>-0.344*** (0.008)</td>
<td>-0.278*** (0.030)</td>
<td>-0.066*** (0.002)</td>
</tr>
</tbody>
</table>

***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively (two-tailed). The null hypothesis for beta (the coefficient estimate on the market excess return) is $H_o: \beta = 1$ except in the difference column, where the null hypothesis is $H_o: \beta = 0$. 

---

Trading Is Hazardous to Your Wealth 797
returns earned by taxable and tax-deferred accounts are both poor, after a reasonable accounting for the small stock tilt of these individuals (see Panel C). The tax-deferred accounts outperform the taxable accounts by about six basis points per month. In short, the general tenor of our results is similar for the taxable and tax-deferred accounts.

Third, Odean (1998a, 1999) documents that most investor trading activity is inconsistent with tax-motivated trading. He observes that investors at a discount brokerage sell profitable investments twice as often as unprofitable investments (during the period 1987 to 1993) and that, relative to their opportunities to do so, these investors are about one and one-half times more likely to realize any gain than any loss. They do engage in tax-loss selling late in the year, but December is the only month in which they realize losses at as fast a rate as they do gains.

Finally, we should emphasize that trading not associated with tax-loss selling will further hurt the after-tax returns of individual investors. Not only does this trading incur trading costs, when done in a taxable account it also accelerates the payment of capital gain taxes that could be otherwise deferred.

D. Gambling

To what extent may a desire to gamble account for the excessive trading we observe? Many people appear to enjoy gambling. Some buy lottery tickets. Others gamble at casinos. We consider two distinct aspects of gambling: risk-seeking and entertainment. Risk-seeking is when one demonstrates a preference for outcomes with greater variance but equal or lower expected return. In equity markets the simplest way to increase variance without increasing expected return is to underdiversify. Excessive trading has a related, but decidedly different, effect; it decreases expected returns without decreasing variance. Thus risk-seeking may account for underdiversification (though underdiversification could also result from simple ignorance of its benefits), but it does not explain excessive trading.

A second aspect of gambling is the entertainment derived from placing and realizing bets. When coupled with the overconfident belief that these bets are expected-wealth enhancing, it is easy to see that the entertainment utility of gambling will fuel greater trading. There is also the possibility that people may trade for entertainment while fully realizing that each trade is more likely than not to reduce their personal future wealth. (Note that this is different from realizing that the trades of others are wealth reducing.) We favor the hypothesis that most investors trade excessively because they are overconfident, or because they are overconfident and they enjoy trading, over the hypothesis that they trade purely for entertainment and expect thereby to lower their wealth. Many studies have established that people are overconfident. We know of no study demonstrating that ordinary investors expect to lower their wealth through trading.
It is possible that some investors set aside a small portion of their wealth with which they trade for entertainment, while investing the majority more prudently. If “entertainment accounts” are driving our findings, we would expect turnover and underperformance to decline as the common stocks in the accounts we observe represent a larger proportion of a household’s total wealth. We are able to test this hypothesis directly and find no support for it. For approximately one-third of our sample, the households reported their net worth at the time they opened their accounts. We calculate the proportion of net worth invested at the discount broker as the beginning value of a household’s common stock investments scaled by its self-reported net worth. We then analyze the turnover and investment performance of 2,333 households with at least 50 percent of their net worth in common stock investments at this discount broker. These households have similar turnover (6.25 percent per month, 75 percent annually) to our full sample (see Table I). Furthermore, these households earn gross and net returns that are very similar to the full sample. The monthly net return, own-benchmark abnormal return, market-adjusted return, CAPM intercept, and Fama–French intercept for these households are 1.285, −0.173, −0.135, −0.221, and −0.285 percent, respectively.

Finally, it is worth noting that the negative relation between turnover and net returns that we document for individual investors also exists in mutual funds (Carhart (1997)). It is unlikely that mutual fund managers buy and sell stocks for the pure joys of trading despite the fact that this trading lowers the expected returns of their shareholders.25

VII. Conclusion

We analyze the returns earned on common stock investments by 66,465 households at a large discount brokerage firm for the six years ending in January 1997. We document that the gross returns (before accounting for transaction costs) earned by these households are quite ordinary, on average. Unfortunately, the net returns (after accounting for the bid-ask spread and commissions paid by these investors) earned by these households are poor. The average household underperforms a value-weighted market index by about 9 basis points per month (or 1.1 percent annually). After accounting for the fact that the average household tilts its common stock investments toward small value stocks with high market risk, the underperformance averages 31 basis points per month (or 3.7 percent annually). The average household turns over approximately 75 percent of its common stock portfolio annually. The poor performance of the average household can be traced to the costs associated with this high level of trading.

24 This estimate is upwardly biased because the account opening date generally precedes our first portfolio position observation and net worth is likely to have increased in the interim.

25 Lakonishok et al. (1992) report a positive relation between turnover and performance for 769 all-equity pension funds, though this finding puzzles the authors.
Our most dramatic empirical evidence is provided by the 20 percent of households that trade most often. With average monthly turnover of in excess of 20 percent, these households turn their common stock portfolios over more than twice annually. The gross returns earned by these high-turnover households are unremarkable, and their net returns are anemic. The net returns lag a value-weighted market index by 46 basis points per month (or 5.5 percent annually). After a reasonable accounting for the fact that the average high-turnover household tilts its common stock investments toward small value stocks with high market risk, the underperformance averages 86 basis points per month (or 10.3 percent annually).

The investment experience of individual investors is remarkably similar to the investment experience of mutual funds. As do individual investors, the average mutual fund underperforms a simple market index (Jensen (1969) and Malkiel (1995)). Mutual funds trade often and their trading hurts performance (Carhart (1997)). But trading by individual investors is even more deleterious to performance because individuals execute small trades and face higher proportional commission costs than mutual funds.

Our main point is simple: Trading is hazardous to your wealth. Why then do investors trade so often? The aggregate turnover of the individual investor portfolios we analyze is about 70 percent; the average turnover is about 75 percent. The New York Stock Exchange reports that the annual turnover of stocks listed on the exchange hovered around 50 percent during our sample period. Mutual funds average an annual turnover of 77 percent (Carhart (1997)). We believe that these high levels of trading can be at least partly explained by a simple behavioral bias: People are overconfident, and overconfidence leads to too much trading.

Based on rational agents free from such behavioral biases, the efficient markets hypothesis has been central to both the theory and practice of investment management. The efficiency research posits that private information is rare. Thus, active investment strategies will not outperform passive investment strategies. Both the theoretical and empirical work on efficiency supporting this view have led to a rise of passive investment strategies that simply buy and hold diversified portfolios (Fama (1991)).

Behavioral finance models that incorporate investor overconfidence (e.g., Odean (1998b)) provide an even stronger prediction: Active investment strategies will underperform passive investment strategies. Overconfident investors will overestimate the value of their private information, causing them to trade too actively and, consequently, to earn below-average returns. Consistent with these behavioral models of investor overconfidence, we provide empirical evidence that households, which hold about half of U.S. equities, trade too much, on average. Those who trade the most are hurt the most.

Appendix A. The Analysis of Trade Timing

In this appendix, we analyze the timing of purchases and sales within a month. The timing of trades within a month is ignored in our main analysis where we assume all purchases and sales are made at month end.
Consistent with the results reported in Odean (1999), we document that the stocks investors buy subsequently underperform the stocks they sell. In aggregate, we estimate that an exact accounting for the timing of purchases and sales would reduce the performance of individual investors by more than two basis points per month (or approximately 0.29 percent annually).

For each account with a beginning-of-month position statement in month \( t \), we identify all purchases in month \( t - 1 \) and sales in month \( t \). For both purchases and sales, we calculate the compound return on the stock from the day following the purchase to the end of the month less the compound return on the value-weighted NYSE/AMEX/Nasdaq market index. Sales turnover and sales abnormal return are analogously calculated. The estimated effect on the monthly abnormal return is the purchase turnover times the purchase abnormal return minus the sale turnover times the sale abnormal return.

### Table AI

**The Gross Abnormal Returns for Stocks Bought and Sold from the Trade Date to the End of the Month**

The sample is account records for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. Purchase turnover is the average value of stocks purchased divided by the average value of stocks held in each month. The purchase abnormal return is calculated by compounding the daily returns on the purchased security from the day following the purchase to the end of the month less the compound return on the value-weighted NYSE/AMEX/Nasdaq market index. Sales turnover and sales abnormal return are analogously calculated. The estimated effect on the monthly abnormal return is the purchase turnover times the purchase abnormal return minus the sale turnover times the sale abnormal return.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Monthly Purchase Turnover (%)</th>
<th>Purchase Abnormal Return (%)</th>
<th>Monthly Sale Turnover (%)</th>
<th>Sale Abnormal Return (%)</th>
<th>Estimated Effect on Monthly Abnormal Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Aggregate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All households</td>
<td>4.92</td>
<td>−0.472</td>
<td>4.93</td>
<td>0.021</td>
<td>−0.0242</td>
</tr>
<tr>
<td><strong>Panel B: Households Partitioned by Beginning Position Value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Small)</td>
<td>6.85</td>
<td>−0.650</td>
<td>6.06</td>
<td>−0.116</td>
<td>−0.0375</td>
</tr>
<tr>
<td>2</td>
<td>5.83</td>
<td>−0.381</td>
<td>5.16</td>
<td>−0.019</td>
<td>−0.0213</td>
</tr>
<tr>
<td>3</td>
<td>5.82</td>
<td>−0.386</td>
<td>5.25</td>
<td>0.437</td>
<td>−0.0454</td>
</tr>
<tr>
<td>4</td>
<td>5.55</td>
<td>−0.445</td>
<td>5.25</td>
<td>0.030</td>
<td>−0.0263</td>
</tr>
<tr>
<td>5 (Large)</td>
<td>4.41</td>
<td>−0.486</td>
<td>4.23</td>
<td>−0.035</td>
<td>−0.0199</td>
</tr>
<tr>
<td><strong>Panel C: Households Partitioned by Turnover</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Low)</td>
<td>0.26</td>
<td>−0.184</td>
<td>0.23</td>
<td>0.068</td>
<td>−0.0006</td>
</tr>
<tr>
<td>2</td>
<td>1.37</td>
<td>−0.176</td>
<td>1.14</td>
<td>−0.089</td>
<td>−0.0014</td>
</tr>
<tr>
<td>3</td>
<td>3.07</td>
<td>−0.126</td>
<td>2.57</td>
<td>0.041</td>
<td>−0.0049</td>
</tr>
<tr>
<td>4</td>
<td>6.46</td>
<td>−0.234</td>
<td>6.13</td>
<td>0.102</td>
<td>−0.0214</td>
</tr>
<tr>
<td>5 (High)</td>
<td>21.81</td>
<td>−0.674</td>
<td>20.75</td>
<td>−0.003</td>
<td>−0.1464</td>
</tr>
</tbody>
</table>

Trading Is Hazardous to Your Wealth
The results of our analysis are presented in Table AI. The second (fourth) columns of this table present aggregate purchase (sale) turnover calculated as the aggregate dollar value of purchases (sales) divided by the aggregate dollar value of positions held. (This turnover measure is slightly different from that used in the main text, where turnover is calculated based on market values contained in position statements and is thus capped at 100 percent per month for each household.) Abnormal returns are calculated for purchases and sales by subtracting the compound return on the CRSP NYSE/AMEX/Nasdaq value-weighted index. The trade-weighted mean abnormal returns are presented in columns 3 (for purchases) and 5 (for sales) of Table AI. In aggregate (Panel A), from the day following the trade to the end of the month, the stocks that investors buy underperform the value-weighted market index by 47 basis points, and those they sell outperform the index by two basis points. Based on these abnormal returns and our estimates of aggregate turnover, we calculate that the results we present in the main text overestimate the performance of individual investors by 2.42 basis points per month.

We also analyze the timing of trades by partitioning households on the basis of account size (Panel B) and turnover (Panel C). In each of the sample partitions, the timing of their trades hurts investors. In short, the results in the main text overestimate the performance of individual investors by ignoring the exact timing of purchases and sales.

Consider how the accounting for the exact timing of trades relates to the return calculations contained in the main text. In Figure A1, we present an example of a security that is purchased in month 1 and sold in month 3. A time line for these transactions is depicted in Figure A1.

In the main text, we calculate the return for this security from $t_1$ to $t_3$. In this appendix, we calculate the return from timing as the return from $t_b^{cl}$ to $t_1$ minus the return from $t_s^{cl}$ to $t_s$. Our estimate of the bid-ask spread is the return from $t_s$ to $t_s^{cl}$ minus the return from $t_b$ to $t_b^{cl}$. When the return from timing is added to the main calculation and the spread is subtracted, one gets the (approximate) return from $t_b$ to $t_s$, the period in which the investor held the stock.

![Figure A1. Time line of returns calculations. The time of purchase (sale) is $t_b$ ($t_s$). The close on the purchase (sale) day is $t_b^{cl}$ ($t_s^{cl}$). The close on the last day of the purchase (sale) month is $t_1$ ($t_3$).](image-url)
Appendix B. The Analysis of Intramonth Trades

In this appendix, we analyze the performance of stocks that are bought and then sold within a calendar month (e.g., purchased on January 3 and sold on January 10). These intramonth trades are excluded from our main analyses, since those analyses are based on monthly position statements. In aggregate, we estimate that intramonth trades would improve the performance of individual investors by less than one basis point per month (or approximately 0.06 percent annually). Though profitable, the aggregate value of intramonth trades accounts for less than one percent of the aggregate value of positions held.

For each account, we identify all purchases followed by a sale within the same month. In accounting for multiple purchases and sales, we assume that the first securities purchased are the first sold. Over our 72-month sample period, we identify 87,095 round-trip intramonth trades worth approximately $27 million per month, on average. In contrast, the average beginning-of-month value of positions held, which we analyze in the main text, is over $2.7 billion.

We calculate the gross returns on these round-trip transactions using the CRSP daily return files assuming the security is purchased and sold at the close of trading on the purchase and sale dates, respectively. We calculate the net returns on these round-trip transactions by subtracting estimates of the bid-ask spread and commissions as is done in the main text for the case of monthly returns. The average round-trip trade involves a purchase of $22,275, is held for 6.16 days, and costs 2.08 percent in commissions and 0.30 percent for the bid-ask spread. (In aggregate, these round-trip trades cost 0.87 percent in commissions and 0.27 percent for the bid-ask spread.) Note that the bid-ask spread is lower than that documented for trades that we analyze in the main text, which have an average round-trip bid-ask spread of one percent (see Table I). This lower spread is likely a result of the intraday return earned by investors from the transaction price through the end of the trading day (which is included in our estimate of the spread) rather than a smaller bid-ask spread for these intramonth trades.

In Table BI, we summarize our analysis of the gross and net returns earned on intramonth trades. In this table, we calculate market-adjusted abnormal returns by subtracting the daily value-weighted NYSE/AMEX/Nasdaq CRSP market index from the return earned on each intramonth trade. Both the gross and net abnormal returns in this table are weighted by the size of each trade, so that we can estimate the aggregate impact of these intramonth trades on the performance of individual investors.

Panel A presents results for all households. In aggregate, the intramonth trades earn impressive gross abnormal returns of 1.64 percent. The net abnormal returns are 0.50 percent. Since these intramonth trades average 0.99 percent of the average value of positions held, we estimate that these intramonth trades would improve the performance of individual investors by 0.49 basis points per month (0.0050 times 0.0099) in aggregate. This small improvement in performance does not affect any of the conclusions that we present in the main text.
We also analyze the profitability of intramonth trades by partitioning households on the basis of account size (Panel B) and turnover (Panel C). In short, none of these results are so dramatic that they would lead us to qualify any of the results that we present in our main text. Those who benefit most from intramonth trades are those who trade most. Their intramonth trades improve their performance by 3.12 basis points per month (last row and last column of Panel C). Yet, we estimate that these investors underperform by a whopping 86 basis points per month (last row, Table V).

In conclusion, we emphasize that the positive net returns earned on intramonth trades do not necessarily imply that individual investors have superior short-term trading ability. If investors have a disposition to sell winning investments and ride losing investments (as proposed by Shefrin and Statman (1985)), we would expect to observe positive abnormal returns on short-term round-trip trades.

### Table BI

The Gross and Net Abnormal Returns earned on Intramonth Trades

The sample is account records for 66,465 households at a large discount brokerage firm from January 1991 to December 1996. The gross abnormal return on intramonth trades is calculated as the compound return from the day following the purchase to the day of the sale less the compound return on a value-weighted NYSE/AMEX/Nasdaq index. The net abnormal return is the gross abnormal return adjusted for the return earned on the day of the purchase or sale, the bid-ask spread, and the commission cost. The intramonth trades as a percentage of total position value are the average monthly value of intramonth purchases divided by the average monthly value of all stocks held. The estimated effect on monthly abnormal return is the net abnormal return times the intramonth trades as a percentage of total position value.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean Trade Size</th>
<th>Gross Abnormal Return (%)</th>
<th>Net Abnormal Return (%)</th>
<th>Intramonth Trades as a Percentage of Total Position Value</th>
<th>Estimated Change in Monthly Abnormal Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All households</td>
<td>$22,275</td>
<td>1.636</td>
<td>0.496</td>
<td>0.99</td>
<td>0.0049</td>
</tr>
<tr>
<td>Panel B: Households Partitioned by Beginning Position Value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Small)</td>
<td>17,459</td>
<td>2.376</td>
<td>0.904</td>
<td>1.42</td>
<td>0.0128</td>
</tr>
<tr>
<td>2</td>
<td>12,579</td>
<td>2.082</td>
<td>0.248</td>
<td>0.92</td>
<td>0.0023</td>
</tr>
<tr>
<td>3</td>
<td>17,173</td>
<td>1.757</td>
<td>0.486</td>
<td>1.17</td>
<td>0.0057</td>
</tr>
<tr>
<td>4</td>
<td>20,255</td>
<td>1.363</td>
<td>0.351</td>
<td>1.33</td>
<td>0.0046</td>
</tr>
<tr>
<td>5 (Large)</td>
<td>28,387</td>
<td>1.563</td>
<td>0.526</td>
<td>0.86</td>
<td>0.0045</td>
</tr>
<tr>
<td>Panel C: Households Partitioned by Turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Low)</td>
<td>10,638</td>
<td>−0.003</td>
<td>−0.026</td>
<td>0.00</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>12,876</td>
<td>3.006</td>
<td>0.200</td>
<td>0.02</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>11,886</td>
<td>1.843</td>
<td>0.220</td>
<td>0.08</td>
<td>0.0002</td>
</tr>
<tr>
<td>4</td>
<td>13,838</td>
<td>2.925</td>
<td>1.378</td>
<td>0.36</td>
<td>0.0050</td>
</tr>
<tr>
<td>5 (High)</td>
<td>23,702</td>
<td>1.545</td>
<td>0.451</td>
<td>6.92</td>
<td>0.0312</td>
</tr>
</tbody>
</table>
REFERENCES


Barber, Brad M., Reuven Lehavy, Maureen McNichols, and Brett Trueman, 1998, Can investors profit from the prophets? Consensus analyst recommendations and stock returns, Working paper, Graduate School of Management, University of California, Davis.


PART 4: FINANCIAL INSTRUMENTS

4.1 INTRODUCTION

This part considers the current regulation of securities and derivatives markets and other investment products and some of the difficulties caused by the current regulatory regimes. It proposes a new integrated regulatory regime for all financial instruments. Elements of the new regime are considered in more detail in the remaining parts of the paper.

4.2 CURRENT REGULATION OF SECURITIES AND DERIVATIVES MARKETS

The current regulation of Australia’s securities and derivatives markets is largely institution and product based. Financial instruments are regulated differently depending on how the instruments are legally classified.

Market regulation draws a distinction between ‘securities’ and ‘futures contracts.’ Generally, financial arrangements falling within the definition of ‘securities’ are traded on a securities exchange, while arrangements within the definition of a ‘futures contract’ are traded on a futures exchange or an OTC futures market. ‘Derivatives’ are not recognised as a distinct category by the Corporations Law and are regulated differently depending on whether they are classified as a security or a futures contract.

32 Corporations Law, section 92 basically defines securities as shares, debentures, bonds, prescribed interests and options over these instruments.
33 The Corporations Law, section 72 provides a technical and elaborate definition of a futures contract. Basically, adjustment agreements (contracts for differences), commodity agreements and options over these two instruments fall within the definition.
34 Banks trading currency swaps, interest rate swaps, forward exchange and forward interest contracts are not regulated by the Corporations Law. Section 72(1)(d) of the Law excludes these contracts from the definition of futures contract.
4.3 PROBLEMS WITH CURRENT REGULATION OF SECURITIES AND DERIVATIVES MARKETS

The current regulation of securities and futures markets has not adequately accommodated market developments or financial innovation. The definition of ‘futures contract’ is widely acknowledged as unsatisfactory and the distinction between securities and futures is challenged by innovative financial products which exhibit characteristics of both types of instruments.

Inhibits Competition

The legal distinction between securities and futures inhibits competition between market providers and creates barriers to entry. Markets in innovative financial instruments which challenge the legal definitions of securities and futures provide a striking example. For example, if the ASX wishes to conduct trading in a derivative which is legally classified as a futures contract it must either:

- seek authorisation as a futures exchange; or
- seek regulations to permit the product to be traded on the ASX as if it were a securities contract.\(^{35}\)

The SFE is in a similar position if it proposes to introduce a derivative which falls within the definition of a security.\(^{36}\)

Regulations to facilitate the trading of financial products on an exchange require time to develop and implement. The opportunity to develop a new product in a timely manner is lost and the delay also permits competitors to take strategic action.

---

35 The Law was amended in 1995 by the *Corporations Law (Futures and Securities) Amendment Act 1994* (‘the Amendment Act’) which commenced on 12 April 1995 to introduce a measure of flexibility in the regulation of new financial products developed by the exchanges. The Amendment Act permits the making of regulations to prescribe certain exchange traded agreements and provide a tailored regulatory regime for those agreements. The Amendment Act was an interim response designed to facilitate innovation and competition within the securities and futures industries pending the CASAC review of the current regulatory requirements for derivatives under the Corporations Law.

36 For example, as a result of *Sydney Futures Exchange Limited v Australian Stock Exchange Limited* (1995) 16 ACSR 148 (the LEPOs case) deliverable share futures and options over shares fall outside the definition of futures contracts.
Market participants have attempted to compete within the existing regulatory framework and have also structured products to take advantage of a particular regulatory regime which may create market distortions.\(^\text{37}\)

**Lack of Certainty**

There is some uncertainty among market participants as to whether new financial arrangements are futures contracts or securities. An illustration of that uncertainty is provided by the recent legal action in relation to the ASX’s derivative initiative, Low Exercise Price Options (LEPOs). The Full Federal Court noted that the definitions of ‘securities’ and ‘futures contracts’ are ill suited in determining the nature of a new financial arrangement.\(^\text{38}\) Nevertheless, the product was considered to be a security, although the Court did recognise that in substance, a LEPO is very similar to a futures contract.

**Traditional View**

The separate regulatory regimes for securities and futures markets may be attributed to the independent development of the Securities Industry Act and the Futures Industry Act. However, there is some support for separate regulation based on the traditional view that there are fundamental distinctions between particular financial instruments.

A new regulatory distinction between ‘securities’ and ‘derivatives’ would result in more consistent regulation of derivative products (eg options, warrants and futures). The development of this distinction would be based on the traditional view that there is a fundamental difference between fundraising (securities) and risk management agreements (futures/derivatives). That is, there is a fundamental difference between the role of stock exchanges in transferring title to securities and the role of futures/derivatives exchanges in transferring price risks through contractual arrangements.

---

37 For example, products may be structured to take advantage of a particular regulatory regime. It may also be possible to structure a futures contract as an option so that it can be traded on a securities exchange. It is also possible to structure products to avoid regulation. For example, swaps where payments are made on a gross rather than a net basis may not be either securities or futures contracts.

However, the traditional view that there is a fundamental distinction between the role of securities and derivatives markets is under challenge from a number of perspectives including that:

- derivatives are traded on both the ASX and the SFE;
- substantially similar products are traded on the ASX as securities and the SFE as futures contracts (eg ASX LEPOs and SFE deliverable share futures contracts);
- some derivatives give rise to a transfer of title to a physical asset if held until expiry (eg equity options, deliverable futures contracts);
- new financial arrangements have been developed which exhibit traditional characteristics of both ‘securities’ and ‘futures’ — eg endowment warrants which permit holders of a derivative instrument to benefit from dividends paid on the underlying security.

Commercial practice does not accord with the traditional view that securities are investment products while derivatives are risk management tools. Currently, investors use related derivatives and securities products as alternatives to achieving a particular investment strategy. For example, an investor wishing to risk capital for future profit based on movements in the price of shares has a number of choices available: to purchase the share, take an option over it, take out a warrant or buy an individual share future. In order to diversify the investment to reduce risk, the same investor may purchase a share portfolio or a share price index futures contract. All transactions represent risk-taking on the future movement of share prices.

The extent to which a person is risk averse will influence the type of financial instrument they wish to purchase. For example, some investors prefer ‘safe’ investments where the capital value is relatively fixed but long term returns tend to be relatively low. Others will invest in riskier investments where the expected long term return is relatively high but there is a risk of losing part of the initial capital investment. Risk can be managed in a number of ways, including portfolio diversification which allows the subdivision and broadening across a range of investments. Alternatively, instruments which limit down-side risk, like options over shares, permit a risk-averse investor to participate in profits while only risking a small premium.

The creation of a new distinction based on ‘securities’ and ‘derivatives’ does not appear to meet the policy objectives outlined in Part 3 of this paper. Regulation based on legal definitions is inherently inflexible and, in time, is likely to become redundant given the rapid changes that have been experienced in financial markets. Further, it would result in inefficiencies, as a market operator would be required to seek dual authorisation as a securities...
exchange and a derivatives exchange in order to conduct trading in both types of instruments (eg shares and options over shares).

**Economic Functions Of Securities And Derivatives**

From an economic perspective, the creation of a new regulatory regime based on the separate regulation of securities and derivatives would be inefficient as it would regulate functionally similar products differently depending on their legal classification.

The economic functions of securities and derivatives markets are similar, notwithstanding the varying characteristics of particular instruments. As noted in Part 3, the fundamental functions of financial markets include price discovery and the management and pricing of risk. Both securities and derivatives markets perform a price discovery function through the exchange and evaluation of information.\(^3^9\) Stock index futures markets now often function as the dominant price discovery mechanism for the stock market.\(^4^0\)

The similar economic functions of securities and derivatives becomes clear from an examination of the intrinsic pricing links between securities and derivatives markets. The prices of an underlying asset and the related derivative contract are simultaneously determined due to arbitrage where differences in one market are quickly transmitted to related markets. The similar price behaviour between derivatives and securities markets provides investors with real alternatives to achieve investment strategies. For example, warrants may be used to achieve capital gains from an equity issue without legally owning the underlying securities. Investors commonly use derivatives as alternatives to holding physical assets.

---


A further pricing link between securities and derivatives is illustrated by put-call parity. This pricing relationship permits investors to synthesise particular financial instruments through a combination of other instruments. Put-call parity is illustrated in Chart 7 which represents the cash flows of a bought gold futures contract for $400 and related put and call options at a $20 premium with a $400 strike price. That is, the outcome in the chart above is that the bought futures contract (A) equals the sum of a bought call (B) and a sold put (C). Represented as an equation, $A = B + C = D$.

Similarly, a securities position may be replicated by a position in a risk-free bond plus an equity futures contract. A bought position in a call can be replicated with a bought position in the security and a bought put. The relationship between various financial instruments means that investors may achieve their investment strategies without being confined to traditional institutional arrangements or products. For example, innovative uses of
derivative instruments may be used to achieve the same economic outcome which was previously achieved through the buying and selling of assets.\footnote{See also Merton, \textit{Financial Innovation and the Management and Regulation of Financial Institutions}, Working Paper No 5096.}

4.4 CURRENT REGULATION OF OTHER FINANCIAL PRODUCTS

The regulation of other financial products is governed by various Acts and Codes of Practice.\footnote{For example, the ABA Code of Banking Practice, Life Circulars Nos. G.I.1 (disclosure requirements) and G.II.1 (code of practice), and the General Insurance Codes of Practice for companies and brokers.} As discussed in some detail in Part 5, regulation is piecemeal and varied in its regulation of intermediaries and product disclosure requirements. This results in unnecessary compliance costs for industry participants as well as additional administration costs for regulators. The different regulatory regimes contribute to investor confusion and create some uncertainty for intermediaries who must satisfy different regulatory requirements depending on the classification of the particular financial product.

The convergence between different financial products is not limited to securities and derivatives. Increasingly, investors have a choice of products available to achieve particular investment strategies, eg market linked life company products, managed investments and public offer superannuation products may all be used to achieve similar economic outcomes. The put-call parity relationship demonstrated in Chart 7 can be rearranged in a number of ways to mimic particular products.\footnote{For example, see Department of the Treasury, \textit{Taxation of Financial Arrangements: An Issues Paper}, December 1996, pp 14-18.} Participants view particular financial products as interchangeable with, or substitutes for, other products.

Distribution of products is not confined to traditional boundaries and institutional arrangements. For example, as noted in Part 3, investors may now purchase insurance products from a broad range of institutions including life companies, life agents and brokers, banks and superannuation funds. The management of risk through guaranteeing asset values is not limited to the products offered by insurance institutions. Options and futures contracts can also be used to achieve the same economic outcome as insurance contracts. The blurring of traditional boundaries is compounded by the rise of
conglomerates which combine activities which were traditionally offered by separate institutions.

4.5 FUNCTIONAL REGULATION

When considered against the underlying policy principles in Part 3, the operation of the current regulatory framework for financial markets is deficient. The inflexibility of the regulatory framework impedes innovation, inhibits competitive forces and encourages regulatory arbitrage. The current regulatory regime for financial instruments is fractured as instruments with the same economic function are regulated differently, depending on how they are classified.

Competition and financial innovation highlight the difficulties with the current regulatory regime and the need to provide a more flexible and efficient framework for financial markets. A more functional approach to the regulation of financial instruments will facilitate competition and financial innovation.

Regulation should treat economically equivalent financial instruments as elements of an integrated financial market. This position is supported by the FSI, which noted that the regulation of financial products should provide similar regulatory treatment for functionally equivalent products to promote a consistent regulatory regime.44 A more uniform approach to financial markets regulation is supported by CASAC as well as the FSI.45

A functional approach to regulating financial markets and products will achieve the following regulatory goals:

- consistent regulation across products or institutions in an environment where product and institutional boundaries are blurring;
- flexibility to accommodate differences which arise between different types of financial markets and products as well as future market developments; and
- certainty to market participants and clarity to investors.

4.6 **HARMONISATION OF FINANCIAL MARKETS REGULATION**

A more efficient and flexible regulatory regime for financial markets will be achieved by developing an integrated regulatory framework for financial instruments.

As discussed in the remaining parts of this paper, it is proposed that a large amount of the current regulation of financial markets and investment products be harmonised. For example, the proposed regime will provide uniform licensing of markets and intermediaries, and harmonise the conduct of business and market misconduct provisions.

In developing a new regulatory regime, the FSI recommendations for more uniform regulation of similar and substitute financial products will be taken into account.\(^{46}\) The FSI recommended the development of consistent and comparable disclosure requirements for retail financial products (deposit accounts, payments instruments, securities, collective investments, superannuation and insurance products)\(^ {47}\) and a single licensing regime for securities and futures intermediaries, foreign exchange dealers, insurance and life intermediaries.\(^ {48}\)

The proposals in this paper regarding consistent and comparable disclosure requirements and a uniform licensing regime set out the broad policy parameters which will apply to all financial intermediaries and financial products. Further consultation with industry and the regulator will take place regarding the development of consistent administration of the disclosure and licensing regimes, harmonised remedies and enforcement mechanisms and how the proposed regulatory framework will be implemented.

4.7 **REGULATORY APPROACH**

Individual financial instruments have varying characteristics notwithstanding the similar economic purpose of financial products. Differences include varying risk profiles and financial obligations. However, by focussing on the

development of key criteria in the legislation which are designed to satisfy regulatory objectives it is likely that the proposed regulatory regime will accommodate these differences at a practical level, without the need for a high level of detail to be prescribed by regulation.

Similarly, the application of the statutory obligations imposed on financial advisers and brokers will vary in some instances. For example, an intermediary may be required to provide monthly accounts for funds held on behalf of clients where client accounts can fluctuate rapidly, such as where money is held for the purposes of margining requirements for derivative transactions. Less frequent reporting will be required for some other products. An intermediary which does not hold funds on behalf of clients, eg in simple insurance arrangements, will not be subject to the obligation.

The regulatory framework will prescribe key standards which need to be satisfied by persons wishing to conduct a market in financial instruments or provide broking and advisory services. In effect a person would demonstrate to the regulator that their proposed arrangements for the conduct of the market or intermediary services will satisfy these standards.

Flexibility and Market Freedom

The advantage of the proposed regulatory framework is that elements of regulation which are required for market integrity can be met in various ways. Participants will have the capacity to adopt systems and procedures which can accommodate differences between certain transactions without unnecessary prescription or other regulatory constraints. The intention is to provide maximum market freedom so that participants can design systems which accommodate their particular operations, provided that the regulatory objectives of market integrity and investor protection are achieved.

For example, an often cited difference between securities and derivatives is the longer term obligations involved in derivative instruments. Clearing arrangements for derivatives need to deal with these longer term obligations, eg through margining requirements.49 However, rather than legislatively prescribing systems to deal with the specific risks of particular instruments, the legislation should require markets to have appropriate clearing arrangements to support their proposed activities. The nature of the proposed market activities will determine what types of clearing arrangements will be

---

49 Note that currently the Law does not require margining for futures contracts. Rather, the systems and procedures have been developed by market participants in order to strengthen market stability.
appropriate. Market operators may develop and adopt systems of their own choice provided they facilitate market integrity and investor protection objectives.

The proposed regulatory approach is similar to the arrangements which currently exist for the authorisation of exchanges under the Corporations Law.\textsuperscript{50} Exchanges may be authorised if their proposed business rules make satisfactory provision for particular matters, including admission and supervision of participants and dispute resolution. The Corporations Law does not prescribe the particular rules or manner in which exchanges must satisfy the key criteria.

It would not be feasible for the Law to prescribe what systems are appropriate in particular circumstances given the varying nature of market operations, the rapid pace of change in markets and the unwarranted interference with market freedom. The regulatory framework must not constrain market participants from responding to change by developing new systems and procedures provided that the objectives of regulation are achieved.

### Varying Application of Regulation

It is proposed to harmonise the current regulation of financial markets to the greatest extent possible. The broad and flexible criteria and statutory obligations proposed in this paper will vary in their application to some markets and intermediaries. It will be necessary to provide greater guidance about how some of the broad requirements will be satisfied in certain instances.

Where there is a clear need for different rules for particular markets or financial instruments, or a need to provide greater guidance on how to satisfy the legislative criteria, it is proposed that the detail may be set out in:

- regulations;
- conditions imposed on a financial intermediary’s licence;
- industry codes adopted by the regulator; and
- the rules developed by market operators which are not disallowed on market integrity grounds.

---

\textsuperscript{50} Corporations Law, sections 769 and 1126.
4.8. **DEFINITION OF FINANCIAL INSTRUMENTS**

A new definition of financial instruments will include all securities and derivatives, superannuation, life and general insurance, foreign exchange and deposit accounts and will be based on the UK Financial Services Act definition of ‘investments’. It will also draw upon the considerable work undertaken by CASAC in developing a definition of ‘derivatives’ to replace the unsatisfactory definition of ‘futures contract’. An indicative list of the types of instruments which will be included in the new concept of ‘financial instrument’ is at Appendix C.

A regulation making power to extend or restrict the definition is proposed in order to provide flexibility. This is required to accommodate financial innovation and to ensure the regime does not extend beyond its intended scope.

4.9 **ADVANTAGES OF HARMONISED REGULATION**

Benefits of a uniform regime for the regulation of financial instruments include:

- simplification of the regulatory framework for the trading of financial instruments by removing unnecessary legal distinctions;
- increased opportunities for competition and financial innovation without the need to seek dual regulatory authorisation (eg as a stock market and a futures market or dual licences to provide advice on securities and life products) and the removal of incentives for regulatory arbitrage; and
- creating flexibility that will accommodate inevitable change and permit market participants to respond in a timely manner to market developments.
Proposal No. 1 — Uniform Regulation of Financial Instruments

A more efficient and flexible regime for financial markets and investment products will be achieved by developing an integrated regulatory framework for financial instruments. The new regulatory regime will provide consistent regulation of functionally similar markets and products.

Financial instruments will include all securities, futures and other derivatives as well as foreign exchange, superannuation, general and life insurance and deposit accounts (see Appendix C).

The existing diverse regulatory arrangements for financial markets and investment products under the Corporations Law, the Insurance (Agents and Brokers) Act, the Insurance Contracts Act, the Banking Act, the Superannuation Industry (Supervision) Act and various industry codes will be harmonised.
The Long-Run Performance of Initial Public Offerings

JAY R. RITTER*

ABSTRACT

The underpricing of initial public offerings (IPOs) that has been widely documented appears to be a short-run phenomenon. Issuing firms during 1975–84 substantially underperformed a sample of matching firms from the closing price on the first day of public trading to their three-year anniversaries. There is substantial variation in the underperformance year-to-year and across industries, with companies that went public in high-volume years faring the worst. The patterns are consistent with an IPO market in which (1) investors are periodically overoptimistic about the earnings potential of young growth companies, and (2) firms take advantage of these “windows of opportunity.”

Numerous studies have documented two anomalies in the pricing of initial public offerings (IPOs) of common stock: (1) the (short-run) underpricing phenomenon, and (2) the “hot issue” market phenomenon. Measured from the offering price to the market price at the end of the first day of trading, IPOs produce an average initial return that has been estimated at 16.4%.1 Furthermore, the extent of this underpricing is highly cyclical, with some periods, lasting many months at a time, in which the average initial return is much higher.2 In this paper, I document a third anomaly: in the long-run, initial public offerings appear to be overpriced. Using a sample of 1,526 IPOs that went public in the U.S. in the 1975–84 period, I find that in the 3 years after going public these firms significantly underperformed a set of comparable firms matched by size and industry.

*Associate Professor of Finance, University of Illinois at Urbana-Champaign. I wish to thank Clifford Ball, Chris Barry, Stephen Buser (the editor), Donald Keim, Josef Lakonishok, Gita Rao, Nejat Seyhun, René Stulz, Michael Vetsuygens, Ivo Welch, Joseph Williams, an anonymous referee, participants in workshops at Illinois, Iowa, Vanderbilt, and Wharton, and especially Harry DeAngelo for helpful suggestions. An earlier version of this paper was presented at the December 1988 AFA meetings, the September 1989 Garn Institute conference on The Capital Acquisition Process, the November 1989 CRSP Seminar on the Analysis of Security Prices, and the October 1990 Q Group meetings. Navin Chopra, Tim Loiighran, and Tae Park provided extensive and very able research assistance. The research is partially supported by a grant from the Institute for Quantitative Research in Finance.

1A few of the many recent studies documenting positive initial returns include Carter and Manaster (1990), Miller and Reilly (1987), Ritter (1984, 1987), and Tinic (1988). Much of the recent work is discussed in Smith (1986) and Ibbotson, Sindelar, and Ritter (1988). The 16.4% average initial return figure is from Ibbotson, Sindelar, and Ritter, where the sample includes 8,668 IPOs going public in 1960–87.

2The “hot issue market” phenomenon is documented in Ibbotson and Jaffe (1975), Ritter (1984), and Ibbotson, Sindelar, and Ritter (1988).
There are several reasons why the long-run performance of initial public offerings is of interest. First, from an investor's viewpoint, the existence of price patterns may present opportunities for active trading strategies to produce superior returns. Second, a finding of nonzero aftermarket performance calls into question the informational efficiency of the IPO market. It provides evidence concerning Shiller's (1990) hypothesis that equity markets in general and the IPO market in particular are subject to fads that affect market prices. Third, the volume of IPOs displays large variations over time. If the high volume periods are associated with poor long-run performance, this would indicate that issuers are successfully timing new issues to take advantage of "windows of opportunity." Fourth, the cost of external equity capital for companies going public depends not only upon the transaction costs incurred in going public but also upon the returns that investors receive in the aftermarket. To the degree that low returns are earned in the aftermarket, the cost of external equity capital is lowered for these firms.

To summarize the empirical findings of this paper, the average holding period return for a sample of 1,526 IPOs of common stock in 1975–84 is 34.47% in the 3 years after going public, where this holding period return is measured from the closing market price on the first day of public trading to the market price on the 3 year anniversary. However, a control sample of 1,526 listed stocks, matched by industry and market value, produces an average total return of 61.86% over this same 3 year holding period. In other words, every dollar invested in a portfolio of IPOs purchased at the closing market price on the first day of trading results in a terminal wealth of $1.3447, while every dollar in the matching firms results in $1.6186, a ratio of only 0.831. In the long run, IPOs underperformed.

Possible explanations for this underperformance include (1) risk mismeasurement, (2) bad luck, or (3) fads and overoptimism. To ascertain whether risk mismeasurement could account for the poor long-run performance, alternative benchmark portfolios are used. To distinguish between the bad luck explanation and the fads and over optimism explanation, various cross-sectional and time-series patterns are documented. The pattern that emerges is that the underperformance is concentrated among relatively young growth companies, especially those going public in the high-volume years of the 1980s. While this pattern does not rule out bad luck being the cause of the underperformance, it is consistent with a scenario of firms going public when investors are irrationally over optimistic about the future potential of certain industries which, following Shiller (1990), I will refer to as the "fads" explanation. Further support for this interpretation is contained in Lee, Shleifer, and Thaler (1991) who find that the annual number of operating companies going public in the 1966–85 period is strongly negatively related

---

3Miller (1977) and Blazenko (1989) present models in which overoptimistic investors are the marginal investors in IPOs. Their models predict the long-run underperformance that is documented here.
The Long-Run Performance of Initial Public Offerings

The discount on closed-end mutual funds, which they interpret as a measure of individual investor sentiment.

At least three published academic studies, plus a series of articles in Forbes magazine, have examined the long-run performance of IPOs. Stoll and Curley (1970), focusing on 205 small offers, found that, “in the short run, the stocks in the sample showed remarkable price appreciation . . . . In the long run, investors in small firms did not fare so well . . . .” (pp. 314-315.)

Ibbotson (1975), using one offering per month for the 10-year period 1960–69, computed excess returns on IPOs with an offer price of at least $3.00 per share. He concludes that the “results generally confirm that there are no departures from market efficiency in the aftermarket.” (p. 265.) However, he does find evidence that there is “generally positive performance the first year, negative performance the next 3 years, and generally positive performance the last [fifth] year,” (p. 252), although the standard errors of his estimates are high enough to make it difficult to reject market efficiency. In his Table 12 he reports initial public offerings underperforming by an average of approximately 1% per month in the second through fourth years of public trading, with positive excess returns in the first and fifth years.4

Buser and Chan (1987) evaluate the two-year performance of over 1,078 NASDAQ/National Market System (NMS)-eligible initial public offerings in 1981–1985.5 Their sample has a positive average initial return of 6.2% and a mean 2-year market-adjusted return of 11.2% (exclusive of the initial return) where they use the NASDAQ Composite Index for their market adjustment.

The cover story of the December 2, 1985 Forbes magazine, entitled “Why New Issues Are Lousy Investments,” reports the results of a study that finds, for the period from January 1975 through June 1985, that initial public offerings have underperformed the market in the long run. Stern and Bornstein (1985) find as reported on p. 152 that “from its date of going public to last month, the average new issue was down 22% relative to the broad Standard & Poor’s 500 stock index.” Forbes analyzed 1,922 issues with an offering price of $1.00 or more. Unlike academic research, which typically uses event time, Forbes used calendar time, so their excess returns are computed over a period of anywhere from 10 years to a few months.

In summary, Stoll and Curley (1970), Ibbotson (1975), and Stern and Bornstein (1985) present evidence which suggests that at some point after going public the abnormal returns on initial public offerings may be negative. Ibbotson conducts the most satisfactory formal statistical tests, but his small sample size (120 issues) results in such large standard errors that he is unable to reject the hypothesis of market efficiency after a stock goes public. Only the Buser and Chan (1987) study does not find evidence of negative aftermarket performance after the initial return period.

---

4In Ibbotson’s (1975) Table 12 (p. 253), the months 43–48 alpha coefficient has a minus sign omitted, as can be seen from comparison with his Tables 11 and 13.
5The primary qualification for National Market System (NMS) eligibility among the NASDAQ stocks is income of $300,000 or more in the most recent fiscal year prior to going public.
The structure of this paper is as follows. Section I describes the data and methodology. Section II presents evidence regarding the aftermarket performance. Section III presents cross-sectional and time-series evidence on the aftermarket performance. Section IV concludes the paper with a summary and interpretation of the findings.

I. Data and Methodology

The sample is comprised of 1,526 initial public offerings in 1975–1984 meeting the following criteria: (1) an offer price of $1.00 per share or more, (2) gross proceeds, measured in terms of 1984 purchasing power, of $1,000,000 or more, (3) the offering involved common stock only (unit offers are excluded), (4) the company is listed on the CRSP daily Amex-NYSE or NASDAQ tapes within 6 months of the offer date, and (5) an investment banker took the company public. These firms represent 85.1% of the aggregate gross proceeds of all firms going public in 1975–84.

Table I presents the distribution of the sample by year, both in terms of the number of offers and the gross proceeds. Inspection of Table I shows that the number and value of IPOs were not evenly distributed over the 1975–84 sample period. Only 143 of the 1,526 sample offers occurred during the first half of the period. Fifty-seven percent ($12,060.4 million of the $21,066.9 million total) of the aggregate gross proceeds in the sample was raised in 1983 alone.

To evaluate the long-run performance of initial public offerings, two measures are used: (1) cumulative average adjusted returns (CAR) calculated with monthly portfolio rebalancing, where the adjusted returns are computed using several different benchmarks, and (2) 3-year buy and hold returns for both the IPOs and a set of matching firms. The matching firms are represented by American and New York stock exchange-listed securities that are to some extent matched by industry and market capitalization with each IPO.

6The gross proceeds numbers are computed based upon the actual number of shares sold, including overallotment options, if exercised. Most firm commitment offers include overallotment options in which the underwriter has the option of selling additional shares (limited to a maximum of 10% prior to August 1983 and 15% thereafter) at the offer price within 30 days of the offering.

7Of the 1,526 sample offers, 36 were initially traded on the American or New York stock exchanges, and the rest on NASDAQ. Of the 1,490 NASDAQ-listed issues, 128 changed to the American or New York stock exchanges within 3 years of going public. Of the 1,526 sample offers, 1,362 used a firm commitment contract, 157 used a best-efforts contract, and 7 used a combination firm commitment/best-efforts contract in going public. Only 13 of the 1,526 sample offers were not listed on the CRSP tapes within 1 month of the offer date. This sample of 1,526 firms is available from the author upon request.

8Details of the matching procedure are provided in the appendix. Because the industry composition of Amex-NYSE firms differs so dramatically from that of the IPOs, only 36% of the matching firms are in the same three-digit SIC code industry as their IPO (57% at the two-digit level). Also, the matching firms have, on average, larger market capitalizations. Sensitivity tests (not reported here) done using subsets that are more closely matched by size and industry show qualitatively similar results to those reported in this paper. See footnote 16.
### Table I

**Distribution of Initial Public Offerings by Year, 1975–84**

The number of total offers is based upon *Going Public: The IPO Reporter*'s listings, after excluding closed-end mutual funds and real estate investment trusts. Gross proceeds calculations are based upon the amount sold in the United States, including the proceeds from overallotment options, if exercised. No price level adjustments have been made in this table.

<table>
<thead>
<tr>
<th>Year</th>
<th>No. of IPOs</th>
<th>Aggregate gross proceeds, $ millions</th>
<th>No. of IPOs</th>
<th>Aggregate gross proceeds, $ millions</th>
<th>No. of IPOs</th>
<th>Aggregate gross proceeds</th>
<th>%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>14</td>
<td>264.0</td>
<td>12</td>
<td>262.4</td>
<td>85.7</td>
<td>99.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>33</td>
<td>237.3</td>
<td>28</td>
<td>213.9</td>
<td>84.8</td>
<td>90.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>32</td>
<td>150.6</td>
<td>19</td>
<td>132.3</td>
<td>59.4</td>
<td>87.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>48</td>
<td>247.3</td>
<td>31</td>
<td>218.4</td>
<td>64.6</td>
<td>88.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>78</td>
<td>429.0</td>
<td>53</td>
<td>347.1</td>
<td>67.9</td>
<td>80.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>234</td>
<td>1,408.3</td>
<td>129</td>
<td>1,097.9</td>
<td>55.1</td>
<td>78.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>438</td>
<td>3,200.3</td>
<td>300</td>
<td>2,689.5</td>
<td>68.5</td>
<td>84.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>199</td>
<td>1,335.0</td>
<td>93</td>
<td>1,104.2</td>
<td>46.7</td>
<td>82.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>865</td>
<td>13,247.8</td>
<td>589</td>
<td>12,060.4</td>
<td>68.1</td>
<td>91.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>535</td>
<td>4,237.1</td>
<td>272</td>
<td>2,940.8</td>
<td>50.8</td>
<td>69.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2,476</td>
<td>24,756.7</td>
<td>1,526</td>
<td>21,066.9</td>
<td>61.6</td>
<td>85.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Returns are calculated for two intervals: the initial return period (normally 1 day), defined as the offering date to the first closing price listed on the CRSP daily return tapes (both NASDAQ and Amex-NYSE), and the aftermarket period, defined as the 3 years after the IPO exclusive of the initial return period. The initial return period is defined to be month 0, and the aftermarket period includes the following 36 months where months are defined as successive 21-trading-day periods relative to the IPO date. Thus, month 1 consists of event days 2–22, month 2 consists of event days 23–43, etc. For IPOs in which the initial return period is greater than 1 day, the month 1 period is truncated accordingly, e.g., if the initial return period is 6 days, month 1 consists of event days 7–22. For IPOs that are delisted before their 3-year anniversary, the aftermarket period is truncated, and the 3-year buy and hold return ends with CRSP’s last listing. Firms which moved from NASDAQ to the American or New York stock exchanges during the 3 years after their offering date are not removed unless they are subsequently delisted from the Amex-NYSE tape. The CRSP NASDAQ daily returns file and the CRSP Amex-NYSE daily returns file are the source of the returns data.

Monthly benchmark-adjusted returns are calculated as the monthly raw return on a stock minus the monthly benchmark return for the corresponding 21-trading-day period. The benchmarks used are (1) the CRSP value-weighted NASDAQ index, (2) the CRSP value-weighted Amex-NYSE index, (3) listed firms matched by industry and size, and (4) an index of the smallest size
decile of the New York Stock Exchange. The benchmark-adjusted return for stock i in event month t is defined as

\[ ar_{it} = r_{it} - r_{mt} \]

The average benchmark-adjusted return on a portfolio of n stocks for event month t is the equally-weighted arithmetic average of the benchmark-adjusted returns:

\[ \text{AR}_t = \frac{1}{n} \sum_{i=1}^{n} ar_{it}. \]

The cumulative benchmark-adjusted aftermarket performance from event month q to event month s is the summation of the average benchmark-adjusted returns:

\[ \text{CAR}_{q,s} = \sum_{t=q}^{s} \text{AR}_t. \]

When a firm in portfolio p is delisted from the CRSP data, the portfolio return for the next month is an equally-weighted average of the remaining firms in the portfolio. The cumulative market-adjusted return for months 1 to 36, \( \text{CAR}_{1,36} \), thus involves monthly rebalancing, with the proceeds of a delisted firm equally allocated among the surviving members of the portfolio p in each subsequent month. For the month in which an IPO is delisted, the return for both the IPO and the benchmark includes just the days from the start of the month until the delisting.

As an alternative to the use of cumulative average benchmark-adjusted returns, which implicitly assumes monthly portfolio rebalancing, I also compute 3-year holding period returns, defined as

\[ R_i = \prod_{t=1}^{36} (1 + r_{it}) \]

where \( r_{it} \) is the raw return on firm i in event month t. This measures the total return from a buy and hold strategy where a stock is purchased at the first closing market price after going public and held until the earlier of (i) its 3-year anniversary, or (ii) its delisting.\(^9\) To interpret this 3-year total return, I compute wealth relatives as a performance measure, defined as

\[ WR = \frac{1 + \text{average 3-year total return on IPOs}}{1 + \text{average 3-year total return on matching firms}}. \]

A \textit{wealth relative} of greater than 1.00 can be interpreted as IPOs outperforming a portfolio of matching firms; a \textit{wealth relative} of less than 1.00 indicates that IPOs underperformed.

\(^9\)Unlike the Amex and NYSE in the 1975–1984 period, where mergers and takeovers are the predominant reasons for delisting, most NASDAQ delistings during this period are due to firms failing to meet the minimum capital requirements for continued listing. Because of early delistings, the average holding period is 34 months, rather than 36 months.
In this paper, I have calculated performance measures without explicitly adjusting for betas. While I do not report them here, the betas of the IPO firms display the same time-series patterns documented in Ibbotson (1975), Chan and Lakonishok (1990), and Clarkson and Thompson (1990), i.e., the average beta is greater than 1.00, and the average betas decline with the length of time since the IPO. This is true when betas are calculated using either the CRSP value-weighted NASDAQ or Amex-NYSE indices. The matching firms also have betas greater than 1.00. For post-issue months 1-12, 13-24, and 25-36, respectively, the average betas for IPOs are 1.39, 1.24, and 1.14 and the average betas for matching firms are 1.14, 1.13, and 1.04, using the CRSP value-weighted Amex-NYSE index. Although the IPO betas are greater than 1.00 on average, the difference in betas between the IPOs and matching firms is too small to have economically significant effects on the conclusions. To the degree that the IPO betas are higher than the betas of control portfolios, computing adjusted returns without explicitly adjusting for beta differences results in conservative estimates of IPO underperformance when the market risk premium is positive, as it is for this paper’s sample.

II. Aftermarket Performance

Table II reports the average matching firm-adjusted returns (AR) and cumulative average matching firm-adjusted returns (CAR) for the 36 months after the offering date for 1,526 IPOs in 1975-84. Thirty-one of the 36 monthly average adjusted returns are negative, with 13 of them having $t$-statistics lower than $-2.00$. The negative average adjusted returns are reflected in a steady decline in the cumulative average adjusted returns, which, after a slight increase in the first 2 months of seasoning, falls to $-29.13\%$ by the end of month 36, exclusive of the initial return, with an associated $t$-statistic of $-5.89$. The underperformance of the IPOs is both economically and statistically significant.

In Figure 1, I have plotted the matching firm-adjusted CAR, where the initial return is also included. Also plotted are four other cumulative average returns with different adjustments. The five series plotted, in order of their

---

10 For all of the beta calculations, I use Ibbotson’s (1975) RATS procedure. As Chan and Lakonishok (1990) document, I also find that the initial return betas are much higher when the market return is positive rather than negative. Rao (1989) notes that one reason for the decline in average betas with the time since the IPO is that riskier firms are more likely to be delisted and so are less likely to be included in the averages the longer the time since the IPO.

11 The average total return, exclusive of the initial return, during the 3 years after going public is 34.47\% for the IPOs in this sample, as reported in Table III of this paper. The average total return that an investor would have earned by rolling over T-bills for 3 years is approximately 28\%. Thus, in spite of most of these IPOs going public before substantial market rises, the IPO investors outperformed T-bills by only about 2\% per year. The betas of the IPOs would have to be implausibly low to reverse the conclusion that these IPOs underperformed in the 3 years after going public.
Table II
Abnormal Returns for Initial Public Offerings in 1975–84
Average matching firm-adjusted returns (ARt) and cumulative average returns (CARj t), in percent, with associated t-statistics for the 36 months after going public, excluding the initial return. The number of firms trading begins at less than 1,526 because some firms have a delay of more than one month after going public before being listed. ARt = \( \frac{1}{n_t} \sum_{i=1}^{n_t} (r_{ipo, it} - \overline{r}_{match, it}) \)
where \( r_{ipo, it} \) is the total return on initial public offering firm \( i \) in event month \( t \), and \( \overline{r}_{match, it} \) is the total return on the corresponding matching firm. The t-statistic for the average adjusted return is computed for each month as ARt \( \cdot \sqrt{n_t / \text{sd}_t} \), where ARt is the average matching firm-adjusted return for month \( t \), \( n_t \) is the number of observations in month \( t \), and \( \text{sd}_t \) is the cross-sectional standard deviation of the adjusted returns for month \( t \). The cross-sectional standard deviations vary from a low of 19.02 percent in month 10 to a high of 25.24 percent in month 16. The t-statistic for the cumulative average adjusted return in month \( t \), CARj t, is computed as CARj t \( \cdot \sqrt{n_t / \text{csd}_t} \), where \( n_t \) is the number of firms trading in each month, and \( \text{csd}_t \) is computed as \( \text{csd}_t = [t \cdot \text{var} + 2 \cdot (t - 1) \cdot \text{cov}]^{1/2} \), where \( t \) is the event month, \( \text{var} \) is the average (over 36 months) cross-sectional variance, and \( \text{cov} \) is the first-order autocovariance of the ARt series. Var has a value of 0.04453 (21.10 percent squared) and cov has a value of 0.02097, representing an autocorrelation coefficient of 0.471.

<table>
<thead>
<tr>
<th>Month of seasoning</th>
<th>Number of firms trading</th>
<th>ARt</th>
<th>t-stat</th>
<th>CARj t</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,512</td>
<td>0.38</td>
<td>0.63</td>
<td>0.38</td>
<td>0.70</td>
</tr>
<tr>
<td>2</td>
<td>1,514</td>
<td>1.49</td>
<td>2.81</td>
<td>1.88</td>
<td>2.02</td>
</tr>
<tr>
<td>3</td>
<td>1,517</td>
<td>-0.12</td>
<td>-0.24</td>
<td>1.75</td>
<td>1.46</td>
</tr>
<tr>
<td>4</td>
<td>1,518</td>
<td>-1.07</td>
<td>-2.91</td>
<td>0.69</td>
<td>0.49</td>
</tr>
<tr>
<td>5</td>
<td>1,519</td>
<td>-0.81</td>
<td>-1.63</td>
<td>-0.12</td>
<td>-0.08</td>
</tr>
<tr>
<td>6</td>
<td>1,519</td>
<td>-0.55</td>
<td>-1.06</td>
<td>-0.67</td>
<td>-0.38</td>
</tr>
<tr>
<td>7</td>
<td>1,518</td>
<td>-1.59</td>
<td>-3.13</td>
<td>-2.27</td>
<td>-1.18</td>
</tr>
<tr>
<td>8</td>
<td>1,516</td>
<td>-1.10</td>
<td>-2.21</td>
<td>-3.37</td>
<td>-1.63</td>
</tr>
<tr>
<td>9</td>
<td>1,514</td>
<td>-1.73</td>
<td>-3.38</td>
<td>-5.10</td>
<td>-2.31</td>
</tr>
<tr>
<td>10</td>
<td>1,513</td>
<td>-1.63</td>
<td>-3.32</td>
<td>-6.72</td>
<td>-2.88</td>
</tr>
<tr>
<td>11</td>
<td>1,508</td>
<td>-1.59</td>
<td>-3.08</td>
<td>-8.32</td>
<td>-3.39</td>
</tr>
<tr>
<td>12</td>
<td>1,501</td>
<td>-1.91</td>
<td>-3.66</td>
<td>-10.23</td>
<td>-3.97</td>
</tr>
<tr>
<td>13</td>
<td>1,496</td>
<td>-0.32</td>
<td>-0.56</td>
<td>-10.55</td>
<td>-3.92</td>
</tr>
<tr>
<td>14</td>
<td>1,492</td>
<td>-0.82</td>
<td>-1.60</td>
<td>-11.37</td>
<td>-4.06</td>
</tr>
<tr>
<td>15</td>
<td>1,486</td>
<td>-1.19</td>
<td>-2.30</td>
<td>-12.56</td>
<td>-4.32</td>
</tr>
<tr>
<td>16</td>
<td>1,478</td>
<td>-1.28</td>
<td>-1.92</td>
<td>-13.82</td>
<td>-4.59</td>
</tr>
<tr>
<td>17</td>
<td>1,469</td>
<td>-0.47</td>
<td>-0.85</td>
<td>-14.29</td>
<td>-4.58</td>
</tr>
<tr>
<td>18</td>
<td>1,463</td>
<td>-0.49</td>
<td>-0.88</td>
<td>-14.78</td>
<td>-4.59</td>
</tr>
<tr>
<td>19</td>
<td>1,449</td>
<td>0.37</td>
<td>0.61</td>
<td>-14.42</td>
<td>-4.43</td>
</tr>
<tr>
<td>20</td>
<td>1,440</td>
<td>0.30</td>
<td>0.55</td>
<td>-14.11</td>
<td>-4.12</td>
</tr>
<tr>
<td>21</td>
<td>1,429</td>
<td>-0.94</td>
<td>-1.66</td>
<td>-15.05</td>
<td>-4.27</td>
</tr>
<tr>
<td>22</td>
<td>1,416</td>
<td>-0.20</td>
<td>-0.33</td>
<td>-15.25</td>
<td>-4.21</td>
</tr>
<tr>
<td>23</td>
<td>1,403</td>
<td>-0.56</td>
<td>-0.92</td>
<td>-15.80</td>
<td>-4.24</td>
</tr>
<tr>
<td>24</td>
<td>1,397</td>
<td>-1.09</td>
<td>-1.97</td>
<td>-16.89</td>
<td>-4.43</td>
</tr>
<tr>
<td>25</td>
<td>1,388</td>
<td>0.30</td>
<td>0.50</td>
<td>-16.59</td>
<td>-4.25</td>
</tr>
<tr>
<td>26</td>
<td>1,372</td>
<td>-0.26</td>
<td>-0.44</td>
<td>-16.85</td>
<td>-4.20</td>
</tr>
<tr>
<td>27</td>
<td>1,354</td>
<td>-1.66</td>
<td>-2.87</td>
<td>-18.51</td>
<td>-4.50</td>
</tr>
<tr>
<td>28</td>
<td>1,347</td>
<td>-1.02</td>
<td>-1.72</td>
<td>-19.54</td>
<td>-4.65</td>
</tr>
<tr>
<td>29</td>
<td>1,339</td>
<td>-0.97</td>
<td>-1.84</td>
<td>-20.51</td>
<td>-4.78</td>
</tr>
<tr>
<td>30</td>
<td>1,324</td>
<td>-1.51</td>
<td>-2.74</td>
<td>-22.01</td>
<td>-5.01</td>
</tr>
<tr>
<td>31</td>
<td>1,309</td>
<td>-1.02</td>
<td>-1.57</td>
<td>-23.03</td>
<td>-5.13</td>
</tr>
<tr>
<td>32</td>
<td>1,296</td>
<td>-0.63</td>
<td>-1.00</td>
<td>-23.66</td>
<td>-5.16</td>
</tr>
<tr>
<td>33</td>
<td>1,283</td>
<td>-1.31</td>
<td>-2.16</td>
<td>-24.96</td>
<td>-5.33</td>
</tr>
<tr>
<td>34</td>
<td>1,270</td>
<td>-1.39</td>
<td>-2.39</td>
<td>-26.35</td>
<td>-5.52</td>
</tr>
<tr>
<td>35</td>
<td>1,260</td>
<td>-1.10</td>
<td>-1.89</td>
<td>-27.45</td>
<td>-5.64</td>
</tr>
<tr>
<td>36</td>
<td>1,254</td>
<td>-1.67</td>
<td>-2.80</td>
<td>-29.13</td>
<td>-5.89</td>
</tr>
</tbody>
</table>
The Long-Run Performance of Initial Public Offerings

Figure 1. Cumulative average adjusted returns for an equally-weighted portfolio of 1,526 initial public offerings in 1975-84, with monthly rebalancing. Five CAR series are plotted for the 36 months after the IPO date: 1) no adjustment (raw returns), 2) CRSP value-weighted NASDAQ index adjustment (NASDAQ-adjusted), 3) CRSP value-weighted Amex-NYSE index adjustment (VW-adjusted), 4) matching firm adjustment (matching firm-adjusted), and 5) lowest decile of NYSE market capitalization index adjustment (small firm-adjusted). Month 0 is the initial return interval.

Carrying first on the raw returns, a positive initial return of 14.32% is followed by monthly average raw returns varying between negative 1.20% and positive 2.96%. The cumulative average raw return peaks at 42.49% in month 34. This rise can be at least partly attributed to the bull market prevailing from August 1982 to August 1987, a period comprising the three post-issue years for the vast majority of sample firms.

Figure 1 also plots cumulative average market-adjusted returns, formed by subtracting the market return each month, for two different market indices: (1) the CRSP value-weighted NASDAQ index, and (2) the CRSP value-weighted index of Amex-NYSE stocks. These indices are nearly identical to the NASDAQ Composite and S&P 500 index returns, respectively. The
resulting CAR's display different patterns, which can be attributed to the different performance of the two indices, especially in 1984. During 1984, the total return on the CRSP value-weighted Amex-NYSE index was 5.02%, whereas the CRSP value-weighted NASDAQ index produced a total return of -8.96%.

The difference in the performance of the various indices sheds light on the discrepancy between Buser and Chan's (1987) findings of positive aftermarket performance and this study's findings of negative performance. Buser and Chan's use of the NASDAQ Composite index as their benchmark portfolio for a sample period in which this index substantially underperformed other indices accounts for part of the difference in findings. Furthermore, since much of the underperformance documented in Figure 1 occurs in the third post-issue year, their use of 2 years of aftermarket data, rather than the 3 years of this study, accounts for another part of the difference in findings. The rest of the difference in findings can largely be attributed to two differences in the sample selection criteria. Their restriction to NMS-qualifying issues removes many of the more speculative issues that this study includes, in which, as I will document in later tables, the poorest long-run performance occurs. Furthermore, a slight survivorship bias in their sample removes some of the issues that were subsequently delisted; these issues display a tendency to perform especially poorly, pulling down the average long-run performance.

In addition to the raw returns, market-adjusted returns, and matching firm-adjusted returns, Figure 1 also plots average small firm-adjusted returns, formed by subtracting a benchmark portfolio of the equally-weighted return on the smallest decile of NYSE stocks from the raw returns. Because many of the firms going public have low market capitalizations (measured in terms of 1984 purchasing power, the median gross proceeds are only $7.59 million, and the median post-issue market capitalization, valued at the closing market price on the first day of trading, is only $28.4 million), a small firm index may be appropriate to use as a benchmark portfolio. Using an equally-weighted index of small stock returns, as represented by the lowest decile of market capitalization stocks trading on the NYSE, the months 1-36 cumulative average small firm-adjusted return is -42.21%.

As Figure 1 shows, while all four adjustments display negative post-initial return performance, the quantitative measurement of the long-run performance of initial public offerings is very sensitive to the benchmark employed. This is not unusual in event studies using long windows, as indicated by Dimson and Marsh (1986). For evaluating the long run performance of IPOs, it is not at all clear what constitutes the appropriate benchmark portfolio. Since the vast majority of the IPOs trade on NASDAQ, a natural candidate would be the NASDAQ index. This index has the advantage that the industry mix more closely matches that of the sample IPOs than does the Amex or NYSE. However, the reason that the NASDAQ index's industry mix so closely matches is because in the mid-1980s so many of the firms in the index had recently gone public. Hence, using the NASDAQ index as a
The Long-Run Performance of Initial Public Offerings

...benchmark would tend to bias the results in favor of finding no abnormal market-adjusted returns.

To have a quantitative measure of long-run performance, some benchmark must be used. Throughout the rest of the paper, I will focus on wealth relatives, defined as the average gross total return on IPOs divided by the average gross total return on the matching firms, where both of these are measured over the 3 years after the IPO, excluding the initial return, as the primary measure of IPO aftermarket performance.\(^{12}\)

Table III reports the distribution of 3-year holding period returns for both the 1,526 IPOs and the matching firms. The median IPO 3-year return is \(-16.67\%\) contrasted with \(38.54\%\) for the median matching firm. The distribution of IPO 3-year holding period returns is more skewed than that of the matching firms, but the mean IPO 3-year holding period return is only \(34.47\%\) compared to a mean of \(61.86\%\) for the matching firms.

The highest 3-year total return of \(3964.43\%\), excluding the initial return of \(-3.23\%\), belongs to This Can't Be Yogurt, Inc. (now TCBY Enterprises), a March 28, 1984 IPO at $7.75 per share. After six 3 for 2 stock splits, its market price on NASDAQ was $27.50 on March 27, 1987, the equivalent of $313.25 on a pre-split basis. For 272 firms that were delisted before their 3-year anniversary, the mean 3-year holding period return, exclusive of their mean 17.02\% initial return, is \(-13.34\%\), with a wealth relative value of 0.581. For the 1,254 firms that were not delisted, the mean 3-year holding period return, exclusive of their mean initial return of 13.42\%, is 44.79\%, with a wealth relative value of 0.880. As one might expect, delisted firms have lower mean gross proceeds than nondelisted firms: $10.9 million versus $16.0 million when measured in terms of 1984 purchasing power.

III. Cross-Sectional and Time-Series Patterns in the Aftermarket Performance of IPOs

A. Aftermarket Performance Categorized by Issue Size and Initial Returns

To investigate possible explanations for the long-run underperformance of IPOs, this section documents various cross-sectional and time-series patterns.

In Table IV, firms are segmented by the gross proceeds of the offer. This permits examination of the generality of the negative aftermarket performance of IPOs. Inspection of Table IV discloses that there is a tendency for the smaller offers, which have the highest average matching firm-adjusted initial returns (henceforth, “adjusted initial returns”), to have the worst aftermarket performance. All gross proceeds categories display long-run underperformance.

In addition to reporting mean initial and aftermarket returns, Table IV also reports median initial and aftermarket returns. For the initial returns, \(^{12}\)For IPOs that are delisted prior to their 3-year anniversary, the total return is computed up to the delisting date.
Table III
Distribution of Three-Year Holding Period Returns, Exclusive of the Initial Return, for 1,526 Initial Public Offerings and Matching Firms in 1975-84

Three-year holding period returns are calculated as \( \prod_{t=1}^{756}(1 + r_{it}) - 1 \times 100\% \) where \( r_{it} \) is the daily return on stock \( i \), with the CRSP daily NASDAQ returns tape and the daily Amex-NYSE returns tape being the source of the daily returns. For initial public offerings that were delisted before the 3-year anniversary, the total return is calculated until the delisting date. If the initial return period lasted for more than 1 day, the total return is calculated from the first CRSP-reported closing price until the 756th trading day after the IPO. The corresponding matching firm's total return is calculated over the same truncated return interval. If the matching firm is delisted early, a second (and possibly third) matching firm's return is spliced onto the first matching firm. For firms with no dividends and no stock splits the total return corresponds to \( \frac{P_3}{P_1} - 1 \times 100\% \) where \( P_3 \) is the price on the 3-year anniversary, and \( P_1 \) is the first closing market price after the IPO.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Initial public offerings</th>
<th>Matching firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (lowest)</td>
<td>-99.02</td>
<td>-94.59</td>
</tr>
<tr>
<td>77</td>
<td>-92.27</td>
<td>-61.11</td>
</tr>
<tr>
<td>153</td>
<td>-85.80</td>
<td>-44.53</td>
</tr>
<tr>
<td>229</td>
<td>-79.16</td>
<td>-31.58</td>
</tr>
<tr>
<td>306</td>
<td>-72.98</td>
<td>-19.05</td>
</tr>
<tr>
<td>382 (25th percentile)</td>
<td>-66.25</td>
<td>-9.49</td>
</tr>
<tr>
<td>458</td>
<td>-57.98</td>
<td>0.00</td>
</tr>
<tr>
<td>535</td>
<td>-48.06</td>
<td>8.79</td>
</tr>
<tr>
<td>611</td>
<td>-38.28</td>
<td>18.75</td>
</tr>
<tr>
<td>687</td>
<td>-28.20</td>
<td>27.67</td>
</tr>
<tr>
<td>764 (median)</td>
<td>-16.67</td>
<td>38.54</td>
</tr>
<tr>
<td>840</td>
<td>-2.54</td>
<td>51.14</td>
</tr>
<tr>
<td>916</td>
<td>13.25</td>
<td>63.14</td>
</tr>
<tr>
<td>992</td>
<td>29.07</td>
<td>75.82</td>
</tr>
<tr>
<td>1,069</td>
<td>46.85</td>
<td>87.56</td>
</tr>
<tr>
<td>1,145 (75th percentile)</td>
<td>69.59</td>
<td>103.16</td>
</tr>
<tr>
<td>1,222</td>
<td>99.96</td>
<td>120.42</td>
</tr>
<tr>
<td>1,298</td>
<td>138.03</td>
<td>148.86</td>
</tr>
<tr>
<td>1,374</td>
<td>205.33</td>
<td>187.01</td>
</tr>
<tr>
<td>1,450</td>
<td>320.53</td>
<td>240.78</td>
</tr>
<tr>
<td>1,526 (highest)</td>
<td>3,964.43</td>
<td>1,268.56</td>
</tr>
<tr>
<td>Mean</td>
<td>34.47</td>
<td>61.86</td>
</tr>
</tbody>
</table>

the median is a positive 4.61%, with only 368 of the 1,526 offers (22.2%) having negative adjusted initial returns.

DeBondt and Thaler (1985, 1987) have presented evidence that, at least for low-capitalization stocks, there is a negative relation between past and subsequent abnormal returns on individual securities using holding periods of a year or more, which they interpret as evidence of market overreaction. Table V tests whether the IPO market is subject to overreaction by comput-


Table IV

Mean Performance Measures for 1,526 IPOs in 1975–84
Categorized by Gross Proceeds

Gross proceeds are measured in dollars of 1984 purchasing power using the U.S. GNP deflator. Initial returns are computed as \( r_{\text{ipo}} - r_{\text{matching firm}} \) over the initial return interval (one day for 1,203 of the 1,526 firms). The three-year holding period return is calculated excluding the initial return. For IPOs that are delisted prior to their three-year anniversary the matching firms' return is ended on the same date as the IPO. Total returns include both capital gains and dividends. The wealth relative is the ratio of one plus the mean IPO 3-year holding period return (not in percent) divided by one plus the mean matching firm 3-year holding period return (not in percent). For the smallest gross proceeds category, 1.1794/1.6754 = 0.704.

<table>
<thead>
<tr>
<th>Gross proceeds, $</th>
<th>Average adjusted initial return %</th>
<th>Excluding initial returns</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IPOs %</td>
<td>Matching firms %</td>
<td>Month 0</td>
</tr>
<tr>
<td>1,000,000– 2,999,999</td>
<td>17.94</td>
<td>67.54</td>
<td>221</td>
</tr>
<tr>
<td>3,000,000– 4,999,999</td>
<td>20.89</td>
<td>58.72</td>
<td>296</td>
</tr>
<tr>
<td>5,000,000– 9,999,999</td>
<td>40.06</td>
<td>69.87</td>
<td>379</td>
</tr>
<tr>
<td>10,000,000– 14,999,999</td>
<td>46.25</td>
<td>55.99</td>
<td>211</td>
</tr>
<tr>
<td>15,000,000– 24,999,999</td>
<td>43.97</td>
<td>50.56</td>
<td>200</td>
</tr>
<tr>
<td>25,000,000–353,950,260</td>
<td>39.81</td>
<td>62.50</td>
<td>219</td>
</tr>
<tr>
<td>All (mean)</td>
<td>14.06</td>
<td>34.47</td>
<td>1,526</td>
</tr>
<tr>
<td>All (median)</td>
<td>4.61</td>
<td>−16.67</td>
<td>1,526</td>
</tr>
</tbody>
</table>

In some respects, the finding that there is a tendency for the offerings with the highest initial returns to do worst in the long run may be a manifestation of a desire by issuers to avoid future lawsuits (see Ibbotson (1975, p. 264) and Tinic (1988)), by not fully exploiting the market’s overoptimism at the time of the offering. This may also shed light on the “partial adjustment” phenomenon which refers to a positive correlation between initial returns and changes in the offering price between the preliminary and final prospectuses (see Ibbotson, Sindelar, and Ritter (1988), Sternberg (1989), and Weiss (1990) for discussions).

\(^{13}\)Carter and Dark (1990) examine the correlation between initial returns and 18-month aftermarket returns for a sample of 911 firm commitment offers that went public between January 1, 1979 and November 11, 1984. They find that the abnormal 18-month aftermarket returns for firms having the highest initial returns tend to be slightly lower than for firms having the lowest initial returns which they interpret as evidence of valuation errors.
Table V

Aftermarket Performance Categorized by Initial Return Quintiles, with Results for Small and Large Offerings, for 1,526 IPOs in 1975-84

Gross proceeds are measured in dollars of 1984 purchasing power. $7.59 million is the median gross proceeds for the 1,526 offerings. The wealth relative is the ratio of one plus the mean IPO 3-year holding period return (not in percent) divided by one plus the mean matching firm 3-year holding period return (not in percent), exclusive of the initial return.

<table>
<thead>
<tr>
<th>Matching firm-adjusted initial return quintile</th>
<th>IPO</th>
<th>Matching firm</th>
<th>Wealth relative</th>
<th>Segment by gross proceeds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td></td>
<td>Proceeds &lt; $7.59 million</td>
</tr>
<tr>
<td>23.70 &lt; IR &lt; 37.39</td>
<td>9.45</td>
<td>61.39</td>
<td>0.678</td>
<td>0.606</td>
</tr>
<tr>
<td>8.10 &lt; IR &lt; 23.70</td>
<td>27.94</td>
<td>65.52</td>
<td>0.773</td>
<td>0.702</td>
</tr>
<tr>
<td>2.37 &lt; IR &lt; 8.10</td>
<td>41.56</td>
<td>55.82</td>
<td>0.908</td>
<td>0.800</td>
</tr>
<tr>
<td>-0.84 &lt; IR &lt; 2.37</td>
<td>45.51</td>
<td>60.88</td>
<td>0.904</td>
<td>0.794</td>
</tr>
<tr>
<td>-92.38 &lt; IR &lt; -0.84</td>
<td>47.95</td>
<td>65.70</td>
<td>0.893</td>
<td>0.829</td>
</tr>
</tbody>
</table>

B. Aftermarket Performance by Industry

Tables VI and VII segment firms by industry classifications based upon three-digit Standard Industrial Classification (SIC) codes. Where two or more SIC codes represent industries that are very similar, I have grouped them into a single industry. The 13 industries for which there were at least 25 IPOs in my sample are listed with the remaining 420 other firms grouped together. As inspection of Table VI demonstrates, companies going public in 1975-84 were not evenly distributed over all industries. Oil and gas firms are heavily represented (with most of these offers conducted in 1980 and 1981) as are financial institutions. Following the deregulation of the airline industry in 1978, several dozen young airlines went public. High technology firms in the computer and biomedical fields also have high representation. On the other hand, very few auto and steel companies went public in 1975-84. The industry representation represents capital flowing into growing industries in a dynamic economy.

14The SIC codes are compiled from the CRSP NASDAQ database, the January 1987 NASDAQ Company Directory, and other sources. Where there are discrepancies between various sources (due, for example, to a company's having changed the nature of its business after going public), I have assigned an SIC code based upon Going Public: The IPO Reporter's description of its business at the time of the offer.

15In the 1975-84 period, only 15 of the 1,526 companies going public represented "reverse LBOs," defined as a company going public that had been involved in a leveraged buyout. Among companies going public in 1986 and later, reverse LBOs have been more common. The industry representation of reverse LBOs is concentrated in mature industries, such as retailing and food companies. See Muscarella and Vetsuypens (1989a) for an analysis of these reverse LBOs.
Table VI

Mean and Median Sales, Gross Proceeds, and Age of 1,526 Sample Offers Categorized by Industry

Both sales and gross proceeds are expressed in terms of dollars of 1984 purchasing power. Sales are measured as 12-month revenues for the most recent 12-month period prior to going public. Gross proceeds are measured including, for firm commitment offerings, the proceeds from overallotment options, if exercised. The age of the issuing firm is measured as the calendar year of going public minus the calendar year of founding. The year of founding is the same or earlier than the year of incorporation or reincorporation. The 39 firms with a founding date prior to 1901 have their age computed as the offer year minus 1901.

<table>
<thead>
<tr>
<th>Industry</th>
<th>SIC codes</th>
<th>Number of offers</th>
<th>Annual sales, $ millions</th>
<th>Gross proceeds, $ millions</th>
<th>Age of issuing firm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>Communications and electronic equip.</td>
<td>366, 367</td>
<td>138</td>
<td>14.16</td>
<td>8.26</td>
<td>11.25</td>
</tr>
<tr>
<td>Oil and gas</td>
<td>131, 138</td>
<td>127</td>
<td>19.05</td>
<td>0.53</td>
<td>9.57</td>
</tr>
<tr>
<td></td>
<td>291, 679</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial institutions</td>
<td>602, 603</td>
<td>125</td>
<td>120.20</td>
<td>49.43</td>
<td>27.41</td>
</tr>
<tr>
<td>(banks and S&amp;L's)</td>
<td>612, 671</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer and data processing services</td>
<td>737</td>
<td>113</td>
<td>16.40</td>
<td>11.50</td>
<td>13.98</td>
</tr>
<tr>
<td>Optical, medical, and scientific instruments</td>
<td>381-384</td>
<td>111</td>
<td>10.89</td>
<td>2.23</td>
<td>9.29</td>
</tr>
<tr>
<td>Retailers</td>
<td>520-573</td>
<td>70</td>
<td>74.70</td>
<td>34.74</td>
<td>17.28</td>
</tr>
<tr>
<td></td>
<td>591-599</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wholesalers</td>
<td>501-519</td>
<td>63</td>
<td>56.58</td>
<td>13.41</td>
<td>12.32</td>
</tr>
<tr>
<td>Restaurant chains</td>
<td>581</td>
<td>54</td>
<td>34.54</td>
<td>10.98</td>
<td>10.34</td>
</tr>
<tr>
<td>Health care and HMOs</td>
<td>805-809</td>
<td>50</td>
<td>35.20</td>
<td>7.42</td>
<td>14.96</td>
</tr>
<tr>
<td>Drugs and genetic engineering</td>
<td>283</td>
<td>44</td>
<td>21.14</td>
<td>1.98</td>
<td>19.70</td>
</tr>
<tr>
<td>miscellaneous business services</td>
<td>739</td>
<td>42</td>
<td>14.27</td>
<td>2.38</td>
<td>7.75</td>
</tr>
<tr>
<td>Airlines</td>
<td>451</td>
<td>25</td>
<td>20.65</td>
<td>14.33</td>
<td>11.68</td>
</tr>
<tr>
<td>All other firms</td>
<td>-</td>
<td>420</td>
<td>61.24</td>
<td>18.07</td>
<td>15.04</td>
</tr>
<tr>
<td>All firms</td>
<td>-</td>
<td>1,526</td>
<td>42.82</td>
<td>11.55</td>
<td>15.06</td>
</tr>
</tbody>
</table>

Also reported in Table VI are the mean and median gross proceeds and annual sales, expressed in terms of 1984 purchasing power, and the mean and median age of the issuing firm, categorized by industry. As can be seen, there are substantial industry differences. The overall median age at the time of issue is only 6 years. For oil and gas firms, however, the median age is only 2 years, while for financial institutions the median age is 49 years. Most of the financial institution IPOs involve mutual savings banks and mutual savings and loan associations converting to stock companies after a 1982 regulatory change. Masulis (1987) analyzes this process. Also noteworthy is the fact that the median oil and gas IPO raised 12 times its annual revenue when it went public.

Table VII reports the long-run performance measures for IPOs, categorized
### Table VII
#### Mean Performance Categorized by Industry

The *wealth relative* is the ratio of one plus the mean IPO 3-year holding period return (not in percent) divided by one plus the mean matching firm 3-year holding period return (not in percent).

<table>
<thead>
<tr>
<th>Industry</th>
<th>Average matching firm-adjusted initial return %</th>
<th>Excluding initial returns</th>
<th>Average 3-year holding period total return %</th>
<th>Wealth relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computers</td>
<td>13.67</td>
<td>19.22</td>
<td>47.84</td>
<td>0.806</td>
</tr>
<tr>
<td>Electronic equipment</td>
<td>14.59</td>
<td>29.93</td>
<td>61.46</td>
<td>0.805</td>
</tr>
<tr>
<td>Oil and gas</td>
<td>30.92</td>
<td>-43.86</td>
<td>34.67</td>
<td>0.417</td>
</tr>
<tr>
<td>Financial institutions</td>
<td>3.69</td>
<td>128.21</td>
<td>59.23</td>
<td>1.433</td>
</tr>
<tr>
<td>Computer services</td>
<td>16.07</td>
<td>13.13</td>
<td>50.38</td>
<td>0.752</td>
</tr>
<tr>
<td>Scientific instruments</td>
<td>20.96</td>
<td>18.14</td>
<td>72.20</td>
<td>0.686</td>
</tr>
<tr>
<td>Retailers</td>
<td>7.60</td>
<td>54.05</td>
<td>113.63</td>
<td>0.721</td>
</tr>
<tr>
<td>Wholesalers</td>
<td>16.95</td>
<td>1.42</td>
<td>47.14</td>
<td>0.689</td>
</tr>
<tr>
<td>Restaurant chains</td>
<td>13.51</td>
<td>73.86</td>
<td>82.36</td>
<td>0.953</td>
</tr>
<tr>
<td>Health care</td>
<td>14.12</td>
<td>36.93</td>
<td>53.25</td>
<td>0.894</td>
</tr>
<tr>
<td>Drugs</td>
<td>14.63</td>
<td>121.69</td>
<td>91.96</td>
<td>1.155</td>
</tr>
<tr>
<td>Miscellaneous services</td>
<td>10.20</td>
<td>26.61</td>
<td>80.50</td>
<td>0.701</td>
</tr>
<tr>
<td>Airlines</td>
<td>6.26</td>
<td>61.62</td>
<td>42.93</td>
<td>1.131</td>
</tr>
<tr>
<td>All other firms</td>
<td>11.13</td>
<td>33.40</td>
<td>64.24</td>
<td>0.812</td>
</tr>
<tr>
<td>All firms</td>
<td>14.06</td>
<td>34.47</td>
<td>61.86</td>
<td>0.831</td>
</tr>
</tbody>
</table>

By industry. As can be seen, the long-run performance of IPOs in different industries varies widely. Financial institutions (almost all of which went public in 1983 and 1984) had the best long-run performance, benefiting from the large drop in interest rates in 1985–86. Oil and gas firms (most of which went public in 1980 and 1981) substantially underperformed the market. As is well-known, oil prices declined substantially during 1981–83, so the underperformance of oil and gas firms does not come as a surprise. However, the long-run underperformance of IPOs is present in all but three of the 14 industry groupings. The underperformance of the IPOs in so many industries relative to other firms in the same industries may be interpreted as evidence that is more consistent with a “fads” explanation than mere bad luck.

---

16 Since only 57% of the matching firms are in the same two-digit industry as the IPOs, it is possible that the imperfect control for industry factors can account for the long-run underperformance. In tests not reported here, I restricted the long-run analysis to the IPOs for which I had a matching firm in the same two-digit industry (or the same industry as defined in Table VI). The wealth relative value for this subsample is 0.866 as contrasted with 0.831 for the entire sample. This indicates that IPOs tend to underperform relative to their industries, which in turn tend to underperform relative to the market in the 3 years after going public.
Table VIII
Performance Categorized by Year of Issuance for Initial Public Offerings in 1975–84

The average real gross proceeds, measured in dollars of 1984 purchasing power, is computed as the product of the U.S. GNP Deflator index and the average nominal gross proceeds. The wealth relative is the ratio of one plus the mean IPO 3-year holding period return (not in percent) divided by one plus the mean matching firm 3-year holding period return (not in percent), exclusive of the initial return.

<table>
<thead>
<tr>
<th>Year</th>
<th>GNP deflator</th>
<th>Nominal</th>
<th>Real</th>
<th>Number of issues</th>
<th>Average gross proceeds, $ millions</th>
<th>Average matching firm-adjusted initial return (%)</th>
<th>Excluding initial returns</th>
<th>Matching firms</th>
<th>Wealth relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>1.76</td>
<td>21.87</td>
<td>38.49</td>
<td>12</td>
<td>38.49</td>
<td>-5.24</td>
<td>59.44</td>
<td>52.51</td>
<td>1.045</td>
</tr>
<tr>
<td>1976</td>
<td>1.67</td>
<td>7.64</td>
<td>12.76</td>
<td>28</td>
<td>6.38</td>
<td>122.58</td>
<td>124.11</td>
<td>0.993</td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>1.58</td>
<td>6.96</td>
<td>11.00</td>
<td>19</td>
<td>8.21</td>
<td>188.35</td>
<td>54.72</td>
<td>1.864</td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>1.47</td>
<td>7.05</td>
<td>10.36</td>
<td>31</td>
<td>31.78</td>
<td>134.60</td>
<td>97.37</td>
<td>1.189</td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>1.38</td>
<td>6.55</td>
<td>9.04</td>
<td>63</td>
<td>22.06</td>
<td>75.98</td>
<td>71.76</td>
<td>1.025</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>1.26</td>
<td>8.51</td>
<td>10.72</td>
<td>129</td>
<td>38.27</td>
<td>46.28</td>
<td>68.56</td>
<td>0.868</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>1.15</td>
<td>8.96</td>
<td>10.30</td>
<td>300</td>
<td>9.98</td>
<td>5.26</td>
<td>60.85</td>
<td>0.654</td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>1.08</td>
<td>11.87</td>
<td>12.82</td>
<td>93</td>
<td>15.18</td>
<td>26.07</td>
<td>119.92</td>
<td>0.573</td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>1.04</td>
<td>20.48</td>
<td>21.30</td>
<td>589</td>
<td>12.56</td>
<td>21.31</td>
<td>52.88</td>
<td>0.793</td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>1.00</td>
<td>10.81</td>
<td>10.81</td>
<td>272</td>
<td>8.40</td>
<td>52.03</td>
<td>47.91</td>
<td>1.028*</td>
<td></td>
</tr>
</tbody>
</table>

All: 13.81 15.06 1,526 14.06 34.47 61.86 0.831

*aIf one outlier (TCBY, Inc.) is removed, the average 3-year raw return falls to 37.59% and the 3-year wealth relative falls to 0.930.

C. Aftermarket Performance by Year of Issuance

In Table VIII, firms are categorized by their year of issuance. The results in Table VIII show that the long-run underperformance is not as general a phenomenon as the short-run underpricing that has been widely documented. The wealth relatives are less than one for only five of the ten sample years. Because the volume of new issues was much heavier in the early 1980s than in the late 1970s, however, the mean wealth relative is only 0.831 when all issues are weighted equally.

The negative relation between annual volume and aftermarket performance that is evident in Table VIII is consistent with the following scenario: firms choose to go public when investors are willing to pay high multiples (price-earnings or market-to-book) reflecting optimistic assessments of the net present value of growth opportunities. The negative aftermarket performance that then typically results is due to disappointing realizations of the subsequent net cash flows. This is due to either (1) bad luck or (2) irra-
Table IX
Aftermarket Performance Categorized by the Age of the Issuing Firm

Panel A includes all 1,526 IPOs. Panel B includes the 1,274 IPOs remaining after excluding the two industries with the most extreme wealth relatives: oil and gas (primarily very young firms) which did poorly, and financial institutions (primarily very old firms) which did well. Oil and gas firms are defined as firms with SIC codes of 131, 138, 291, and 679, representing oil and gas exploration, production, servicing, refining, and holding companies. Financial institutions are defined as firms with SIC codes of 602, 603, 612, and 671, representing commercial banks, savings banks, savings and loans, and bank holding companies. The wealth relative is the ratio of one plus the mean IPO 3-year holding period return (not in percent) divided by one plus the mean matching firm 3-year holding period return (not in percent).

<table>
<thead>
<tr>
<th>Age in years</th>
<th>Sample size</th>
<th>Average matching firm-adjusted initial return %</th>
<th>Average 3-year holding period total return</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Panel A: All 1,526 firms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample size</td>
<td>IPOs %</td>
</tr>
<tr>
<td>0-1</td>
<td>252</td>
<td>29.42</td>
<td>5.34</td>
</tr>
<tr>
<td>2-4</td>
<td>381</td>
<td>14.51</td>
<td>15.69</td>
</tr>
<tr>
<td>5-9</td>
<td>328</td>
<td>13.15</td>
<td>28.47</td>
</tr>
<tr>
<td>10-19</td>
<td>312</td>
<td>9.05</td>
<td>40.74</td>
</tr>
<tr>
<td>20-up</td>
<td>253</td>
<td>5.42</td>
<td>91.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Panel B: Excluding oil and gas firms and financial institutions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample size</td>
<td>IPOs %</td>
</tr>
<tr>
<td>0-1</td>
<td>177</td>
<td>23.87</td>
<td>16.19</td>
</tr>
<tr>
<td>2-4</td>
<td>338</td>
<td>14.87</td>
<td>19.22</td>
</tr>
<tr>
<td>5-9</td>
<td>305</td>
<td>13.71</td>
<td>33.01</td>
</tr>
<tr>
<td>10-19</td>
<td>300</td>
<td>9.32</td>
<td>42.97</td>
</tr>
<tr>
<td>20-up</td>
<td>154</td>
<td>5.41</td>
<td>63.76</td>
</tr>
</tbody>
</table>

D. Aftermarket Performance by Age

In Table IX, firms are segmented on the basis of their age at the time of going public, computed as the year of the offer minus the year of founding. There is a strong monotone relation between age and aftermarket performance. For the initial return, there is a strong monotone pattern in the other direction, consistent with the notions that risky issues require higher average initial returns and that age is a proxy for this risk. The initial return and aftermarket performance patterns are much clearer in Table IX, using age as a measure of both ex ante uncertainty and investor optimism, than are

Muscarella and Vetsuypens (1989b) also document a negative relation between initial returns and age.
Ordinary Least Squares Regression Results with the Three-year Total Return as the Dependent Variable,
for 1,526 IPOs in 1975-84

Returnₙ = b₀ + b₁IRₙ + b₂Log(1 + ageₙ) + b₃ Marketₙ + b₄Volₙ + b₅Oilₙ + b₆Bankₙ + εₙ.

Returnₙ is the raw three-year return, measured from the first aftermarket closing price to the earlier of the three-year anniversary or its CRSP delisting date. IRₙ is the market-adjusted initial return, calculated using the CRSP value-weighted index of Amex-NYSE stocks as the market index. Log(1 + ageₙ) is the natural logarithm of one plus the difference between the year of going public and the year of founding, with firms founded before 1901 assumed to be founded in 1901. Marketₙ is the CRSP value-weighted market return for the same return interval as the dependent variable. Volₙ is the annual volume of IPOs in the year of issuance, divided by 100. The gross number of IPOs, given in Table I, is used. Oilₙ is a 0,1 dummy variable taking on the value of 1 if the issuing firm has an SIC code of 131, 138, 291, or 679, representing oil and gas production, exploration, refining, and service companies, or oil and gas holding companies. Bankₙ is a 0,1 dummy variable taking on the value of 1 if the issuing firm has an SIC code of 602, 603, 612, or 671, representing banks, savings and loans, and associated holding companies. Bootstrapped p-values are in parentheses.

### Panel A: Parameter estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intercept</th>
<th>IR</th>
<th>Log(1 + age)</th>
<th>Market</th>
<th>Vol</th>
<th>Oil</th>
<th>Bank</th>
<th>R² adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>0.238</td>
<td>-0.206</td>
<td>0.127</td>
<td>0.841</td>
<td>-0.109</td>
<td>-0.765</td>
<td>0.825</td>
<td>0.070</td>
</tr>
<tr>
<td>(0.186)</td>
<td>(0.143)</td>
<td></td>
<td>(0.010)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: Summary statistics of variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return</td>
<td>0.345</td>
<td>-0.167</td>
<td>1.902</td>
</tr>
<tr>
<td>IR</td>
<td>0.141</td>
<td>0.040</td>
<td>0.309</td>
</tr>
<tr>
<td>Log(1 + age)</td>
<td>2.009</td>
<td>1.946</td>
<td>1.079</td>
</tr>
<tr>
<td>Market</td>
<td>0.566</td>
<td>0.580</td>
<td>0.246</td>
</tr>
<tr>
<td>Vol</td>
<td>5.520</td>
<td>5.350</td>
<td>2.831</td>
</tr>
<tr>
<td>Oil</td>
<td>0.083</td>
<td>0.000</td>
<td>0.276</td>
</tr>
<tr>
<td>Bank</td>
<td>0.082</td>
<td>0.000</td>
<td>0.274</td>
</tr>
</tbody>
</table>

the patterns in Table IV, in which firms are segmented by gross proceeds. The patterns using gross proceeds are not as clear because of two confounding effects—larger issues are typically issued by more established firms, but a given firm will choose to float a larger issue when the market conditions are characterized by strong demand.

A potential problem with interpreting Panel A of Table IX is that many of the oldest firms are financial institutions, which had exceptionally good aftermarket performance during this period, and many of the youngest firms are oil and gas firms, which had exceptionally poor aftermarket performance, as documented in Table VII. Thus, the pattern of aftermarket performance documented in Table IX is strengthened by these industry effects. To control for these effects, I report in Panel B of Table IX the initial and aftermarket performance for firms categorized by their age at the time of issue, with the 125 financial institutions and the 127 oil and gas firms deleted. The patterns...
present in Panel A are still present in Panel B, demonstrating that the lack of underperformance by established companies is not merely a manifestation of strong aftermarket performance by financial institutions.\textsuperscript{18} I interpret the poor long-run performance of the younger IPOs, which typically have higher market-to-book ratios than more established firms, as evidence consistent with the overoptimism and fads story.

\textbf{E. Regression Results}

The cross-sectional patterns documented in Tables IV through IX are not independent of each other. Among other correlations, the worst-performing industry in the long run (oil and gas) has the lowest median age and the highest average initial return, while the best-performing industry in the long run (financial institutions) has the highest median age and the lowest average initial return. To disentangle the effects, Table X reports the results of a multiple regression using the raw 3-year total return on IPOs as the dependent variable. The explanatory variables are the market-adjusted initial return, the 3-year total return on the market, the logarithm of one plus age, the volume of IPOs in the year of issuance, and dummy variables for the oil and financial institutions industries.\textsuperscript{19}

The Table X results generally support the conclusions from earlier tables. The adjusted coefficient of determination is rather low at only 7%. Because the dependent variable (3-year total returns) is so skewed, the residuals are also highly nonnormal. Consequently, bootstrapped \( p \)-values are reported.\textsuperscript{20} With the exception of the initial return, all of the coefficient estimates are statistically significant at conventional levels. The parameter estimates are also economically significant. The coefficient on annual IPO volume (divided by 100) of \(-0.109\), for instance, indicates that the difference in 3-year total returns for a firm going public in a low volume year such as 1976 (33 offerings) rather than in a high volume year such as 1983 (865 offerings) is 0.907 (90.7%), ceteris paribus. The coefficient on the market return of 0.841

\textsuperscript{18}In results not reported here I have prepared tables analogous to Panels A and B of Table IX using sales and market-to-book ratios rather than age for categorizing firms with similar results.

\textsuperscript{19}Several additional variables were also included in other regressions (unreported) that were run, with no boost in the adjusted coefficient of determination. Among these other insignificant variables are the logarithm of sales and a dummy variable accounting for the use of a best-efforts contract.

\textsuperscript{20}The approximate randomization bootstrapping procedure described in Noreen (1989) creates a coefficient vector under the null hypothesis of no relation by randomly reordering the 1,526 dependent variable observations (sampling without replacement) and running an OLS regression. This is repeated 10,000 times, creating a distribution of least-squares coefficient vectors. The bootstrapped \( p \)-values are calculated by finding the location of the original coefficient vector in the ranked empirical distribution, variable by variable. The two-tailed \( p \)-values reported are calculated by doubling the percentile location. Intuitively, this simulation procedure answers the question “How likely is it to observe a value at least as large (in absolute value) as the original least squares coefficient estimate if there is no true relation, given the empirical distribution of the dependent variable?” The bootstrapped \( p \)-values that are reported are similar to the ordinary least squares values.
is surprisingly low. I would have expected that the average beta would be slightly above 1.0, given the findings of Clarkson and Thompson (1990).

IV. Summary and Conclusions

This paper has documented the time- and industry-dependence of the long-run performance of initial public offerings. A strategy of investing in IPOs at the end of the first day of public trading and holding them for 3 years would have left the investor with only 83 cents relative to each dollar from investing in a group of matching firms listed on the American and New York stock exchanges. Younger companies and companies going public in heavy volume years did even worse than average. I have attempted to shed some light on the reason for this underperformance. In particular, do the firms in this sample underperform merely due to bad luck, or does the market systematically overestimate the growth opportunities of IPOs? The evidence presented here is broadly consistent with the notion that many firms go public near the peak of industry-specific fads. It should be noted, however, that since the sample involves IPOs going public in only a 10-year period, alternative interpretations cannot be ruled out.

With 20-20 hindsight, investors in the 1,526 IPOs in this sample were overoptimistic about the firms' prospects. There are other securities markets in which investors in new issues have systematically lost money. Weiss (1989) and Peavy (1990) document that investors in new issues of closed-end funds in 1985-87 suffered substantial losses as the funds moved from premiums over net asset value at the time of issue to substantial discounts 6 months later. Elton, Gruber, and Rentzler (1989) document that publicly offered commodity funds going public in 1979-83 performed poorly, in spite of extremely high monthly returns reported in their offering prospectuses. Uhlir (1989) documents a pattern of returns of IPOs of common stock in West Germany that is almost identical to that presented here for the 12 months after going public.

The finding that initial public offerings underperform, on average, implies that the costs of raising external equity capital are not inordinately high for these firms. The high transaction costs of raising external equity capital in an IPO, documented in Ritter (1987) and Barry, Muscarella, and Vetsuypens (1990), are partly offset by the low realized long-run returns, at least for those firms going public at times when investor sentiment is optimistic. Consequently, the small growth companies that predominate among firms going public do not necessarily face a higher cost of equity capital than is true for more established firms.

For issuers, it appears that the concentrations in volume in certain years are associated with taking advantage of “windows of opportunity.” Kim and Stulz (1988) present evidence that issuers take advantage of differences in borrowing costs that periodically arise between the domestic and Eurobond markets. Lee, Shleifer, and Thaler (1991) present evidence that closed-end
funds are issued more frequently in periods when discounts are unusually small. Thus, evidence exists in several markets that issuers successfully time offers to lower their cost of capital.

Several issues have been left unresolved. In particular, I have analyzed the stock market returns in the 3 years after going public without finding any tendency for the underperformance to eventually end. My suspicion, however, is that the underperformance does not extend much beyond 3 years, based upon Ibbotson (1975) and Rao’s (1989) findings. Ibbotson finds no underperformance in the fifth year after going public, the last year that he analyzes. Furthermore, Rao finds negative earnings announcement effects in the first 3 years after going public, but not in years 4 through 6.

A second issue that is unresolved is the generality of my findings. Only by extending the sample period beyond the 10 years of this paper can additional evidence be gained regarding some of the patterns that have been documented. In this regard, Aggarwal and Rivoli (1990) report that IPOs issued in the high-volume years of 1985 and 1986 had negative market-adjusted returns, using a NASDAQ index as the market, during their first year of trading.

A third issue that is unresolved is the relation of the long-run underperformance to the short-run underpricing phenomenon. It has always been somewhat of a mystery why IPOs are priced in a manner that results in such large positive average initial returns. This paper’s evidence indicates that the offering price is not too low, but that the first aftermarket price is too high. If issuers and their investment bankers set the offering price in a manner that reflects the firm’s underlying fundamental value, however, it is even more of a mystery why some offerings have extremely high initial returns.

Appendix: Matching Firm Selection Procedure

To select matching firms for the 1,526 IPOs in 1975–84, the following procedure was employed: Among firms listed on the American and New York Stock Exchanges, their market values were computed on the dates December 31 of 1974, 1980, and 1983. Within each three-digit SIC code, these firms were ranked by market value. For firms going public in 1975–80 in a given three-digit industry, the listed firm with the closest (as of December 31, 1974) market value was chosen as the matching firm, with a matching firm used only once until 3 years had passed. If a matching firm in the same industry was not available, then a small firm in another industry was chosen, with preference given to firms in similar industries. For companies going public in 1981–83, the market value of listed firms at the end of 1980 was used. For firms going public in 1984, the market value of listed firms at the end of 1983 was used. This procedure resulted in 1,526 matching firms, of which 543 (36%) were in the same three-digit industry. An additional 328 firms (21%) were matched by either two-digit SIC codes or by the industry groups as defined in Table VI, resulting in a total of 57% of IPOs matched with a firm
Table AI
Distribution of Market Values for 1,526 IPOs and Matching Firms in 1975-84

Market values for IPOs are calculated using the post-offering number of shares multiplied by the CRSP-reported closing market price on the first day of trading. Market values for matching firms are calculated using the CRSP-reported number of shares for the prior December 31 multiplied by the market price on the date of the IPO. No price level adjustments have been made.

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Market values, $</th>
<th>Percentile</th>
<th>Market values, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>971,014</td>
<td>10th</td>
<td>7,024,060</td>
</tr>
<tr>
<td>10th</td>
<td>10,810,811</td>
<td>20th</td>
<td>14,033,019</td>
</tr>
<tr>
<td>20th</td>
<td>18,724,091</td>
<td>30th</td>
<td>25,987,887</td>
</tr>
<tr>
<td>30th</td>
<td>35,654,362</td>
<td>40th</td>
<td>50,818,888</td>
</tr>
<tr>
<td>40th</td>
<td>67,816,000</td>
<td>50th (median)</td>
<td>76,545,169</td>
</tr>
<tr>
<td>50th (median)</td>
<td>106,260,736</td>
<td>60th</td>
<td>130,024,824</td>
</tr>
<tr>
<td>60th</td>
<td>181,082,000</td>
<td>70th</td>
<td>314,312,192</td>
</tr>
<tr>
<td>70th</td>
<td>314,312,192</td>
<td>80th</td>
<td>719,505,920</td>
</tr>
<tr>
<td>80th</td>
<td>35,028,414,500</td>
<td>90th</td>
<td>35,028,414,500</td>
</tr>
<tr>
<td>90th</td>
<td>106,260,736</td>
<td>100th</td>
<td>35,028,414,500</td>
</tr>
<tr>
<td>100th</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

in roughly the same industry. The low rate of matching industries is attributable to the large difference in the industry mix between the IPOs and the listed companies. Also, as shown in Table AI, in spite of the overrepresentation of small firms among the matching firms, the matching firms (with market values calculated at the time of the IPO, rather than merely at three discrete dates) tend to be larger than the IPOs.

For 277 out of the 1,526 IPOs, the original matching firm was delisted before the earlier of (1) the 3-year anniversary date, or (2) the delisting of the IPO. For these firms, a second matching company was chosen, using the same criteria as above, for the remainder of the aftermarket performance interval. For 61 companies, a third matching firm was needed for the remainder of the interval due to the delisting of the second matching company. A given matching company could be matched with several different IPOs that went public more than 3 years apart. This procedure allowed matching firms with low market capitalizations in certain industries to be used multiple times.

For all of the matching firm choices, care was taken to avoid “survivorship bias.” This was accomplished by choosing a matching firm regardless of when it was delisted, with some matching firms being delisted as soon as a week after the offering date of its matched IPO. Almost all of the Amex-NYSE delistings occurring during this period are due to takeovers and management buyouts. For IPOs that were delisted before the 36 month aftermarket period ended, the last month of returns involves fewer than 21 days. The matching
company's returns were matched up to end on exactly the same day as the IPO.

REFERENCES


Barry, Christopher B., Chris J. Muscarella, and Michael R. Vetsuypens, 1990, Underwriter warrants, underwriter compensation, and the costs of going public, Unpublished working paper, SMU and TCU.


Buser, Stephen A. and K. C. Chan, 1987, NASDAQ/NMS qualification standards, Ohio registration experience and the price performance of initial public offerings, Columbus, Ohio Department of Commerce and National Association of Securities Dealers, Inc.


Chan, Louis and Josef Lakonishok, 1990, Robust measurement of beta risk, Unpublished working paper, University of Illinois.


The Long-Run Performance of Initial Public Offerings


—., 1990, Investor demand for initial public offerings and the relationship of the offer price to the preliminary file range, Unpublished working paper, University of Michigan.
制度内卷、行政干预与共犯结构
——重新解释中国股市困境的一个新视角

葛晋亮

金融 0303 （0304200120）

楼迎军（指导教师）

内容摘要: 中国股票市场在九十年代国企改革面临巨大“资金饥渴”与“还贷风险”的背景下被推上了历史舞台。由于制度设计中的严重缺陷，阻碍了中国股市在寻求富有效率的增长道路上的积极演进。在计划经济制度遗产——强大的行政干预以及行为者“策略选择”的双重作用下，中国股市陷入了一种“制度内卷化”的“动态停滞”中，即整体上表现出“缺乏效率的增长”的变迁轨迹。本文认为，正是由于政府低效的行政干预和企业内部控制人的“共犯结构”，使得中国股票市场十余年发展呈现出低效困顿的景象，并提出积极推进市场化改革才是变“内卷”为“演进”的有效解决途径。

关键词: 行政干预; 制度内卷; 股票市场; 市场化改革; 共犯结构

一、导论

任何一个有效的股票市场必须遵循市场化的原则，在充分发挥自我扩张，自我收缩，自我选择与自我协调的功能下，引导社会资金的转移进而实现社会资源的优化配置与合理组合。市场化教会交易各方如何界定和使用自己的权利，构建起市场参与者从进入到退出详尽有效的游戏规则，并最终引领股票市场在上述完善合理的制度下富有效率地积极演进。

然而中国股票市场 14 年的发展道路未能沿着这样一条道路迈进，它以自己的繁华与艰辛走出了一条边缘于中国经济发展的躁动之路。关于中国股票市场实验不成功的原因所在，历来是讨论争辩的焦点，许多专家学者均作过不同的分析揭示，各种曲直，莫衷一是。

在笔者看来，中国股票市场的发展不成功，关键在于整个市场从诞生开始便卷入了计划经济的制度遗产以及多方利益博弈形成的“共犯结构”二者错综复杂的旋涡之中。

制度遗产置之于我国改革开放前所固守的大背景下便指的是无所不在的政府行政干预和政策主导。当代主流的行政干预理论认为，即便市场的力量足以承担经济活动中
的大部分功能（甚至是一些本应由政府承担的功能），由于市场机制本身的缺陷，有的时候无法使得资源配置达到最优，即使存在所谓的“市场失灵”，需要政府发挥拾遗补缺的作用。因此在这种情况下，政府是市场的替代和补充。政府通过政策的制定与执行，纠正某些由市场带来的经济缺陷。但事实上，这一理论是建立在西方发达国家相对纯粹的市场经济体系之上的。它缺乏对于不同经济环境，尤其是我国这样由传统计划经济体制转轨而来的经济体系的普遍适应性。

首先，作为一种制度安排，我国市场经济体制的产生和发展具有与西方发达国家不同的初始状态和约束条件。它并不是一个伴随现代经济发展应运而生的自发演变过程，而是一个以设计和干预为主导的制度创新和突变过程。在这样的客观制度环境下，政府需要介入与干预的地方往往很多。频繁的干预亦使得政府作用与市场作用的有效边界趋于模糊。在我国，政府部门既是社会经济的管理者，又是国有产权的主体。1作为前者，政府应本着无所有制歧视的态度提供各项公共服务，而作为后者，政府又不得为实现国有资产的保值增值而向国有企业实行一定的政策倾斜。所以，这种双重的身份使得政府在作出决策时，往往无法做到合理公平，集中表现为其在干预过程中的“越位”、“错位”与“缺位”现象。

其次，市场失灵并不意味着可以推出政府干预的必然有效性。正如Samuleson所讲的“政府矫正市场失灵的企图可能使之更坏或引起其他问题的出现。”即由于“统治者的偏好和有界理性，意志形态刚性，官僚政治，集团利益冲突和社会科学知识的局限性”，导致政府干预的无效或低效，主要表现为政府管制、官僚主义和寻租(rent-seeking)行为2。这种干预非但不能弥补市场失灵，反而抑制了市场机制的正常运作。我国股票市场的设计者与监管层正是通过类似强大的行政干预人为地介入于股市发展的方方面面，并最终导致了市场的低效。在大量的行政干预下，从政府财政，银行到股市，其只是在一味地复制旧有行为模式或关系来应对制度内卷(Institutional Involution)的困境，致使在股市发展过程中纠结出许多复杂的对立力量，最主要的便是后者所提及的上市公司，集团公司与地方政府三者。它们作为市场参与的主体，由于彼此之间复杂敏感的利益纠结，逐渐形成了一股阻碍股市市场化进程的“共犯结构”，最终形成了我国股票市场中上市公司经营能力与治理结构水平的低下。这种制度遗下强大的行政干预与“共犯结构”的相互

1张维迎，2001年， 《产权、政府与信誉》生活·读书·新知三联书店，  
2R. 科斯，A. 阿尔钦，D. 诺斯等著： 《财产权利与制度变迁——产权学派与新制度学派译文集》，上海三联书店，1994年版，第397页
作用，使得我国的股票市场无论是组成结构还是运行机制都充满了政治痕迹，在发展到一定程度后便无法向更高水平发展，形成“缺乏发展的增长”及“功能弱化”，即本文所谓的“制度内卷化”倾向。

二、制度内卷化(Involution)

（一）“内卷”的概念及渗透

关于历史演进的复杂过程，文化人类学家塞维思（Elman Rogers Service）认为，人类历史的发展是在“革命(Revolution)”、“演化(evolution)”、“内卷(involution)”等三种动力的循环作用下前进的。Involution源于拉丁语的involutum，愿意为“转或卷起来”，既有复杂的、错杂的、卷成螺旋状等意思，又含内旋、衰退和消散之意，表达了一种“演化过程中复杂的退缩力量”。

“内卷”概念的最初使用始自人类学家葛尔茨在研究印尼爪哇水稻农业后的发现。长久以来爪哇地区的地理环境及耕种传统的稳定性使得爪哇的农业生产长期以来只是不断的重复简单再生产，农产量的增长凭借的是对传统生产方式的不断强化而并非现代技术的革新。这样的农业体制作用于整个爪哇经济，则表现为在达到一定发展阶段后便停滞不前，无法向更高一级的模式转化，形成“内卷”现象。

英国汉学家黄宗旨（P.Huang）在葛尔茨的研究基础上，对长江三角洲和华北小农经济运行逻辑进行了分析与解释。他指出，人们对土地的压力和耕地的缩减使农民趋于“过密化”。即以劳动单位日边际报酬递减为代价换取单位面积劳动力投入的增加。生产越是“过密化”，就越难于把劳动力抽出去走通过资本化提高劳动生产率的道路。黄宗旨称此种“没有发展的增长”为“内卷化”或“过密化”，改革后长江三角洲农村正是通过乡镇企业才改变了这一运行了几百年的运行逻辑。

近代学者将“内卷化”的概念逐步渗透到国家的体制变革之中，做了更高层次上的理论剖析。美国学者杜赞奇（Prasenjit Duara）借用“内卷化”的概念研究了二十世纪前期中国国家政权的扩张及现代化过程。他从更广泛的意义上总结了国家政权的内卷化是指国家机构不再是靠提高旧有或新增机构的效益，而是靠复制或扩大旧有国家社会关系

---

3 Elman Rogers Service 著，黄宝伟等译，《文化进化论》，北京华夏出版社，1991 年版，第 9 页

4 P. Huang, 1990; The Peasant Family and Rural Development in Yangzi Deltal 1350-1988, Stanford, Stanford University Press

5 杜赞奇，《现代化的陷阱——1900~1942 年中国国家政权的扩张对华北农村社会的影响》，战略与管理，1994 年第四期，第 38-51 页
来扩大其行政职能。二十世纪后半期处于转型阶段的中国同样身处“制度内卷”的困扰。社会主义的制度遗产形成了巨大的“路径依赖”（path dependence），其与行动者在其中的“策略选择”一道，构成了社会主义国家创新与转型的复杂命题。

今天，“内卷化”同样渗透在催生中国股票市场的主体——国有企业之中。大陆学者李培林与张冀在组织了对我国十大城市500余家国有企业的问卷调查后，结合各方佐证提出，国有企业作为理性决策者，在追求综合福利最大化的逻辑促动下，其行为不自觉地朝着功能内卷化及人员过密化的方向展开：福利保障作为利润的替代指标引导企业将福利功能向企业内部转移和扩展，同时这样的功能内卷又强化了企业人员的过密化。即在企业总产值递增的情况下，由于人员及相应福利支出的增多而出现人均效益产出递减，企业劳动生产率降低，经营效益恶化。

（二）“制度内卷化”的成因

诺斯曾指出，现在和未来是通过一个独立制度的连续性与过去联系在一起的……过去只有在被视为一个制度演进的历程时，才可以理解。将制度整合到经济理论与经济研究中去，是推进理论与历史的实质性一步。这也是为什么“制度规则”在当今世界的经济发展中受到如此重视的原因。所谓“制度”包括了正式规则（宪法，法律，规定）与非正式的限制（惯例，行事准则，行为规范）以及上述规则与限制的有效执行者三方面。一些符合特定历史条件的经济制度可以提供鼓励持续创新的激励机制，但在技术变迁，信息成本与相对价格不断发生变化时，制度便需要进行调整，即所谓的“制度变迁”。但是任何一种新制度的构建都要受到前一时期所遗留下来的制度遗产的制约，它所产生的一种向后内卷，卷曲的力量便是“转型学”中最重要的概念之一——路径依赖（path dependence）。路径依赖实际上是一种历史的变化或惯性特征。就制度而言，制度变迁一旦走上了某一条路径，制度就会沿着既定的路径进入良性循环轨道，或顺着原来的错误路径前行，甚至被固化在某种无效率的状态下难以改变。

同时，由于制度本身需要通过组织展现其功能与绩效。因此，制度与组织（及其中的行动者）之间便形成了一种相互影响的关系。如果该组织（行为者）是原有制度下的既得利益集团，一旦有新的制度，哪怕更有效率，一旦影响利益格局（原由路径），他们便会反对这种改革和创新。“路径依赖”和行动者“策略选择”二者的共同作用及由此而形成的乘数效应便构成了“制度内卷”的主要成因。

张炜，2004年，《中国金融——制度结构与制度创新》，中国金融出版社，第25页
中国股票市场作为典型的金融市场，其运行与演进也同样内置于一个特定具体的制度框架下，因此也同样受到制度遗产及行为者“策略选择”的双重影响。中国股票市场在中国改革开放后由计划经济体制向市场经济体制转轨的大背景下诞生，因此计划经济的制度遗产在很大程度上作用于股市的发展演变，而市场主体——上市公司及其集团、公司与地方政府所共同缔结成的“共犯结构”也以多种方式影响着市场决策层的“策略选择”。事实上，行政干预及政策主导作为计划经济下制度遗产最为典型的代表同时也是上述行为者“策略选择”得以实施的前提条件，串联着“制度内卷化”成因的两大方面。为此笔者在本文中，试图通过对“行政干预”在整个中国股市发展道路上所扮演的重要角色的剖析探究中国股市演进过程中的困顿所在，从而进一步揭示推进市场化改革在变“内卷”为“演进”过程中的重要意义。

三、行政干预与中国股票市场

提及中国股票市场的产生，就不得不谈到九十年代的国企改革。中国国有企业改革的历史进程大致可以分为三个阶段：第一阶段是从 1979 年开始推行的扩大国营企业自主权，增加企业留利的改革，包括随后实行的承包经营责任制与厂长（经理）负责制；第二阶段是以 1992 年 7 月《全民所有制工业企业转换经营机制条例》的颁布与实施为标志，开始进行的“转机建制”，它确立了国企的法人地位，实行自主经营，自负盈亏的改革模式；第三阶段始于 1997 年，中共十五大重新阐述了对公有制和国有经济主体地位的认识，国有经济在一些领域实行了“战略性退出”，实行大规模改制和资产重组，实现投资主体多元化和股权结构多元化。

回顾国有改革的历史不难发现，国企改革的过程实际是“路径依赖”的。第一阶段锁定产权改革的道路是受到旧有的经济体制束缚及当事人（改革者）认知能力的局限。而在“扩权让利”的过程中，由于行政管理体制尚未改革，政府职能尚未转变，中央下放给企业的权利也被既得利益集团层层瓜分。为避免地方政府和行业组织对企业的干预，中央将承包经营制引入国企改革。承包制虽然暂时解放了企业的生产力，但从长远看导致了机会主义和短期行为，企业竞争乏力的弊端在市场体制的作用下马上显露无遗。经过十几年改革开放，随着决策者认知能力的大幅度提高，股份制经营被引入。但

—周叔莲，1998 年，《二十年来中国国有企业改革的回顾与展望》，《中国社会科学》，第 6 期，P44-48 页
由于原有制度结构的惯性作用和利益集团的内部冲突，国企改制大都采取了国有独资和
国有控股的方式，产生了经营者大量的“寻租”行为。国家为了割断政府与企业的联系，
减少政府直接治理的成本，又开始实施股权私有化，并从一些行业中退出。然而股权稀
释后的国企由于目标不一致，信息不对称，外部股东和国家利益时常受到损害，资产流
失，作假成风。鉴于此，“公司治理”的理念开始被引入并受到逐步重视。

从上述分析中可以看出，国企改革之所以形成“路径依赖”与政府的强大干预是紧
密相连的。应该说，国有企业作为国有资产的最大承担者对国民经济发展有着不可估量
的影响，国企要改革，脱离国家的宏观指导是不现实的，但问题是中国国有企业改革的
道路不是经过科学规划的。形象一点讲，国有企业是被政府推着走上了一条由决策者选
择的路径。在“摸着石头过河”的过程中，政府似乎比国企走得更痛苦，需要随时通过
干预保证国企改革在既定的道路上前进。久而久之这种稳固的力量成为了一种惯性，一
种自然，并渗透到市场经济转轨过程中的方方面面，它在换来改革表面平静的同时却失
去了许多部门通过自我调整适应市场的机会。

不幸的是，中国股票市场的演进再一次重演着这一历史。国企改革中面临两大问题，
一是资金来源，二是国有资产保值增值。事实上，随着国有企业资产负债率的不断攀高，
巨大的资金缺口及持续增加的贷款风险已使得从中央财政到银行均再无力承担，各方都
急需寻找一条新的融资渠道解决资金及风险两大问题。中国股市正是在这样的初衷下
被催生的。作为不同于银行债务融资的一种方法，通过股市融资的主要成本是当期的股
息支付和投资者预期的未来股息增长。而中国股票市场上的投资者大都并不指望通过获
取公司的派息来得到投资回报，而更注重在市场的短期投机行为中获得资本利得。因此
对上市公司而言，在获得源源不断融资的同时，却只需承担一种“软性”的机会成本。
这便有效地减轻了国企还贷的压力，同时也将中央财政与银行从先前的贷款压力下释放
了出来。由此被视为能较好地解决政府，企业，银行等多方矛盾的中国股市在一既定
目标的促发下开始了它的困顿演进。

四、 “制度内卷”下的中国股市困顿与演进

(一) 中国股票市场的总体发展状况

8刘美平，2000 年， 《论国有企业与股票市场之间的整合》 , 《改革》 , 第 4 期，P41-49 页
1、中国股票市场的现状

中国股市自1992年诞生至今，经过十余年的高速发展，其市场规模已相当可观。截止2004年6月，沪深两市上市公司（家数）共计1429家，其中A股上市公司1318家，B股上市公司111家。A股发行总额达到6202亿股，流通市值13004亿元人民币。上市公司横跨机械，电子通信，石化，汽车，交通运输，房地产，金融，能源，生物医药，纺织等数十个对国民经济具有举足轻重作用的行业板块。此外，目前已成立的证券公司已达133家，证券经营网点约3000个，证券公司总资产约人民币5700亿元。截止2004年2月，沪深两市的开户总数达7012万户。

另外，中国股市在自身发展壮大的十余年中为国有企业所做的贡献同样有目共睹。从融资角度看，近1400家上市公司在境内通过IPO，增发，配股的融资总额为9300多亿元。这其中，近900家国有控股及参股的上市公司融资6500亿，约占总融资额的70%。其次，从国有资产保值增值的角度看，今天近900家国有上市公司所形成的3474亿股在成立时所投入的净资产总额为3982亿元，平均每股净资产在1.15元左右。而截止2005年6月，我国上市公司平均每股净资产为2.84元，而上市公司的市场平均价格为4.95元。若以2.84元/股的净资产价计算，近900家国有上市公司的国有帐面价值为9866亿元，资产增值2.48倍。若以4.95元/股的市场价格计算，所有国有股的市值为1.72万亿元，资产增值为4.32倍。

2、中国股票市场的发展历程

由于决策者确立以“筹资”而非以“投资回报”作为股市的初始定位，就必然要求以赶超作为股市发展的主基调，超常规地推动国有企业上市便是其中的典型代表，如图1所示。

![中国股市上市公司家数增长趋势(1994--2003)](image)

资料来源：《中国证券市场统计年鉴2003》
从上图中不难发现，从 1994 年至 2003 年，上市家数以每年 的速度增长，总市值与流通市值亦不断攀高。进一步分析可以发现，从 1999 年下半年开始至 2001 年上半年这两年的股市大牛市期间，新增的上市家数及流通市值均达到顶峰，监管层借股市繁荣之机扩大筹资规模的行政干预行为表露无遗。

（二）“行政干预下” 中国股市的躁动之旅

中国股市的发展现状从表面上看规模不断扩大，影响也不断深化，如图 2 所示。但其背后却隐藏着太多的行政因素。一张一弛之间，无不透露出外压之下的躁动。

图 2 中国股市流通市值与总市值增长趋势（1993--2003）

可以发现，中国股市过往十余年的发展历程，是一个层次分明，棱角清晰的阶梯式运动过程，与全球股市尤其是发达股市通常表现出的和缓渐进式延伸大为不同。从 1991 年到 2001 年期间，中国股市主要是靠七波宏阔的上涨浪潮，成就了 10 年牛市辉煌。除却这七波，中国股市就只有一身憔悴，两袖清风了。

此外，进一步分析中国股市最大涨跌幅及股市的波动性可以发现，中国股市不是象发达市场中普遍和正常的情况那样，不知不觉中缓慢地积蓄力量，而是靠着个别交易日的突飞猛进，实现了量子般跳跃式的行进，如表 1 所示。

注：高潮生，2005 年，《重解中国股市》，《财经》总第 128 期，P28-35 页
表1 中国股市的波动性

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>最高点</td>
<td>1429</td>
<td>1558</td>
<td>1052</td>
<td>925</td>
<td>1258</td>
<td>1509</td>
<td>1422</td>
<td>1756</td>
<td>2125</td>
<td>2245</td>
<td>1748</td>
<td>1650</td>
</tr>
<tr>
<td>最低点</td>
<td>292</td>
<td>750</td>
<td>325</td>
<td>524</td>
<td>512</td>
<td>870</td>
<td>1042</td>
<td>1047</td>
<td>1361</td>
<td>1514</td>
<td>1311</td>
<td>1307</td>
</tr>
<tr>
<td>振幅</td>
<td>389%</td>
<td>107%</td>
<td>223%</td>
<td>76%</td>
<td>145%</td>
<td>73%</td>
<td>36%</td>
<td>67%</td>
<td>56%</td>
<td>48%</td>
<td>33%</td>
<td>26%</td>
</tr>
</tbody>
</table>

注：振幅=（全年最高指数-全年最低指数）/全年最低指数

如果将历年股市的行情置之于政策面的框架下，便可清楚地看到，几乎所有“井喷”行情背后都是相应的重大政策措施的出台。因此，与其说中国股市的十年辉煌是在“富有中国特色的道路上”自行演绎的，不如说是政策与监管层苦心经营，频频干预所造就的。今天，中国股市是“政策市”、“市场边缘化”、“股市监管呼唤理念”的种种强烈呼声让我们不得不重新将目光投回整个股市的体制设计与监管层功能定位的角度上。而事实上，这正是“内卷化”成因之一，计划经济体制下的制度遗产——“行政干预”的后果。中国股市的监管层带有计划经济体制下行政部门的浓厚色彩，对其监管对象仍有要者的心理倾向和行为习惯，监管和掌管的角色常常混淆错位。监管部门背负着维护股市健康向上运行，从而不断扩大融资规模的行政目标。因此，市场的稳定与繁荣，指数的稳健上行是始终萦绕其心头的重任之重。当市场低迷时，监管层对于投机活动就比较容忍，甚至对某些明显的违规违法行为视而不见，借此凝聚“人气”，激活市场；相反，市场火爆时，又惊慌失措，不是采取宏观调控的治本方法，而是频频出台微观意义上的行政措施压抑市场的正常活动。久而久之，上市公司开始倾向于迎合监管层的意图，投资者“政策救市”理念固化，股市在追求健全合理制度下自我发展的途径趋于末路，转而陷入了“制度内卷化”缺乏效率的发展困境之中。

（三）行为者的策略性选择

在制度变革初显端倪时，主流组织的领导者会对变革进行自己的预期。作为制度的执行者，必然会将自己的利益得失与这种预期相权衡，从而通过所掌握的决策权对变迁进行“策略选择”。如果说行政干预是股市发展过程中的具体表现，那么“策略选择”则是行为者先于股市运行而确定的制度规范。它们从一开始便构成了股市演进中的制度与结构缺陷，是行政干预的固化体现，同时也为市场的“内卷”化倾向埋下了伏笔。

9
1、行为者“策略选择”之一：上市及退市制度

从 90 年上交所成立至今，证券的发行审核方式可谓变化多端，从额度管理、指标管理、到通道制，再到保荐制。虽然形式上做了多种调整，却始终只是一种权利的转移或变迁，地方政府与上市公司间“合谋共犯”的实质并没有改变。早期额度管理和指标管理，都是以指令性计划为特点的“审批制”，地方政府及一些部委在股票发行上拥有很大的权利，故渴望上市的企业必须与其主管部门及地方政府建立紧密的关系。只要二者利益趋同，企业便可获得宝贵的上市指标。其后由于该行政化的审批制度引发了严重的权利“寻租”，证监会于 2001 年 3 月开始实施了“核准制”下的“通道制”，也就是向各类综合类券商下达可推荐拟公开发行股票的企业家数。但实际上，由于大部分证券公司都是地方政府所有，因此原先地方政府过度参与企业上市的行为并没有根本性变化，“跑部钱进”的腐败现象仍在继续。

与上市审核制度的多变相对应的却是退市机制的“停滞不前”。从 2001 年退市条例公布至今，上市公司真正实行退市的数量稀少。归根结底其原因是一样的，上市公司隶属于地方政府，对于隶属于哪个政府的上市公司退出，不仅为政者的业绩形象会有所损害，而且会将很大的一块利益损失掉。因此在地方政府看来，其管辖下的每一家上市公司都有继续存在下来的理由，也都有得到充分保护的理由，作为和地方政府同级的证监会要推行并实施退市机制会有多困难也就不言而喻了。

2、行为者“策略选择”之二：股权分置

中国股票市场建立和发展的初期，公有制经济仍占主导地位，国有股绝对控股是市场的前提规划。为防止国有资产流失，相关部门出台了国有股、法人股“暂不流通”的规定，从而形成了中国股票市场的另一突出特点——股权分置。虽然在市场运行的初期这样的流通结构促进了证券市场的快速发展，却也构成了整个市场重大的制度缺陷，造成了诸多弊端。

首先，股权分置形成了非流通大股东在上市、配股、增发等一系列过程中，仅用少量原始出资就可以将大量流通股东的资金圈入自己囊中，使股票市场成为一个“圈钱陷阱”。不同股东无法实现“同股同权，同股同利”，市场公平性受到破坏。第二，造就了中国股市高度的投机性。在“圈钱模式”下发行的股票，其上市公司根本不具有投资价值，没有一个战略投资者可以从长期的持股中获益，这样的市场必然充满了投机性。第三，由于大股东的股票不流通，也没有真正的市场价格。大小股东的价值取向完全背
离，股价与上市公司的价值也完全脱离，导致大股东根本没有动力去劳神费力地搞好经营回报普通投资者。市场沦为少数大股东的“抽水机”。

在今天的市场上，关于股权分置改革是点击率最高的字眼之一，而其中就如何对流通股股东进行赔偿的问题成为改革进程中最为棘手的焦点。在笔者看来，其实就股权分置本身而言，假如政府不是绝大部分上市公司的最大股东，股民也不会提出要求政府在解决分置问题时对股民进行赔偿。因为可以肯定的说，一个私营企业的老板宁可看到自己股票的价格大幅下跌，也不会掏腰包为小股东买单，而反过来，理性的中小股东也不会要求他这么做。所以，股权分置反映的是中国经济深层次的结构矛盾，其实质是政府作为市场管理者与最大的市场参与者双重身份的矛盾。

3、行为者“策略选择”之三：“行政干预”渗透下的公司治理

其实，政府的行政干预不仅仅体现在笔者上文所提及的上市与退市制上，同时也渗透在上市公司内部治理的方方面面。政府在经济生活中同为管理者和股东的双重身份为其参与到上市公司的微观经济活动中奠定了基础。政府特别是地方政府作为大部分上市公司的控股股东对上市公司治理结构所产生的深刻影响成为我国上市公司治理体制中极具“中国特色”的一点。

一方面，地方政府往往从地方局部利益或某些集团利益的私利出发，过多地插手或干预上市公司的设立、重组和增资扩股工作。当企业为达到上市条件进行所谓的“资产重组”时，他们把并购重组限定在自己的地区或行业内，不允许本地和本行业的上市公司去购并其他的公司，无论是买与卖都必须是在本地区或本行业间进行。这种行为严重束缚了上市公司进行资产重组的积极性，阻挠了上市公司和其他企业的的发展，歪曲了国企改革的本意，延缓了国企扭亏的进程。此外，政府还运用各项政策手段进一步影响并购本身，通过税收，融资等方法间接引导企业重组。对本地地区本行业的重点上市公司实行优惠政策，对其增发新股，配股等进行特许，对其筹集所急需的大量资金。这样的政策倾斜往往并不从发展经济，结构调整的角度出发，而是表现为地方保护主义的不公平竞争与压制。这与并购重组本身从调整中实现经营效益的提高和资产质量改善的目的无疑是背道而驰。

另一方面，由于许多上市公司是由原国有企业作为主要发起人设立的，原国有企业或国有资产管理部门就理所当然地成为公司的第一大股东，在上市公司的董事会及股东大会中具有多数表决权，至此，上市公司的最高管理层便要深深受到政府行政干预的左
通过对董事长、总经理等高层职位的重叠任命，进一步提高对企业控制效率。职位重叠的直接后果便是企业治理的真空状态，形成“内部人控制”。见表2所示。严重的“内部人控制”使得股东大会流于形式，董事会治理机制虚置，监督会难以形成对财务状况，董事及经理层的有效监督，并直接导致了我国上市公司内部职务消费过分、信息披露不规范、短期行为、过度投资和耗费资产、转移国有资产、置小股东利益于不顾、不分红或少分红、大量拖欠债务等等违规违法行为。

### 表2 企业“内部人控制率”调查

<table>
<thead>
<tr>
<th>内部人控制率</th>
<th>占样本数比例</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>20.4%</td>
</tr>
<tr>
<td>70%—99%</td>
<td>21.2%</td>
</tr>
<tr>
<td>50%—69%</td>
<td>35.7%</td>
</tr>
<tr>
<td>30%—49%</td>
<td>13.3%</td>
</tr>
<tr>
<td>0—30%</td>
<td>9.4%</td>
</tr>
</tbody>
</table>

资料来源：范黎波，李自杰，《企业理论与公司治理》，对外经济贸易大学出版社，2001年。

因此，所谓的“包装改制上市”的国企，实是以筹资为最主要目的的。公司治理结构只是形式上的改变，或是蜻蜓点水，或是片面空洞，而其中地方政府的默许甚至是共谋，更是强化了这一点。

综上所述，由于计划经济体制下的制度遗产——“行政干预”理念在企业从内到外，从地方到中央的环环渗透，促使决策者“策略选择”趋于低效或是无效。它造成了股市困顿演进过程中制度上的根本性缺陷，使得股市在发展到一定规模后无法向更高形式转化。从严重缺乏激励机制的上市退市制度到现今市场最大症结所在的股权分置，再到混乱不堪的上市公司的治理情况，都是行政干预下的路径依赖及利益博弈下地方政府与企业“共谋”的最典型体现。他们不断的纵深演进，使中国股市的制度变革形成了一种“动态停滞”，并最终形成一股了“内卷化”力量迫使中国股市朝向衰败的方向前进。

中国股市十余年发展所教会投资者的东西要远比教给上市公司或政府部门的来得多的多。二级市场高达近万亿的市值缩水不仅仅让7000余万股民尝到巨额亏损的切肤之痛，更让他们在跳动的股价与指数背后对股市投资的真正含义有了逐渐清醒的认识。今天的股市真实地反映了我们的企业，企管和员工队伍的现实素质，如表3。
表 3 中国上市公司历年业绩（1992-2003）

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>每股收益（元）</td>
<td>0.34</td>
<td>0.31</td>
<td>0.33</td>
<td>0.25</td>
<td>0.29</td>
<td>0.25</td>
</tr>
<tr>
<td>净资产收益率（%）</td>
<td>13.86</td>
<td>13.82</td>
<td>13.1</td>
<td>11.03</td>
<td>16.05</td>
<td>10.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>年份</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>每股收益（元）</td>
<td>0.19</td>
<td>0.21</td>
<td>0.2</td>
<td>0.12</td>
<td>0.15</td>
</tr>
<tr>
<td>净资产收益率（%）</td>
<td>7.62</td>
<td>8.28</td>
<td>7.72</td>
<td>5.47</td>
<td>6.46</td>
</tr>
</tbody>
</table>

资料来源：上海证券交易所公开资料

由表3可见，除了个别年份如1996年和2003年外，我国上市公司的整体经营业绩是趋于下降的。另据调查，在1994年以前上市的177家A股企业，从1993年以来净利润逐年下降，93年实现净利润127.35亿人民币，96年为105.22亿，99年为98亿，净利润减少的企业有88家，占全部公司的49.7%，亏损22家，占12.43%。另外还有40家公司实际亏损，是靠其他营业外收入才得以维持账面平衡，占22.59%。换言之，177家公司上市公司六年后，只剩27家公司（占15.25%）本业的净利润总额维持增长。

五、“制度内卷”到“制度外卷”

由于上市公司经营业绩蕴涵了太高的投资风险，让逐渐恢复理性投资观的股民感到迷惘与心寒。过高的风险和过低的盈利能力加速了投资者追求股票真实价值的步伐，而价值回归的唯一途径便是股价的下跌。从2001年下半年起至今，沪深股市经历了近四年的下跌，从最初的2240点跌至目前的1100点左右，跌幅近50%，并曾一度洞穿千点大关。市场市值也从01年6月的18866亿元减少到目前约10000亿元，缩水近8000亿元。

与投资者惨痛却积极地寻求价值回归的行为形成鲜明对比的，是政府与监管层更为频繁的行政干预，但遗憾的是，政策的效应正在明显递减\(^\text{12}\)。仔细分析这些政策可以发现，监管层仍在延续按供求决定价格的传统思维方式进行运作。几乎所有的利好政策和措施仅仅指向了股票的供给与需求，而并非旨在提升上市公司的经营绩效或是健全完善

\(^{12}\)从“国九条”公布，暂停新股发行，到鼓励保险基金入市，允许银行设立基金公司，再到酝酿市场“平准基金”，下调50%的印花税等等。
市场运行制度，因而无法对股票的内在价值产生实质性的影响。

由此可见，在一个已近形成“内卷”的市场中再延续之前的旧有路径，不断强化“行政干预”的力度与密度，非但不能引导股市走出低谷反而会使其在“缺乏效率”的演进过程中越陷越深，越“卷”越乏。事实上，在股市经历了自2001年夏以来近四年的熊市之后，越来越多的人对市场有着前所未有的清醒，“中国股市告别救世主”，正成为众多投资者的心声。在笔者看来，“告别救世主”就是要告别过度的“行政干预”，扭转市场上尚存的“救市思维”，结束“井喷”式幻想，坚定对股市进行基础性改革的决心和信心，从“制度内卷”的困境中摆脱出来。而推进市场化改革——变“内卷”为“演进”的唯一出路。

中国股市是90年代末以来中国国民经济各部门中唯一一个没有推进市场化改革，反而向相反的行政集权的管制体制大幅退缩的部门，这也正是股市的走势与国民经济改革发展的走势明显背道而驰的原因。实践以有力地证明，这样高度行政集权的市场管理方式不但不能有效地控制股市风险，反而在不断加大风险，降低市场运行效率，致使市场陷于“制度内卷”的状态中。要彻底改变这种“内卷”的状态，只有从中国股市的症结所在——过度的“行政干预”出发，坚定不移地推进市场化改革，尊重证券市场的运行规律，才能变“内卷”为积极演进，走出一条积极向上，富有效率的发展道路。

其实近几年来，政府及监管部门亦开始重视起市场化改革的必要性，并在努力付诸于实践。但由于先行政管体制的市场基调下形成了多方错综复杂的利益纠葛，市场化改革的每一步都走的很不容易。众所周知，在股市的流通领域，股权分置问题已成为制约和困扰二级市场发展最关键动因之一。大量上市公司的股权不能在证券市场上流通，金融资本无法通过股票的买卖实现对部分上市公司的控制权的转移，产业资本也无法通过证券交易进入和退出上市公司。分置的股权结构割裂了股票与其背后的实体经济间的套利关系，使市场的参与者某种程度上都构成了股票的投机商，股市整体的高风险由此可见一斑。自“国九条”明确提出“积极稳妥解决股权分置问题”的原则以来，围绕着股权分置解决办法的各种设计方案层出不穷，在一定程度上反映了市场的急躁心态。对股权分置改革，应和中国股市市场化改革的原则下进行规划设计，单纯以技术层次推动改革很有可能走偏方向。

本着市场化改革的原则推进股权分置改革，应努力作到以下几点：首先，切忌闭
门造车，要尊重市场各方的意见，少用主观的行政意志去作判断。目前市场上通过“分
散决策”比较好的体现了这一点。改革的具体方案由各上市公司流通与非流通股股东之
间一对一谈判解决，可以充分兼顾上市公司与股东的权利。其次，要真正从促进资本市
场发展的大目标、大战略出发，避免在进行改革程序，原则设计时人为地增加垄断，保
护少数机构利益。这里笔者想提一提国资委在处理流通股股东赔偿问题上的态度。国资
委作为国有资产的所有者，国有股全面流通是其所希望看到的，但这一过程却应付出相
应的代价。事实上面对中国股市登记在册的7000余万股民，其中真正参与市场的股民
个人人数则少之又少；相反，很多利益集团，机构投资者在其中有大量的投资，而国
有资产是全民共有的，一旦在国有股流通的问题上出现不合理的补偿，必将损害全体人
民的利益。因此，国资委在表明其对于流通股股东如何进行补偿的问题上应先充分考
虑这一点，给出公平合理、具有可操作性的建议。最后，在方案决策中要真正体现“公
平公正公开”的原则，制定科学的股东投票机制。特别要考虑广大中小股东专业知识不足，
极其分散等特点，提出真正反映股民意愿的投票方案。

在一级市场的融资领域，从2005年1月起，证监会改革了新股发行制度，开始实
施旨在推进市场化改革的询价制。首次公开发行股票的公司及其保荐机构应向证券
投资基金公司、证券公司、信托投资公司、财务公司、保险公司和机构投资者（QFII）等询价对象询价的方式来确定股票的发行价格。

本文之前已经提出，在新股上市的过程中，以前的“审批制”蕴涵了太多的行政干
预，加之上市公司、集团公司的利益，导致“寻租”泛滥，一级市场上“圈钱”成灾。作为改革中国股市制度缺陷中的重要一环，笔者
认为询价制的推出是新股定价市场化改革的重大突破，对改变困扰股市的“制度内卷化”倾
t向大有裨益。首先，询价制改变了过去一级市场由发行者与监管层力量悬殊的单
一博弈局面，形成由发行人、承销商和机构投资者之间的博弈态势。这样的局面可以形成比较准确地反映出股票内在价值的发行价
格，促使新股定价的理性回归。其次，询价制试行以后，以往为争夺一级市场而“跑部
攻关”，“寻租成风”的现象将得到有效改善。由于一旦对拟投资新股的价位定高或定低
均会给博弈双方造成损失，利益鞭策的机制，只能促使机构投资者对该上市公司进行全方
位的研究，并对其盈利能力作出客观的长期预测。这种客观细致的分析预测反过来也会
give the company pressure. After all, the amount of IPO financing needs to be determined by the company itself.
构投资者主动的证明。这样双向的压力可以促进一、二级市场趋于均衡，从而在真正意义上提高证券市场的资源配置效率。

由于询价制推出的时机正值沪深股市持续下跌的深渊，加上股权分置改革悬而未决，市场更多地以一种悲观的“救市”心态对其作出了消极回应。鉴于大盘的颓势，新股发行暂时停止，询价制的进一步完善亦被迫搁置。但询价制作为积极倡导市场定价的IPO方式，其方向是正确无疑的，关键在于要给予询价制运行以一个良好稳定的市场环境。随着对信息披露监管的进一步加强，合规机构投资者的培养以及股权分置改革的有效推进，询价制的实行也会在市场化的原则下逐步走向成熟。

中国股市在过度的行政干预下，在单一注重融资功能的政策驱使下忽视了对自身制度缺陷以及运行弊端的重视与完善，最终陷入了“制度内卷化”下“缺乏效率”的发展困境中。尽管如此，一个经济持续较快发展中的大国是离不开股票市场的辅助作用的。长期而言，中国股市的远景是光明的，是充满希望与活力的。短期而言，中国股市正陷于“制度内卷”的困顿之中，市场各方正在经历泡沫过后的寂寞清醒，但这也恰为我们重新建设股市的基本制度，完善各项法律法规提供了反思的契机。

六、结论

中国股票市场的建立是九十年代继政府与银行以后，解决国有企业改革“资金饥渴”的重要制度设计，也是改革者设计在为中国经济快速发展寻求持续资本供给下的“策略选择”。作为由计划经济向市场经济转轨过程中体制改革的一个部分，中国股市从最初定位到运行始终都被计划经济下的制度遗产——“行政干预”所深深影响与控制着。受制于强大行政干预的股票市场，大量信息难以发挥其内在传导作用，造成市场行为扭曲，市场运行效率过低。此外，不同机会利益者在股市体制变革或制度变迁中利用不完善的制度安排来获取潜在的机会利益，进而进一步损耗了市场效率。二者相互作用下产生的乘数效应，使股市在发展到了一定规模后便无法再寻求高效率的增长，转而陷入了一种“动态停滞”当中，走出了本文所提的“制度内卷化”的困境中。

中国股市在层层监管及频频的政策干预下发展了十余年，走出了自己的辉煌，也走到了困顿的边缘。一味重视融资功能的市场成了国有企业及利益集团的“提款机”，“寻
租场”，泡沫过后，我们看到的只有损失惨重的中小投资者，获利能力低下，造假不断的上市公司，以及存在严重缺陷的股市构架。

中国股市是被行政干预主导下的集权式管理方式带进了“制度内卷化”的深渊的，所以要真正摆脱这一困境就必须重新认识并发挥市场自身的作用。笔者认为，市场在近四年的持续下跌后，经历了泡沫终结的过程，正迎来前所未有的清醒与理性。各方因此为反思契机，重新建立股市制度，完善各项法律法规，积极全面地推行市场化改革，尊重市场自身的选择。惟有如此，中国股市才能真正发挥作为一个资本市场最为关键的资源配置功能，变“内卷化”为积极演进，开始全新的富有效率的发展之旅。

参考文献：
[1] 张维迎，2004年，《中国金融——制度结构与制度创新》，中国金融出版社
[4] 王信贤，2000年，《大陆国企改革的组织同形主义》，《中国大陆研究》，第44卷，第9期
[6] 张维迎，2001年，《产权、政府与信誉》生活·读书·新知三联书店，
[7] 刘伟、魏杰，2002年，《经济学导论》，中国发展出版社。
[8] 李培林，张翼，2000年，《国有企业社会成本分析》，社会科学文献出版社
[9] 杜赞奇，1994年，《现代化的陷阱——1900-1942年中国国家政权的扩张对华北乡村社会的影响》，《战略与管理》，第四期，P38-51页
[12] 刘美平，2000年，《论国有企业与股票市场之间的整合》，《改革》，第4期，P41-49页
[14] Duara,Prasenjit,1988: Culture, Power, and The State ;Rural North China


Analysis of Institutional Involution

Abstract: Chinese stock market was pushed up to the historic stage on the background of deep “hungry for fund” and “the risks of returing loan” which generated from the reform of State-Own enterprise (SOEs) in 90th. As an efficient way to solve the problems among the state finance, locate governments and SOEs, the Chinese stock market developed a way filled with splendor and languish under the strong pressure of administration. Because of the serious flaws of the institution design, Chinese stock market was failed to evolve positive and resultful. Under the both effects of the low efficiency of the adminstration interfere and the “complicity construction” which appeared inside and outside the enterprises, the Chinese stock market fall into a jam called “institution involution” which performs as a state of “dynamic stagnancy”. This paper just try to use the concept of “institution involution” to analyze the reasons of the inefficient development of Chinese stock market, also on the basis of this, giving some advises about how to change the trend of “institution involution”.

Key words: administration interfere; institution involution; stock market; market-reform complicity construction
Duration and Convexity

The price of a bond is a function of the promised payments and the market required rate of return. Since the promised payments are fixed, bond prices change in response to changes in the market determined required rate of return. For investor's who hold bonds, the issue of how sensitive a bond's price is to changes in the required rate of return is important. There are four measures of bond price sensitivity that are commonly used. They are Simple Maturity, Macaulay Duration (effective maturity), Modified Duration, and Convexity. Each of these provides a more exact description of how a bond price changes relative to changes in the required rate of return.

Maturity

Simple maturity is just the time left to maturity on a bond. We generally think of 5-year bonds or 10-year bonds. It is straightforward and requires no calculation. The longer the time to maturity the more sensitive a particular bond is to changes in the required rate of return. Consider two zero coupon bonds, each with a face value of $1,000. Bond A matures in 10 years and has a required rate of return of 10%. The price of Bond A is $376.89, where

$$P_A = \frac{1,000}{(1 + .10/2)^{20}} = 376.89$$

Bond B has a maturity of 5 years and also has a required rate of return of 10%. Its price is $613.91 or

$$P_B = \frac{1,000}{(1 + .10/2)^{10}} = 613.91$$

By convention, zero coupon bonds are compounded on a semi-annual basis. Since almost all US bonds have semi-annual coupon payments, this note will always assume semi-annual compounding unless otherwise noted.
If the required rate of return for each bond was to increase by 100 basis points to 11%, the prices would then be $342.73 for Bond A and $585.43 for Bond B. This translates into a -9.1% change in price for Bond A and -4.6% for Bond B.

Just from the pricing formulation, it is clear that any change in interest rates will have a much greater impact on Bond A than Bond B. This is reinforced in Figure 1, where the price curve for the 10-year bond (Bond A) is much steeper than that for the 5-year bond (Bond B). Thus, for zero coupon bonds simple maturity can be used to compare price sensitivity.

**Macaulay Duration (Effective Maturity)**

The relationship between price and maturity is not as clear when you consider non-zero coupon bonds. For a coupon-paying bond, many of the cash flows occur before the actual maturity of the bond and the relative timing of these cash flows will affect the pricing of the bond. In order to deal with this, Frederick Macaulay\(^2\) in 1938 suggested that investors use the effective maturity of a bond as a measure of interest rate sensitivity. He called this duration and defined it as a value-weighted average of the timing of the cash flows. The easiest way to see this is to use an example. Consider a six-year bond with face value of $1,000, and a 6.1% coupon rate (semi-annual payments). If the current yield to maturity is 10%, the value of the bond is found by discounting each of the semi-annual payments. This is shown in Exhibit 1.

Exhibit 1

Macaulay Duration takes the present value of each payment and divides it by the total bond price, $P$. By doing this, one has a percentage, $w_t$, of the total bond value that is received in each period, $t$.

$$w_t = \frac{C_t}{P \cdot (1 + y)^t}$$

The duration or effective maturity for the bond could then be estimated by multiplying the weight, $w_t$, times the time, $t$ and then summing all of the weighted values, or

$$Duration = \sum_{t=1}^{T} \frac{C_t}{P \cdot (1 + y)^t} \cdot t = \sum_{t=1}^{T} w_t \cdot t.$$

This measure takes into account the relative timing of the cash flows. Calculation of the Macaulay Duration measure is fairly straightforward but can be somewhat tedious. Exhibit 2 shows how a semi-annual duration for the example shown above would be calculated.

---

3 Excel offers a worksheet function DURATION(), which calculates the Macaulay Duration.
Exhibit 2

The semi-annual duration for this bond is 10.014 six-month periods. We usually use annual duration and we annualize the semi-annual duration simply by dividing by 2 (the number of six month periods in a year). In this case, the annualized duration would be 5.007 years. Note that the Macaulay Duration for a 5-year zero coupon bond is the same as the simple maturity, 5.0 years. Hence, we can expect that the original 6-year, 6.1% coupon bond when interest rates change to behave in a manner similar to a 5-year zero coupon bond, since their effective maturity (Macaulay Duration) is essentially the same.

Modified Duration

If we want a more direct measure of the relationship between changes in interest rates and changes in bond prices, we can use Modified Duration. Modified Duration, D, is defined as the following

\[ D = -\frac{1}{P} \frac{\Delta P}{\Delta y} \]

where \( P \) is the bond price, \( \Delta P \) is the change in bond price and \( \Delta y \) is the change in the required rate of return (yield to maturity). For those with a math background, \( \frac{\Delta P}{\Delta y} \) is the first derivative of the bond price with respect to yield to maturity. The basic pricing formulation for bonds is
\[ P = \sum_{t=1}^{T} \frac{C_t}{(1+y)^t} \]

where \( C_t \) is the cash payment received in time period \( t \) and \( y \) is the semi-annual yield to maturity. Taking the derivative of \( P \) with respect to \( y \),

\[ \frac{\Delta P}{\Delta y} = -\frac{1}{(1+y)} \cdot \sum_{t=1}^{T} t \cdot \frac{C_t}{(1+y)^t} \]

Inserting this into the formula for Modified Duration yields,

\[ D = -\frac{1}{P} \left[ -\frac{1}{(1+y)} \sum_{t=1}^{T} t \cdot \frac{C_t}{(1+y)^t} \right] \]

Rearranging the above slightly,

\[ D = \frac{1}{(1+y)} \cdot \sum_{t=1}^{T} C_t \cdot \left(\frac{1}{(1+y)^t} \right) \]

Comparing this to the definition of Macaulay Duration and using that definition we can write Modified Duration as

\[ \text{Modified Duration} = D = \frac{1}{(1+y)} \cdot \text{Macaulay Duration} \]

While it is easy calculate Modified Duration once you have Macaulay Duration the interpretations of the two are quite different. Macaulay Duration is an average or effective maturity. Modified Duration really measures how small changes in the yield to maturity affect the price of the bond. In fact, from the definition of Modified Duration we can write the following relationship:

\[ \frac{\Delta P}{P} = -D \cdot \Delta y \]

or \% change in bond price = - Modified Duration times the change in yield to maturity.

For example, the six-year 6.1\% coupon bond above had a yield to maturity of 10\% and a semi-annual Macaulay Duration of 10.014 (5.007 annual Macaulay Duration). The Modified Duration of this bond is \( \frac{10.014}{(1.05)} \), or 9.537 on a semi-annual basis or \( \frac{9.537}{2} = 4.77 \) years on an
Assuming that the yield to maturity of 10% increases by 25 basis points to 10.25%, based on the Modified Duration of 4.77 years the price of the bond should change by $\frac{\Delta P}{P} = -D \cdot \Delta y = -4.77 \cdot (.25\%) = -1.19\%$. The bond price should drop by 1.19% from $827.17$ to $817.31$ ($827.17 \times (1-.0119) = 817.31$). The actual calculated price at a yield to maturity of 10.25% is $817.38$.

Exhibit 3 shows the Modified Duration price change and the actual calculated price change for different changes in yield to maturity.

**Exhibit 3**

<table>
<thead>
<tr>
<th>New Yield to Maturity</th>
<th>Change in Yield</th>
<th>-D*Change in Yield</th>
<th>Predicted Price</th>
<th>Actual % change*</th>
<th>Actual Price**</th>
<th>Difference***</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.00%</td>
<td>2.00%</td>
<td>-9.54%</td>
<td>748.26</td>
<td>-9.01%</td>
<td>752.68</td>
<td>4.42</td>
</tr>
<tr>
<td>11.75%</td>
<td>1.75%</td>
<td>-8.35%</td>
<td>758.12</td>
<td>-7.94%</td>
<td>761.52</td>
<td>3.40</td>
</tr>
<tr>
<td>11.50%</td>
<td>1.50%</td>
<td>-7.16%</td>
<td>767.99</td>
<td>-6.85%</td>
<td>770.50</td>
<td>2.52</td>
</tr>
<tr>
<td>11.25%</td>
<td>1.25%</td>
<td>-5.96%</td>
<td>777.85</td>
<td>-5.75%</td>
<td>779.61</td>
<td>1.76</td>
</tr>
<tr>
<td>11.00%</td>
<td>1.00%</td>
<td>-4.77%</td>
<td>787.71</td>
<td>-4.63%</td>
<td>788.85</td>
<td>1.13</td>
</tr>
<tr>
<td>10.75%</td>
<td>0.75%</td>
<td>-3.58%</td>
<td>797.58</td>
<td>-3.50%</td>
<td>798.22</td>
<td>0.64</td>
</tr>
<tr>
<td>10.50%</td>
<td>0.50%</td>
<td>-2.39%</td>
<td>807.44</td>
<td>-2.35%</td>
<td>807.73</td>
<td>0.29</td>
</tr>
<tr>
<td>10.25%</td>
<td>0.25%</td>
<td>-1.19%</td>
<td>817.31</td>
<td>-1.18%</td>
<td>817.38</td>
<td>0.07</td>
</tr>
<tr>
<td>10.00%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>827.17</td>
<td>-0.00%</td>
<td>827.17</td>
<td>-</td>
</tr>
<tr>
<td>9.75%</td>
<td>-0.25%</td>
<td>1.19%</td>
<td>837.03</td>
<td>1.20%</td>
<td>837.10</td>
<td>0.07</td>
</tr>
<tr>
<td>9.50%</td>
<td>-0.50%</td>
<td>2.39%</td>
<td>846.90</td>
<td>2.42%</td>
<td>847.18</td>
<td>0.28</td>
</tr>
<tr>
<td>9.25%</td>
<td>-0.75%</td>
<td>3.58%</td>
<td>856.76</td>
<td>3.66%</td>
<td>857.40</td>
<td>0.64</td>
</tr>
<tr>
<td>9.00%</td>
<td>-1.00%</td>
<td>4.77%</td>
<td>866.63</td>
<td>4.91%</td>
<td>867.78</td>
<td>1.15</td>
</tr>
<tr>
<td>8.75%</td>
<td>-1.25%</td>
<td>5.96%</td>
<td>876.49</td>
<td>6.18%</td>
<td>878.31</td>
<td>1.82</td>
</tr>
<tr>
<td>8.50%</td>
<td>-1.50%</td>
<td>7.16%</td>
<td>886.35</td>
<td>7.47%</td>
<td>889.00</td>
<td>2.64</td>
</tr>
<tr>
<td>8.25%</td>
<td>-1.75%</td>
<td>8.35%</td>
<td>896.22</td>
<td>8.79%</td>
<td>899.84</td>
<td>3.62</td>
</tr>
<tr>
<td>8.00%</td>
<td>-2.00%</td>
<td>9.54%</td>
<td>906.08</td>
<td>10.12%</td>
<td>910.84</td>
<td>4.76</td>
</tr>
</tbody>
</table>

* Actual % change is based on the calculated price relative to the price of $827.17.
** Actual price is the calculated price based on the yield to maturity.
*** Difference is Actual Price - Predicted Price.

Modified Duration assumes that the price changes are linear with respect to changes in the yield to maturity. From Exhibit 3, the true relationship between the bond's price and the yield to maturity is not linear. The Column with the differences is always positive and increases

---

4 If the original compounding basis on the bond was semi-annual, the Modified Duration must first be calculated on a semi-annual basis and then annualized. You can not use the annual Macaulay Duration to calculate the Modified Duration.
as we move away from a yield to maturity of 10%. The actual relationship between the bond price and the yield to maturity is shown in Figure 2.

The curved line is the actual price curve. The straight line is the price relationship using Modified Duration. Everywhere the actual price curve is above the Modified Duration relationship. This is exactly what we saw in Exhibit 3. The difference was always positive, i.e., actual calculated price was greater than the new price using the Modified Duration relationship. In addition, the percentage changes in price are not symmetric. The percentage decrease in price for a given increase in yield is always less than the percent increase for the same decrease in yield. This property is referred to as convexity. Note that the two prices are quite close for small changes in the yield to maturity but the difference grows as the change in yield to maturity becomes bigger.

**Convexity.**

From Figure 2 it is clear that the Modified Duration relationship does not fully capture the true relationship between bond prices and yield to maturity. In order to more fully capture this, practitioners use Convexity. The definition of Convexity is

$$
\text{Convexity} = CV = \frac{1}{P} \left( \frac{\Delta P}{\Delta y} \right)^2
$$

Once again those with a math background will recognize the last term on the right as the second derivative of price with respect to yield to maturity. The actual definition of Convexity that we can use is
Exhibit 4 shows the calculation of the semi-annual convexity for the six-year 6.1% coupon bond.

### Exhibit 4

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$30.50</td>
<td>0.9524</td>
<td>$29.05</td>
<td>3.51%</td>
</tr>
<tr>
<td>2</td>
<td>30.50</td>
<td>0.9070</td>
<td>27.66</td>
<td>3.34%</td>
</tr>
<tr>
<td>3</td>
<td>30.50</td>
<td>0.8638</td>
<td>26.35</td>
<td>3.19%</td>
</tr>
<tr>
<td>4</td>
<td>30.50</td>
<td>0.8227</td>
<td>25.09</td>
<td>3.03%</td>
</tr>
<tr>
<td>5</td>
<td>30.50</td>
<td>0.7835</td>
<td>23.90</td>
<td>2.89%</td>
</tr>
<tr>
<td>6</td>
<td>30.50</td>
<td>0.7462</td>
<td>22.76</td>
<td>2.75%</td>
</tr>
<tr>
<td>7</td>
<td>30.50</td>
<td>0.7107</td>
<td>21.68</td>
<td>2.62%</td>
</tr>
<tr>
<td>8</td>
<td>30.50</td>
<td>0.6768</td>
<td>20.64</td>
<td>2.50%</td>
</tr>
<tr>
<td>9</td>
<td>30.50</td>
<td>0.6446</td>
<td>19.66</td>
<td>2.38%</td>
</tr>
<tr>
<td>10</td>
<td>30.50</td>
<td>0.6139</td>
<td>18.72</td>
<td>2.26%</td>
</tr>
<tr>
<td>11</td>
<td>30.50</td>
<td>0.5847</td>
<td>17.83</td>
<td>2.16%</td>
</tr>
<tr>
<td>12</td>
<td>1,030.50</td>
<td>0.5568</td>
<td>573.82</td>
<td>69.37%</td>
</tr>
</tbody>
</table>

Bond Value $827.17  Semi-annual Convexity 110.88

We can annualize the semi-annual convexity of 110.88 by dividing\(^5\) it by 2\(^2\) or 4. Here it would be 27.72. Convexity is useful to practitioners in a number of ways. First it can be used in conjunction with duration to get a more accurate estimate of the percentage price change resulting from a change in the yield. The formula\(^6\) is

\[
% \Delta \text{Price} = - \text{Modified Duration} \times \Delta y + \frac{1}{2} \times \text{Convexity} \times (\Delta y)^2
\]

\(^5\) Convexity is annualized by dividing the calculated Convexity by the number of payments per year squared.

\(^6\) For those with a math bent, this formula is based on using a Taylor series expansion to approximate the value of the percentage change in price.
Adding the convexity adjustment corrects for the fact that Modified Duration understates the true bond price. For example, in Exhibit 3, at a yield of 12% the percentage price change using only Modified Duration was -9.54%, while the actual was -9.01%. If we use the Convexity value we just calculated, the predicted percentage price change would be

\[ \% \Delta \text{Price} = -4.77 \times (.02) + \frac{1}{2} \times 27.72 \times (.02)^2 = -0.0954 + 0.00554 = -0.0899. \]

This is -8.99%, which is much closer to the actual percentage price change of -9.01%.

The pricing aspect of Convexity is much less important now since most people have access to calculators and computers that can do the pricing. The more important use of convexity is that it provides insight into how a bond will react to yield changes. Again from Exhibit 3 and Figure 2 we see that the price reaction to changes in yield is not symmetric. For a given change in yield, bond prices drop less for a given increase in yield and increase more for the same decreases in yield. The downside is less and the upside is more. This is clearly a desirable property. The higher the Convexity of a bond the more this is true. Thus, bonds with high convexity are more desirable.

Summary

Each of the price sensitivity measures discussed in this note is part of the everyday language and thinking of fixed income investors. They are relative risk measures that help investment professionals think about the risks they face.
NON-QUANTITATIVE MEASURES IN COMPANY EVALUATION

ÁGNES HORVÁTH
Institute of Business Sciences, University of Miskolc
3515 Miskolc-Egyetemváros, Hungary
vgthagi@gold.uni-miskolc.hu

Field of research: information need of enterprises

Abstract: In order to explore the information needed for evaluating companies both factors affecting the value of a company and sources of information needed for this purpose have to be considered. The usual documents that are expected to be examined are the accounting documents (balance sheets, income statements, annual reports, cash flows, etc.). Several questions arise regarding the accuracy of figures provided in these documents and reflecting the past activity of the company, i.e. whether these figures are realisable enough to give a real overview of the value and performance of a company or whether to start the evaluation on the basis of financial data or whether to find out which subjective and hardly measurable factors affect the company value. This study provides criticism of the accounting information, the most popular evaluation methods and their criticism, non-quantitative factors determining company value (value drivers), and accentuated role of human factors (work force and organisation, loyalty, business relationships) in company evaluation.

Criticism of the accounting information

In the past few decades market evaluations of companies have become more and more independent from accounting data. Book values of companies do not play as important role in this process as they used to, they are usually lower than market value. The accounting approach greatly differs from the market one as the former is past-oriented and the company value is highly earnings-related. In the evaluation of the real value of a company it is not only its earnings-related abilities measured in money that should be taken into account.

Several researchers highlight the distorting factors of accounting systems and the deficiencies of financial indicators taken from accounting documents:

- Ehrbar (2000) considers R+D expenditures to be costs that are to be met and lays stress on goodwill. According to him all expenses that contribute to the future income have to be capitalised.
- The Stern Steward&Co. Consultancy (2002) reported more than 120 potential distortions in the GAAP. With deficiencies in the internal accounting processes the number of corrections reached 160 cases. They divided the most important modifications into 8 groups: accounting of the R+D, strategic investments and acquisitions, recognition of the expenditures, depreciations, expenses spent on reorganisation and restructure, taxes and modifications in the balance sheets. On the basis of their experience they drew a conclusion and said that it is enough to
make about 15 modifications in a company to get accurate values. While selecting the 15 most important modifications it is important to bear in mind the features of companies (Dorgai 2003).

- Black and his co-authors (2001) lay emphasis on differences in accounting statements issued for stakeholders, management and investors.
- Rappaport (2002) puts emphasis on the deficiencies in profit indicators, accounting returns on investments indicators (ROI) and accounting returns on equity indicators (ROE). The most serious problem is that these indicators do not count with capital requirements and time value of money. Drawbacks of these indicators are also examined by other researchers, among others Katits (2002), Dorgai (2003), Brealey-Myers (1999).
- Copeland-Koller-Murrin (1999) thinks that in the accounting approach it is the accounting profit of companies that counts. The shortcoming of this approach is that it does not consider the investments needed to create the profit and neglects timing. They suggest working out a refined accounting model.
- According to Wimmer (2004) the most serious problem is that accounting gives only subsequent screening and does not contribute to the decision-making process at all. Information is aggregated on unsuitable structure. This information is not reliable enough to provide analysis of factors affecting the profit. Dorgai (2003) shares the same ideas.

The most popular valuation methods and their criticism

There are several evaluation methods, but some of them are given preferences. As all methods have more or less deficiencies, applying them without selection and criticism may result in serious problems. Different accounting systems in different countries apply different categories. So does Hungary. In some cases it is almost impossible to find the required data on the basis of the method used in the Hungarian accounting system either. Another problem is that these methods can usually be applied for evaluating large companies and not SMEs. They are not suitable for evaluation of SMEs in spite of all the efforts made in this respect.

Discounted cash flow

Discounted cash flow is the most popular evaluation method. According to this method the value of a company is the total earnings (in cash) that a company realises in its business activity during its operation in the long run. The value of a company is a discounted value of its cash flows expected in the future.

Shortcomings of the discounted cash flow method:

- It can only be applied successfully when a company operates in a stable environment and is in the maturity of its life cycle. Cash flows for the next year can be forecasted more or less exactly. In the introduction phase forecasting has no real basis.
- In a dynamic environment or in the period of launching a new branch of industry it is impossible to determine the potential revenue or the free cash flow.
Most companies make losses in their first year of operation, thus their cash flows are unfavourable.

Due to deficiencies of stock exchange data it is difficult to determine the WACC.

Valuation with multiplication indicators
This type of valuation is based on comparison of company indicators. The most widely used indicators are Price/Earnings (P/E), Price/Sales (P/S), P/EBIT, P/EBITDA. In the case of listed companies these indicators are easy to get access to. As for values of companies not listed on the Stock Exchange, average indicators of listed companies and the ones operating in the same sectors are taken into account, as their indicators are easy to compute.

Shortcomings of this method are as follows:
- Multiplication evaluation based on figures to be compared is not possible to make as the company starts its operation on a completely new business model and there are no companies with a similar profile on the market.
- This method can be applied on a developed capital market where there are a lot of companies the average indicators of which have been computed for a long time and can be used as multiplicators.

Evaluation based on economic value added (EVA)
The Economic Value Added is Net Operating Profit Less Adjusted Taxes reduced by (Invested Capital*Cost of Capital). The main advantage of EVA is that it takes into account opportunity costs of capital. The main essence of this method is as follows: when a certain amount of capital is invested with a particular aim, we lose the returns we could have realised and invested in something else at the time of investment. The revenues exceeding the expenditures do not really meet the conditions of profitability, as a very important factor namely the fact that the cost of the invested capital has to meet the return objectives is not taken into account.

It was the Stern Steward&Co. Consultancy that registered the EVA method. Both the Eva and the DCF methods have become main methods in monitoring the creation of shareholder value. EVA is a measure of the value a company has created during a particular period of time. This method provides similar results to the ones computed by DCF. The drawbacks of this method arise from accounting.

It is worth studying another approach that deals with company valuation. A question arises: Whose interests should be taken into account when value is created? There are three theories related to this issue.

Shareholder value theory
The supporters of this theory think that it is enough to set an objective to maximise the shareholder value as this objective can be achieved through the interests of other stakeholders. The condition of this is to ensure long-term satisfaction of consumers, employees, suppliers and creditors.
Stakeholder theory

In the centre of this theory lies a collective interest of all stakeholders. The satisfaction of shareholders is only a required condition. The main objective of a company is to reach the stakeholders’ goals.

Double value creation

Chikán (2003) also deals with this theory. According to his theory of double value creation companies target double value creation. On the one hand, a company creates its consumers value by producing its own products and providing its own services (the product has to meet customer satisfaction and expectation), and on the other hand, it creates its shareholders value by selling its products and services.

In spite of having different approaches to the same issues both theories have come to the conclusion that company value can be increased only if the interests of both shareholders and stakeholders are taken into account.

Value drivers

It has been justified that when company value is determined, several factors have to be taken into consideration besides the potential revenue creating ability of a company.

Besides current investments such factors as expected potential investments, cash flows and opportunities for growth also create values. A company can be valuable not only because it possesses assets which will produce free cash flow sooner or later, but because it can obtain them in the future as well (Damodaran 2001).

Juhász (2004) classified factors leading to differences between book value and market value into categories (see Figure 1). He takes book value as a starting point in determining company value.

- The first category includes the replacement value of assets. He starts out from the supposition that obtaining the assets recorded in the balance sheet costs more than their book value is.
- The second category contains the assets that can be sold independently and are not shown in the accounting statements, but there is no doubt that they represent value for the company, for example: its own brand name, its secret manufacturing process.
- The third category includes synergic effects of resource-combinations. This is the added value of a company, for example: management, employees, organisation and knowledge.
I would like to summarise factors determining the company value in Figure 2. I started out from the fact that a company does not exist in complete isolation. There are several elements in the environment that contribute to its operation. It is essential to focus not only on factors closely related to the company operation, but on micro and macro factors as well.
Macro environmental factors

Features of the country (on the basis of the PEST analysis):
- Political stability, risk of country,
- Regional position, technological development,
- Economic situation, cultural environment, state of development of capital market,
- Macroeconomic forecasts, macroeconomic trends, legal environment

Micro environmental factors

Factors in strong connection with the company

Creating value of functional area:
- Trademarks, brand names, patents, certificates,
- Commitment for the quality, developed channel of distribution,
- Special manufacturing processes, available production factors, and its price, business secrets

Market factors
- Market size, size of market segment, (relative) market share, market growth rate, calculability, product life cycle

Human resource – organisation:
- Intellectual capital, knowledge, culture, motivation, encouragement,
- Team-working, creativity of organisation, key person, experience

Loyalty of employees

Loyalty of customers and investors

Business relationships:
- Customer and buyer relationship, relation with partners and authorities

Factors in connection with the industry branch:
- Life cycle, outlooks, opportunities, driving forces, core competences, results factors, attractive force of branch of industry, restrictions to enter into the market, competition in the branch, connection between sales and welfare

Figure 2. Factors determining company value
Accentuated role of human factors

In this paper I suppose that human elements do not include only the system of human resources of a company, but also all the factors that are human-related. This paper deals with the three most important elements namely the work force with its organisation factors, the loyalty that can be strong among employees, customers and investors and finally the business relationships, which affect the daily operation of a company and largely contribute to company value.

Work force and organisation

There is a lot of contradiction in human resources. On the one hand, labour force with its wages and contributions is a cost factor; on the other hand, it is the only factor that can influence its own performance. Employees make their knowledge and abilities available for the company and in return expect encouragement, motivation, financial security and etc.

Value drivers related to human resources and examined can be as follows:

- Motivation and satisfaction leading to commitment and loyalty, which affect operation efficiency, company performance and finally market evaluation of a company.
- Knowledge, skills and abilities of employees. Knowledge has different forms. Professional knowledge belongs to the basic category and can be obtained on the basis of certain qualifications by specific trainings. This can be completed by basic education, specific literacy and national as well as regional knowledge. It is the knowledge that companies expect from their employees. Apart from this an employee can acquire sector, market, company-specific and working knowledge. All these forms of knowledge can become stronger with time and work experience. Their overall effect on value creation makes up the value of human resources. The importance of the obtained knowledge differs and depends on the industrial sector, company and even position taken. All these factors should be taken into account while measuring the values of human resources.
- Relationship capital. This is the personal relationship of employees with customers, suppliers, authorities, financial institutions, investors, business partners and so on.
- Key persons: They are the people who have the necessary knowledge, the required skills and the special business relationship with suppliers and customers and who the other employees are loyal to.
- Value creation effects of the organisation include organisational structure, organisational creativity and atmosphere at work.

The things necessary for creating values of human resources are as follows: strong culture, qualified workforce, investments into employees, division of information, fair allowances and motivation systems, good management, good working conditions and effective organisation of work.
Loyalty

According to Reichheld and Teal (1996) loyalty creates value, which includes loyalty of employees, partners and customers as well as investors. Their model is shown in Figure 3.

Figure 3. The loyalty effect by Reichheld and Teal (Reichheld-Teal 1996)

Consumer satisfaction is determined by the experienced organisation image, consumer expectations, experienced quality and the experienced value. According to this approach quality results in satisfaction that leads to loyalty. However it is difficult to measure the variables of the model or they cannot be measured all. Researchers developed some measurable indicators in order to operationalise the latent variables:

- Consumer satisfaction can be measured by asking such questions as: 'How content are you with the company? Does it meet your expectations and to what extent? How close do you think this company is to an ideal? (The answers are marked and a weighted average is computed (customers’ satisfaction index))

- Consumer loyalty is measured on the basis of such indicators as intent or willingness to repurchase, commitment (purchasing different products from the same company), price sensitivity, offering the product or company to other buyers.

In her studies Hetesi (2003) notes that there are arguments as for the clarity of the chain of quality-satisfaction-loyalty-profitability. In order to achieve profitability there is a need for good quality, consumer satisfaction and loyalty, but they are not enough to ensure profitability (Némethné 2000).

Business relationships
It is not simple to measure the value of business relationships. Economic, social and time dimensions play an important role in this. Relationships of people concerned have different
values. Mandják-Simon-Lantos (2004) conducted extensive interviews in ten companies with different profiles, sizes, market positions, and owners and wanted to get an answer to the question what practising professionals thought of values of business relationships and what the value depended on. They examined both consumer and supplier relationships. According to them value of the relationship depends on the value of its management, strategy and operative decision-making. The evaluation of relationship depends on whether the relationship is considered important or less important.

The authors came to the conclusions that respect, recurring business opportunities, sales revenues, reduction of commercial expenses, sustainability of the relationship, contribution to capacity utilisation, risk level, references given to others, number of references, the role of developed routines and the security of supply contribute to relationship-based sources of value. In certain sectors of industry (service and consultancy) special attention is laid on personal relationships because there is competition in relationship as well.

In order to get a clear picture about the value of an enterprise it is worth collecting information on the above-mentioned factors and analysing their effect on value creation.

**System measuring performance is a source of information**

Performance evaluation of a company also provides important information for company evaluation if the method applied by the company is known. Wimmer (2004) elaborated the analysis framework for performance evaluation of a company in Figure 5.

![Figure 5. Characteristical features of evaluation of company performance – analysis framework (Wimmer 2000)](image)

Wimmer was interested how performance measurements could really serve value creating processes. On the basis of her model the most important moments in value creation processes are as follows:

- Performance measurements should provide information supporting decision-making and ensure feedback.
• It is important to know what sort of information the company regularly collects and on what. The source of information (internal or external), its character (objective or subjective), harmonization of different methods used for analysis, the experienced importance (usefulness), coherency (whether it is important, less important or worth monitoring, why), their conformity with strategy and objectives are very essential indicators.

Methods applied

As it has already been mentioned in the criticism of accounting information in the case of accounting approach the financial information is given preference. There is a great need for methods on the basis of which not only financial indicators, but also the relationship between financial and operational performances can be analysed. Comprehensive and complex strategical performance measuring systems could serve this purpose. The methods are as follows: performance prism, Scandia Navigator, Balanced Scorecard, Shareholder value net and etc.

References


Vállalatértékelés pontosan, megbízhatóan, Verlag Dashöfer Kft. Gazdasági és Jogi Szakkiadó

股票基本分析——财务操纵案例

一、财务操纵的特殊会计环境

与其他国家和地区的上市公司相比，我国上市公司所处的历史条件和经济环境以及所受政策约束有较大差异，特别是在实行了几十年计划经济体制后，目前正处于向市场经济体制转化的过渡时期。在这一特定的历史时期中，许多市场行为带有明显计划经济的痕迹。例如股份公司发行新股和股票上市实行计划额度制，每年由国家计委制定，然后再按类型和行政隶属关系分配到各省、市、自治区及国务院各部委。对于拥有数万亿元资产的全国国有企业而言，每年一百亿元的新股额度无异于杯水车薪，能够争到新股额度，企业自然十分珍惜。在新股发行数量是常数的情况下，要利用这难得的机会募集到更多的资本，只有尽量提高新股发行价格这一变量。新股发行价格受到发行市盈率的限制，一般在 20 倍左右。由于《公司法》规定，股份有限公司向社会公众发行的股票不得低于公司总股本的 25%，因此大型国企在新股额度有限的情况下，只能将原有资产中的一部分剥离出来折合成发起人股。

这部分剥离出来的资产历史上属于原来的会计实体，按会计实体假设将全部资产进行确认和计量，按会计期间假设把一部分资产费用化并与营业收入进行配比以确定利润。

在发行新股前，为了将社会公众揭示这部分剥离资产的盈利能力，会计师不得不将这部分剥离资产假设为一个新的虚拟的会计实体，并且假设其已经存在了三个或三个以上会计期间。然后根据历史资料，从原来会计实体中剥离出一部分营业收入和费用归虚拟会计实体，并据以确定该虚拟会计实体在各个会计期间的利润。按照会计常识，只要是可以辨认的资产都能够从总资产中剥离出来单独计价，但是总资产的盈利能力却不是各单项资产盈利能力的简单相加。这种从总资产中剥离部分资产，并模拟计算其营业收入和费用，再据以确定该部分剥离资产产生的利润的方法，不仅违背了会计实体和会计期间的基本假设，而且给股份有限公司上市前的财务包装提高了许多机会。1997 年新上市的公司，其招股书披露的前三年净资产收益率普遍在 40% 以上，个别公司年度的净资产收益率甚至高达 100% 以上。形成鲜明对照的是，同期全国国有企业的净资产收益率平均值不足 7%。由于上市前的过度包装，导致上市后公司必须在提高净利润和降低净资产两个方面进行利润操纵，以使上市后的净资产收益率不会比上市前陡然降低。

二、利润操纵的案例分析

1. 提前确认营业收入的案例分析

提前确认营业收入的情况多见于房地产业上市公司或上市公司控股的房地产业子公司。

L 公司被出具的保留意见

该意见称：“1994 年销售英达花园以售楼合同金额及其相应的成本入帐，与现行房地产开发企业财务制度对销售收入确认的规定不相一致，其销售收入 129 171 827.85 元，及相应成本人民币 87 762 114 元列示于后附的合并会计报表中。”

点评：根据现行房地产开发企业财务制度规定，房地产的销售应在办理相应的产权移交手续，开具发票或结算单后方能确认为销售收入。房地产的开发周期往往需要几年，按照确认营业收入的会计理论，房地产企业在预售房屋或签订售楼合同后，按工程进度确认销售收入和与之相对应的销售成本也不无道理。现行制度的规定过于严厉，但 L 公司以售楼合同金额确认为当年销售收入的做法似乎也不够谨慎，毕竟合同义务的履行还刚刚开始，如果
将售楼合同金额按施工进度分期确认为销售收入好象还说得过去。至少从推迟交纳所得税的角度来说，按
现行制度在产权移交并开具发票后确认销售收入似乎也是对公司有利的。

以 L 公司为例，1994 年以英达花园售楼合同金额及其相应的成本入帐，导致利润增加 41 955 713.85
元（＝129 717 827.85－87 762 114），按 L 公司适用所得税率 15% 计，1994 年将为此多缴纳所得税 620
万元。

L 公司显然不会不知道推迟缴纳所得税百万元的好处，并且甘冒被注册会计师出具保留意见的风险，
L 公司坚持按售楼合同金额确认 1994 年的销售收入一定有难言的苦衷。

**H 公司被出具的保留意见**

该意见称："经查，贵公司于 1995 年将业已出租给某公司之物业出售给另一公司，1995 年获销售收入
4 100 万元。"

**点评**：未等出租物业到期，便匆匆忙忙出售给另一公司，销售收入 4100 万元占当年营业收入 10 013.6 万元
的 41% ，H 公司之匠心由此可见一斑。

**J 公司被出具的说明意见**

该意见称："贵公司将南澳县国土局出让给贵公司的部分土地使用权转让给深圳辉创实业股份有限公
司、深圳市新鸿泰投资发展有限公司和南澳县国土开发总公司。根据深圳市华商律师事务所于 1995 年 4 月
18 日出具的法律意见书，贵公司尚应完善有关的法律手续。"

**点评**：J 公司将 1995 年 4 月 18 日尚未办完法律手续的转让收入计入 1994 年损益表，是否有点太早？

**T 公司被出具的保留意见**

该意见称："贵公司 1996 年度的主营业务收入中计有 4 106 508.00 元人民币及相应的主营业务额 453
901.04 元人民币，系贵公司在 1996 年 12 月 31 日之前开具销售发票，而于 1997 年 1 月 15 日之前办理产品
出库手续。"

**点评**：销售实际在 1996 年，产品出库却在 1997 年，时间虽相差只有 15 天，却跨越两个会计年度，看来无需注解，已能知道其中奥秘。

**X 公司被出具的保留意见**

该意见称："1995 年贵公司按债务重整方案，以拥有的在建楼宇华乐大厦中的部分产权抵偿；人民币 30 612
839.60 元，抵偿所欠中国建设银行深圳市分行人民币 166 585 723.22 元的债务，由此产生利润人民币
135 972 883.62 元。后又向建设银行深圳市分行以人民币 166 585 723.22 元购回相同产权。我们认为上述业
务的会计处理和中国现行会计制度的有关规定不一致。鉴于上述情况，我们对后附会计报表的比较数字，
即相关的经营所得的利润、所得税、固定资产、可分配利润、股东权益及 1996 年度的固定资产、可分配利
润、股东权益不能确定。"

**点评**：X 公司保留意见中所披露的问题过于复杂，我们作一解释：
1 X 公司拥有在建楼宇华乐大厦的部分产权，该产权的帐面价值为 30 612 839.60 元。
2 X 公司欠建设银行深圳市分行的债务为 166 585 723.22 元。
3 X 公司以帐面价值仅 30 612 839.60 元的资产抵偿 166 585 723.22 元的债务，并将差额 135 972 883.62
元作为利润入帐。
4 X 公司后来又以 166 585 723.22 元的现金向建设银行深圳市分行购回上述产权，并将上述产权列作 X
公司固定资产。
5 X 公司的会计师十分聪明，他首先创造了一笔交易：以一项资产抵偿 5 倍于该项资产帐面价值的债务，
交易的结果当然是公司获得了 1.35 亿元的利润。然后他又制造了一笔交易：以相当于原来所欠债务的金额向债权人买回抵债债务的那项资产，交易的结果当然是债权人全部收回了借款。最后他又以 1.66 亿元的价值将原来只值 3000 多万元的在建楼宇的产权作为 X 公司的固定资产入帐。上述交易的结果是债权人和债务人皆大欢喜：债权人如数收回全部借款，而债务人则获得了 1.35 亿元的账面利润。本来一项很简单的偿还欠款的交易，经过精心包装后竟然会产生巨额利润，堪称财务包装的杰作。也许是会计师自己也觉得这样做太过分，于是 1996 年（即次年）又将原值 1.66 亿元的固定资产调低为 1.33 亿元。

2. 推迟确认本期费用

与提前确认营业收入一样，推迟确认本期费用同样具有增加本期利润的作用。

S 公司被出具的说明意见

该意见称：“百大宾馆已经使用的空调系统 2 014 914.29 元和客房、餐厅装潢费 375 164.37 元，仍列在建工程。”

点评：按现行会计制度规定，在建工程完工后应立即转入“固定资产”账户，并从转入之日起计提固定资产折旧。S 公司的做法显然是推迟确认本期费用，并导致本期利润相应增加。

B 公司被出具的保留意见

该意见称：“1995 年二电炉分厂因设备故障而停产发生的费用 8 194 625.06 元，历史遗留的工程设备大修理支出 6 282 661.75 元，列入待摊费用和其他应收款，未计入当年损益。”

点评：上述两项费用如在本期确认，将使税前利润因此减少 14 477 286.81 元。B 公司 1995 年利润总额为 6375 万元，上述两项费用推迟确认致使 1995 年利润增加 1447 万元，占利润总额的 22.7%。

Z 公司被出具的保留意见

该意见称：“1995 年末，公司资产负债表的存货中，在产品成本为 174 783 884.31 元，其中包括在产品定额成本差异 52 515 657.52 元。此项差异应在期末在产品、库存产成品和本期销售产品之间进行分配。其中，本期销售产品成本应负担此项差异额为 24 342 287.21 元，贵公司并未就该项差异额作出记录和反映。我们认为，除存在上述因转转在产品定额成本差异影响本期销售成本外 ...”


E 公司被出具的保留意见

该意见称：“公司 1995 年度利息净支出较长短期借款金额及合同规定的利率计算的利息要少 660 万元。”

点评：经检索，E 公司 1995 年度利润总额为 858.24 万元，少计利息费用而导致的利润虚增 660 万元，占当年利润总额的 77%。
N 公司被出具的保留意见

该意见称：“贵公司在待摊费用、递延资产两个科目的使用及其摊销上不够规范，不符合有关会计制度的规定，影响了经营成果。”

点评：据检索：公司被出具具有保留意见审计报告的会计年度，利润总额 346.82 万元，仅为上年的 15%。

X 公司被出具的保留意见

该意见称：“贵公司本年度调整以前年度已售罄的怡都大厦成本及华乐大厦在建成本，因此调整以前年度损益计人民币 10 255 760.57 元及在建工程成本 10 255 760.57 元。因华乐大厦在建成本尚未核定，故我们不能确认办理完毕。因此，我们对后附利润表中的销售收入人民币 87 336 000.00 元，成本人民币 32 756 059.28 元，及净利润 54 579 940.72 元予以保留。”


3. 潜亏挂帐的案例分析

SH 公司被出具的保留意见

SH 公司与三家房地产开发公司发生房地产纠纷，法院终审判决 SH 公司败诉，SH 公司为此须向三家房地产开发公司赔偿 2786 万元。该被出具的保留意见称：“公司未将上述终审判结果计入当期损益，计影响当年利润 2786 万元。”

点评：法院终审判决具有法律效力，公司纵然表示不服，也不能因此而拒不执行终审判决，即使是出于谨慎考虑，亦应将损失计入当年利润表。SH 公司 1996 年净亏损 7845.60 万元，每股收益-0.516 元。

G 公司被出具的保留意见

该意见称：“贵公司会计报表还存在以下问题，待处理流动资产净损失期末余额 4.93 亿元，系分别从
存货和应收帐款转入，是一笔需要确认的亏损。由于核算口径的问题，这笔需要确认的亏损应归属1996年，哪些应归属以前年度无法确认。


SL公司被出具的保留意见

该意见称：“我们注意到，贵公司三江下属联营企业本年出现大量亏损共计达人民币103,560,678.05元，按持股比例，贵公司应承担损失人民币51,780,339.03元。因持股比例不超过50%，所以这三家公司的投资收益按成本法核算，也未纳入合并范围。”

点评：根据上述SL公司应承担的亏损测算，SL公司在这三家子公司的持股比例正好是50%。从理论上说，持股50%并非绝对控股，用成本法核算长期投资也勉强说得过去。但是以稳健原则而言，被投资企业发生如此巨额的亏损投资方应调整长期投资的帐面价值，也不确认发生的巨额亏损，似乎就说不过去了。SL公司1996年利润总额为1,216.62万元，SL公司应承担的投资亏损高达5,178万元，是利润总额的4倍以上，根据重要性原则，也应当在年度报告中有所披露。

HD公司被出具的保留意见

该意见称：“本报告所附合并资产负债表存货余额21,219,989.25元中包括由于种种原因以借条形式反映的存货余额3,188,691.69元，由公司清理出的存货计2,588,295.60元，两项金额合计5,246,987.29元，占存货总额的比例为27%，其中可能包括潜在的损失，尚无法量化。在公司合并报表内有未抵冲内部往来，本报告所附合并资产负债表内有应收款内尚有382.90万元借方余额未抵冲，有潜在的损失需要清理。”

点评：根据检索，HD公司被出具保留意见年度（1995年）利润总额为620.74万元，而潜在损失竟高达957万元。如果潜在损失当年计入损益的话，HD公司将进入亏损企业行列。HD公司1996年的利润总额为743.03万元，基于同样的原因，HD公司仍未将潜在损失计入当年损益，并再次被注册会计师出具保留意见。

4. 会计方法变更或会计处理错误的案例分析

F公司被出具的保留意见

该意见称：“我们认为，除公司的机器设备折旧的计提方法由平均年限法改为工作量法外，其他会计处理方法的选用遵循了一贯性原则。”

点评：F公司是一家生产建筑材料的公司，由于房地产市场不景气导致建材销售疲软，被出具保留意见年度净利润仅为上一年度的50%，由于开工不足，机器设备按工作量法提取的折旧，当然会小于按平均年限法提取的折旧，利润自然会相应增加。

SN公司被出具的保留意见

该意见称：“根据我们的审查，贵公司从1996年6月1日起采用了经董事会批准的变更后的固定资产折旧政策，由于此项政策的变更，导致评估表中的累计折旧计数减少人民币15,269,172元，同时对合并报表之税前利润产生影响，影响数为人民币10,674,824元（扣除少数股东权益影响数人民币4,494,348元）。”

点评：根据SN公司1996年年报披露：公司自1996年6月1日起，将房屋建筑物的折旧年限由20年改为
40 年，将机器设备的折旧年限由 10 年改为 20 年，运输工具的折旧年限由 5 年改为 10 年。由于资料有限，我们无法判断上述折旧政策的变更是否合理，但原先估计的固定资产使用年限与变更后的折旧年限竟然会相差一倍，不能不令人感到吃惊。

L 公司被出具的保留意见

该意见称："如附注 32 所述，贵公司本年度根据深圳光大木材工业有限公司各方股东认可的因股权变更（新增参股股东）而进行的资产评估报告所确认 1995 年末该公司净资产（此评估结果未进行财务处理），按持股比例调增权益人民币 13 601 500.00 元。而按成本法核算，贵公司多计权益为人民币 6 874 014.81 元。"


ZF 公司被出具的保留意见

该意见称：“公司 1995 年 12 月 19 日收购康恩贝制药公司包括收购日前的全年净利润，又按权益法列为‘投资收益’21 522 234.38 元。”


5. 帐证不符或帐实不符的案例分析

AD 公司被出具的保留意见

该意见称：“贵公司联营公司 AD 集装箱有限公司和 AD 货运有限公司有计折合 134 187 530.88 元人民币的应收帐款，AD 船务有限公司有计折合 2 543 505.78 元人民币的其他应收款，AD 石油化工实业有限公司有计折合 16 600 000 元人民币的其他应收款。由于我们对债务人发函询证未能收到回函，故我们对此等应收款项可回收性难以确认。”

点评：另据 AD 公司 1995 年年报披露，AD 公司对 AD 集装箱有限公司、AD 贸运有限公司和 AD 船务有限公司均持有 75%股权，对 AD 石油化工实业有限公司持有 63%股权，按权益法并根据 AD 公司对上述公司的持股比例核算长期投资，AD 公司长期投资将有 11500 万元的投资损失，对于当时注册资本仅 1 亿多元的 AD 公司而言，如此巨额的损失显然是无法承受的，因此此等难以确认可回收性的应收款项只能挂在帐
上。然而令人遗憾的是，直至 AD 公司 1995 年报公布日为止，AD 公司的控股子公司并未对上述拖欠巨额债务并拒绝承诺偿债的债务人提起诉讼，也未表示将采取何种行动来设法收回上述应收款项。据此人们有理由怀疑：AD 公司控股子公司的上述债权是否真的存在？这些应收款项确认的营业收入是否真的存在？

JX 公司 1995 年和 1996 年被出具的保留意见

该意见称：“贵公司在大亚湾胜景实业发展公司投资 4788 万元（含短期投资 1000 万元），已累计投资收益为 563 万元，并已列入本年度的净利润中。但我们未能获得被投资公司年度会计报表以及有关投资项目的详情。” JX 公司认为其对大亚湾胜景实业发展公司投资，其中本年从其他应收款、短期投资科目转入 3 024 512.92 元，我们未能取得有关投资原始凭证。”

点评：JX 公司对大亚湾胜景实业公司的长期投资按权益法核算，1995 年将未经审计的大亚湾胜景实业公司的利润按投资比例确认为投资收益，1996 年大亚湾胜景实业公司增加持股比例却没有投资原始凭证，这样算出来的投资收益叫人如何相信？

FY 公司被出具的保留意见

该意见称：“贵公司的附属公司 FY 汽车配件有限公司于 1996 年 12 月 31 日的房屋建筑物原值人民币 9 075 000 元部分和相关的其他递延支出原值人民币 27 225 000 元，为无法提供有关的工程结算书，我们无法确认这些资产的实际成本和账务处理上分摊的合理性。”

点评：由于 FY 公司没有提供将上述 3630 万元支出确认为资本性支出的合法凭证，人们如果提出 FY 公司是将本该费用化的支出予以资本化，从而导致 FY 公司当年利润虚增的嫌疑，似乎也是无可指责的。

WS 公司被出具的保留意见

该意见称：“深圳文联饮料包装有限公司由 WS 公司与香港鹏兆发展有限公司共同经营，注册资本为人民币 500 万元，中方出资比例为 48%，外方出资比例为 52%。上海文联公司长期投资中均未反映对该公司的投资。1996 年 6 月 24 日，《深圳特区报》登载了拟申请注销并请各债务人持有关凭证在 30 天内到公司清算组办理手续一事，我们未能取得该公司目前的一切财务会计资料。”

点评：WS 公司作为持有深圳文联饮料包装有限公司 48%股权的大股东，居然从该公司开张之日起就未向注册会计师提供该公司的财务会计资料，令人不得不怀疑该黑洞的存在对 WS 公司历年利润确定的影响。

GX 公司被出具的保留意见

该意见称：“贵公司的产品销售收入和成本审核时发现部分保洁产品缺乏原材料采购过程和生产过程原始凭证，其销售收入计人民币 2 021.65 万元，贵公司未作帐项调整，我们所难以确认，导致相关利润难以确认。”

点评：GX 公司的保洁产品缺乏原材料采购过程和生产过程原始凭证，不知 GX 公司是如何计算产品成本，并与销售收入进行配比以确定本期利润的。

6. 关联交易影响利润的案例分析

SH 公司被出具的说明意见称：“1995 年 12 月 20 日，贵公司（以下简称甲方）与上海建国社会公益基金会（以下简称乙方）签署《法人股转让协议书》。根据甲乙双方共同协商，甲方将所持上海浦东大众法人股 300 万股（每股面值 1 元）转让给乙方，并办理正式股票过户手续，获利 3 354 000.00 元。另甲乙双方
同日签署的《补充协议》规定，如上海浦东大众法人股未能在两年内上市，则甲方应以原转让价加上乙方已付款项按 15% 利率计算的利息之和购回。

点评：上海市建国社会公益基金会是以 SH 公司董事长名字命名的慈善基金，所以上述甲乙双方之间的交易属关联交易。由于中国大陆证券市场中的法人股在数年内并无上市之可能，因此上市甲乙双方签署的协议，实质上是一份以法人股作质押物的抵押贷款。SH 公司通过将法人股出售给关联方，两年后又以 15% 的利率和乙方已经付款按 7.5% 的利率计算的利息之和购回。不仅获得了低息贷款（1995 年银行两年期贷款年利率为 13%），而且在 1995 年度获得了 335 万元的交易收益，并计入当年净利润。这项投资收益为 SH 公司 1995 年业绩作出了巨大贡献，如果没有这笔 335 万元的投资收益，SH 公司 1995 年的净资产收益率将低于 10%，从而失去配股资格。

三、利润操纵的动机分析

1. 在发行市盈率受到限制时为提高发行价格而进行财务包装

尽管没有明文规定，但是 1999 年以前发行市盈率若超过 15 倍一般很难获得证券监管部门的批准。在发行市盈率常量的情况下，要提高每股发行价格惟有在每股收益这个变量上做文章。1996 年以前计算发行市盈率的公式是“发行市盈率 = 每股发行价格 ÷ 发行新股年度预测的每股收益”，于是 1996 年以前发行新股的不少公司多在历史数据上做文章。有的公司在发行新股前把位于闹市区的一块土地出售，使发行新股前一会计年度获得巨额投资收益。有的公司把不能直接产生盈利的资产剥离，以较低的费用与营业收入进行配比而使利润增加。对于大中型企业而言，由于受发行新股额度的限制，只能从原有总资产中剥离一部分资产折股作为发起人股，对这部分资产的盈利能力只能根据历史数据进行模拟。模拟计算的利润无须缴纳所得税，但却是制定发行价格的依据，其结果自然可想而知。

1996 年 8 月，中国证监会对 SD 公司进行通报批评，并剥夺其 3 年内申请配股的资格。SD 公司原注册资本 12000 万元，若按发行市盈率 2.5:1 的比例进行缩股，缩股后注册资本变更为 48 万元，总股本由 12000 万股缩至 4800 万股，每股净资产和每股净收益随之提高 1.5 倍，新股发行价格亦随之大幅提高。令人吃惊的是：对于如此巨大的股本变动，SD 公司在招股说明书和上市公告书中居然只字未提，会计师事务所居然出具了无保留意见的审计报告。

NT 公司于 1994 年 5 月发行股票并上市，公司 1993 年净收益 2305.6 万元，每股收益高达 0.55 元，公司在上市公告书预测 1994 年净利润为 2175 万元。然而 1994 年报公布后却令投资者大吃一惊：1994 年度实现净利润 1738.89 万元，仅为原预测值的 80%。这还不算，1738.89 万元净利润中还要调整前期损益－1455.4 万元，调整后的净利润仅为 283.49 万元，1994 年度每股收益仅为 0.0458 元。NT 公司为了提高新股发行价格，不仅在盈利能力上做文章，而且在历史数据上做文章，所谓的前期损益调整如果在 1993 年列入损益表的话，则 NT 公司是没有资格发行新股的。

SQ 公司下属的酒店实行承包经营，1996 年 3 月 8 日因合同纠纷经法院判决解除承包合同，然而 SQ 公司在 1996 年 3 月 12 日发布的上市公告书中，仍然根据承包合同预测 1996 年公司全年利润中的 25% 将来自酒店。结果是，SQ 公司 1996 年实现净利润仅为预测值的 8.7%。
2. 为获得配股资格而进行财务包装


首先，HT 公司和 LG 公司签署协议，由 LG 公司出资 16 000 万元收购 HT 公司属下的电表公司，收购价格为电表公司净资产账面价值的两倍，收购所需款项的一部分以 LG 公司对 HT 公司的长期负债挂账，3 年后偿还，另一半则以现金支付。

然后，LG 公司匆忙于 1995 年 12 月 22 日召开股东大会批准上述协议，并于 1995 年 12 月 25 日发布收购公告。

按合并报表理论，以现金收购股权方式受让的子公司应采用购受法编制合并报表，即把子公司被收购日以后的净利润列在合并利润表。按照当时的会计制度，于年末收购的子公司一般不纳入合并报表范围，这样做也符合重要性原则。与 LG 公司同在一个交易所上市的 ZF 公司，1995 年合并报表中将子公司购买日以前的净利润列在合并利润表而被注册会计师出具了保留意见。


SY 公司 1995 年度报告被出具的说明意见称：“公司年末帐上有应付福利费贷方余额 409 958.54 元，职工教育基金 285 110.43 元，不结转 1996 年继续使用，而是在年末全数冲减当年成本。” SY 公司被出具说明意见中的数值较小，仅数十万元，本来并不引人注目。然而 SY 公司在 1995 年度报告摘要中未按国家规定披露审计报告全文，将上述说明意见段遗漏，被上海证券交易所强制要求以补充公告的形式披露注册会计师的说明意见。这一点使 SY 公司的盈利和其他收益均不在合并报表范围内，对利润的影响也不过 3%左右，何苦偏要坚持错误的账务处理，还要在年报摘要中隐瞒自己的错误呢？我们不妨对 SY 公司的行为作一番分析。SY 公司 1995 年报表披露的净利润为 10.03%，勉强超过 10%，为使 1995 年度净利润收益率达到配股资格线，SY 公司 1995 年末进行了三项调整：

**调整一**：将应付福利费和职工教育基金的贷方余额冲减当年成本，使利润总额比冲减前增加 695 068.97 元，若按 15%所得税率计算，税后利润可增加 59 万元。但是这一做法似乎有悖常理：应付福利费和职工教育基金系按职工工资总额为基数，按国家规定比例提取，用于职工福利和职工教育，提取后即形成公司对全体职工的负债，在资产负债表中列在流动负债项下。当年未用完的应付福利费和职工教育基金理应结转下年使用，将上述两项基金的贷方余额冲减成本，意味着将对公司职工的债务一笔勾销，显然是对职工利益的侵权。然而 SY 公司在 1995 年报中坚持采用上述错误的账务处理，因为如果不这样做的话，SY 公司 1995 年的净资产收益率仅 9.7%。但是 SY 公司在 1996 年中对上述账务处理错误作了更正，因为公司 1996 年已经获准并实施了配股。

**调整二**：将过去投入子公司徐行药厂但未作价的专有技术，经评估后作价 30 万元，增加对徐行药厂的持股，使持股比例达到 51%。根据 SY 公司的会计政策，对于控股 51%（含 51%）的子公司按权益法核算长期投资并纳入合并报表。于 1995 年年度盈利的徐行药厂当年即按权益法核算长期投资，显然按权益法核算的净利润要高于按成本法核算的结果。

资产收益率作出了贡献。SY 公司的会计师煞费苦心，终于使 1995 年的净资产收益率达到了 10.03%。

3．为避免连续 3 年亏损公司股票被摘牌而进行财务包装

The Efficient Market Hypothesis
and Its Critics

Burton G. Malkiel

Abstract

Revolutions often spawn counterrevolutions and the efficient market hypothesis in finance is no exception. The intellectual dominance of the efficient-market revolution has more been challenged by economists who stress psychological and behavioral elements of stock-price determination and by econometricians who argue that stock returns are, to a considerable extent, predictable. This survey examines the attacks on the efficient-market hypothesis and the relationship between predictability and efficiency. I conclude that our stock markets are more efficient and less predictable than many recent academic papers would have us believe.
A generation ago, the efficient market hypothesis was widely accepted by academic financial economists; for example, see Eugene Fama’s (1970) influential survey article, “Efficient Capital Markets.” It was generally believed that securities markets were extremely efficient in reflecting information about individual stocks and about the stock market as a whole. The accepted view was that when information arises, the news spreads very quickly and is incorporated into the prices of securities without delay. Thus, neither technical analysis, which is the study of past stock prices in an attempt to predict future prices, nor even fundamental analysis, which is the analysis of financial information such as company earnings, asset values, etc., to help investors select “undervalued” stocks, would enable an investor to achieve returns greater than those that could be obtained by holding a randomly selected portfolio of individual stocks with comparable risk.

The efficient market hypothesis is associated with the idea of a “random walk,” which is a term loosely used in the finance literature to characterize a price series where all subsequent price changes represent random departures from previous prices. The logic of the random walk idea is that if the flow of information is unimpeded and information is immediately reflected in stock prices, then tomorrow’s price change will reflect only tomorrow’s news and will be independent of the price changes today. But news is by definition unpredictable and, thus, resulting price changes must be unpredictable and random. As a result, prices fully reflect all known information, and even uninformed investors buying a diversified portfolio at the tableau of prices given by the market will obtain a rate of return as generous as that achieved by the experts.
The way I put it in my book, *A Random Walk Down Wall Street*, first published in 1973, a blindfolded chimpanzee throwing darts at the *Wall Street Journal* could select a portfolio that would do as well as the experts. Of course, the advice was not literally to throw darts but instead to throw a towel over the stock pages – that is, to buy a broad-based index fund that bought and held all the stocks in the market and that charged very low expenses.

By the start of the twenty-first century, the intellectual dominance of the efficient market hypothesis had become far less universal. Many financial economists and statisticians began to believe that stock prices are at least partially predictable. A new breed of economists emphasized psychological and behavioral elements of stock-price determination, and came to believe that future stock prices are somewhat predictable on the basis of past stock price patterns as well as certain “fundamental” valuation metrics. Moreover, many of these economists were even making the far more controversial claim that these predictable patterns enable investors to earn excess risk-adjusted rates of return.

This paper examines the attacks on the efficient market hypothesis and the belief that stock prices are partially predictable. While I make no attempt to present a complete survey of the purported regularities or anomalies in the stock market, I will describe the major statistical findings as well as their behavioral underpinnings, where relevant, and also examine the relationship between predictability and efficiency. I will also describe the major arguments of those who believe that markets are often irrational by analyzing the “crash of 1987,” the “Internet bubble” of the fin de siecle, and other specific irrationalities often mentioned by critics of efficiency. I conclude that our stock markets
are far more efficient and far less predictable than some recent academic papers would have us believe. Moreover, the evidence is overwhelming that whatever anomalous behavior of stock prices may exist, it does not create a portfolio trading opportunity that enables investors to earn extraordinary risk adjusted returns.

At the outset, it is important to make clear what I mean by the term “efficiency”. I will use as a definition of efficient financial markets that they do not allow investors to earn above-average returns without accepting above-average risks. A well-known story tells of a finance professor and a student who come across a $100 bill lying on the ground. As the student stops to pick it up, the professor says, “Don’t bother—if it were really a $100 bill, it wouldn’t be there.” The story well illustrates what financial economists usually mean when they say markets are efficient. Markets can be efficient in this sense even if they sometimes make errors in valuation, as was certainly true during the 1999-early 2000 internet bubble. Markets can be efficient even if many market participants are quite irrational. Markets can be efficient even if stock prices exhibit greater volatility than can apparently be explained by fundamentals such as earnings and dividends. Many of us economists who believe in efficiency do so because we view markets as amazingly successful devices for reflecting new information rapidly and, for the most part, accurately. Above all, we believe that financial markets are efficient because they don’t allow investors to earn above-average risk-adjusted returns. In short, we believe that $100 bills are not lying around for the taking, either by the professional or the amateur investor.

What I do not argue is that the market pricing is always perfect. After the fact, we know that markets have made egregious mistakes as I think occurred during the recent
Internet bubble. Nor do I deny that psychological factors influence securities prices. But I am convinced that Benjamin Graham (1965) was correct in suggesting that while the stock market in the short run may be a voting mechanism, in the long run it is a weighing mechanism. True value will win out in the end. And before the fact, there is no way in which investors can reliably exploit any anomalies or patterns that might exist. I am skeptical that any of the “predictable patterns” that have been documented in the literature were ever sufficiently robust so as to have created profitable investment opportunities and after they have been discovered and publicized, they will certainly not allow investors to earn excess returns.

A Non-Random Walk Down Wall Street

In this section, I review some of the patterns of possible predictability suggested by studies of the behavior of past stock prices.

Short-term Momentum Including Underreaction to New Information

The original empirical work supporting the notion of randomness in stock prices looked at such measures of short-run serial correlations between successive stock-price changes. In general, this work supported the view that the stock market has no memory – the way a stock price behaved in the past is not useful in divining how it will behave in the future; for example, see the survey of articles contained in Cootner (1964). More recent work by Lo and MacKinlay (1999) finds that short-run serial correlations are not zero and that the existence of “too many” successive moves in the same direction enable
them to reject the hypothesis that stock prices behave as random walks. There does seem to be some momentum in short-run stock prices. Moreover, Lo, Mamaysky and Wang (2000) also find, through the use of sophisticated nonparametric statistical techniques that can recognize patterns, some of the stock-price signals used by “technical analysts” such as “head and shoulders” formations and “double bottoms”, may actually have some modest predictive power.

Economists and psychologists in the field of behavioral finance find such short-run momentum to be consistent with psychological feedback mechanisms. Individuals see a stock price rising and are drawn into the market in a kind of “bandwagon effect.” For example, Shiller (2000) describes the rise in the U.S. stock market during the late 1990s as the result of psychological contagion leading to irrational exuberance. The behavioralists offered another explanation for patterns of short-run momentum – a tendency for investors to underreact to new information. If the full impact of an important news announcement is only grasped over a period of time, stock prices will exhibit the positive serial correlation found by investigators. As behavioral finance became more prominent as a branch of the study of financial markets, momentum, as opposed to randomness, seemed reasonable to many investigators.

However, there are several factors that should prevent us from interpreting the empirical results reported above as an indication that markets are inefficient. First, while the stock market may not be a mathematically perfect random walk, it is important to distinguish statistical significance from economic significance. The statistical dependencies giving rise to momentum are extremely small and are not likely to permit investors to realize excess returns. Anyone who pays transactions costs is unlikely to
fashion a trading strategy based on the kinds of momentum found in these studies that will beat a buy-and-hold strategy. Indeed, Odean (1999) suggests that momentum investors do not realize excess returns. Quite the opposite – a sample of such investors suggests that such traders did far worse than buy-and-hold investors even during a period where there was clear statistical evidence of positive momentum. This is so because of the large transactions costs involved in attempting to exploit whatever momentum exists. Similarly, David Lesmond, Michael Schill, and Chunsheng Zhou (2001) find that the transactions costs involved in undertaking standard “relative strength” strategies are not profitable because of the trading costs involved in their execution.

Second, while behavioural hypotheses about bandwagon effects and underreaction to new information may sound plausible enough, the evidence that such effects occur systematically in the stock market is often rather thin. For example, Eugene Fama (1998) surveys the considerable body of empirical work on “event studies” that seeks to determine if stock prices respond efficiently to information. The “events” include such announcements as earnings surprises, stock splits, dividend actions, mergers, new exchange listings, and initial public offerings. Fama finds that apparent underreaction to information is about as common as overreaction, and post-event continuation of abnormal returns is as frequent as post-event reversals. He also shows that many of the return “anomalies” arise only in the context of some very particular model, and that the results tend to disappear when exposed to different models for expected “normal” returns, different methods to adjust for risk, and when different statistical approaches are used to measure them. For example, a study, which gives equal-weight to post-announcement returns of many stocks, can produce different results
from a study that weight the stocks according to their value. Certainly, whatever momentum displayed by stock prices does not appear to offer investors a dependable way to earn abnormal returns.

The key factor is whether any patterns of serial correlation are consistent over time. Momentum strategies, which refer to buying stocks that display positive serial correlation and/or positive relative strength, appeared to produce positive relative returns during some periods of the late 1990s but highly negative relative returns during 2000. It is far from clear that any stock-price patterns are useful for investors in fashioning an investment strategy that will dependably earn excess returns.

Many predictable patterns seem to disappear after they are published in the finance literature. As Schwert (2001) points out, there are two possible explanations for such a pattern. One explanation may be that researchers are always sifting through mountains of financial data. Their normal tendency is to focus on results that challenge perceived wisdom, and every now and again, a combination of a certain sample and a certain technique will produce a statistically significant result that seems to challenge the efficient markets hypothesis. Alternatively, perhaps practitioners learn quickly about any true predictable pattern and exploit it to the extent that it becomes no longer profitable. My own view is that such apparent patterns were never sufficiently large or stable to guarantee consistently superior investment results and certainly such patterns will never be useful for investors after they have received considerable publicity. The so-called January effect, for example, seems to have disappeared soon after it was discovered.
Long-run Return Reversals

In the short-run, when stock returns are measured over periods of days or weeks, the usual argument against market efficiency is that some positive serial correlation exists. But many studies have shown evidence of negative serial correlation – that is, return reversals – over longer holding periods. For example, Fama and French (1988) found that 25 to 40 percent of the variation in long holding period returns can be predicted in terms of a negative correlation with past returns. Similarly, Poterba and Summers (1988) found substantial mean reversion in stock market returns at longer horizons.

Some studies have attributed this forecastability to the tendency of stock market prices to “overreact.” DeBondt and Thaler (1995), for example, argue that investors are subject to waves of optimism and pessimism that cause prices to deviate systematically from their fundamental values and later to exhibit mean reversion. They suggest that such overreaction to past events is consistent with the behavioral decision theory of Kahneman and Tversky (1982), where investors are systematically overconfident in their ability to forecast either future stock prices or future corporate earnings. These findings give some support to investment techniques that rest on a “contrarian” strategy, that is, buying the stocks, or groups of stocks, that have been out of favor for long periods of time and avoiding those stocks that have had large run-ups over the last several years.

There is indeed considerable support for long-run negative serial correlation in stock returns. However, the finding of mean reversion is not uniform across studies and is quite a bit weaker in some periods than it is for other periods. Indeed, the strongest empirical results come from periods including the Great Depression – which may be a
time with patterns that do not generalize well. Moreover, such return reversals for the market as a whole may be quite consistent with the efficient functioning of the market since they could result, in part, from the volatility of interest rates and the tendency of interest rates to be mean reverting. Since stock returns must rise or fall to be competitive with bond returns, there is a tendency when interest rates go up for prices of both bond and stocks to go down, and as interest rates go down for prices of bonds and stocks to go up. If interest rates mean revert over time, this pattern will tend to generate return reversals, or mean reversion, in a way that is quite consistent with the efficient functioning of markets.

Moreover, it may not be possible to profit from the tendency for individual stocks to exhibit patterns of return reversals. Fluck, Malkiel and Quandt (1997) simulated a strategy of buying stocks over a 13-year period during the 1980s and early 1990s that had particularly poor returns over the past three to five years. They found that stocks with very low returns over the past three to five years had higher returns in the next period, and that stocks with very high returns over the past three to five years had lower returns in the next period. Thus, they confirmed the very strong statistical evidence of return reversals. However, they also found that returns in the next period were similar for both groups, so they could not confirm that a contrarian approach would yield higher-than-average returns. There was a statistically strong pattern of return reversal, but not one that implied an inefficiency in the market that would enable investors to make excess returns.

Seasonal and Day-of-the-Week Patterns

A number of researchers have found that January has been a very unusual month for stock market returns. Returns from an equally weighted stock index have tended to
be unusually high during the first two weeks of the year. The return premium has been particularly evident for stocks with relatively small total capitalizations (Keim, 1983). Haugen and Lakonishok (1988) document the high January returns in a book entitled *The Incredible January Effect*. There also appear to be a number of day-of-the-week effects. For example, French (1980) documents significantly higher Monday returns. There appear to be significant differences in average daily returns in countries other than the United States (Hawawini and Keim, 1995). There also appear to be some patterns in returns around the turn of the month (Lakonishok and Smidt, 1988), as well as around holidays (Ariel, 1990).

The general problem with these predictable patterns or anomalies, however, is that they are *not* dependable from period to period. Wall Street traders often joke that now the “January effect” is more likely to occur on the previous Thanksgiving. Moreover, these non-random effects (even if they were dependable) are very small relative to the transactions costs involved in trying to exploit them. They do not appear to offer arbitrage opportunities that would enable investors to make excess risk-adjusted returns.

**Predictable Patterns Based on Valuation Parameters**

Considerable empirical research has been conducted to determine if future stock returns can be predicted on the basis of initial valuation parameters. It is claimed that valuation ratios, such as the price-earnings multiple or the dividend yield of the stock
Predicting Future Returns from Initial Dividend Yields

Formal statistical tests of the ability of dividend yields (that is, dividend-price ratios) to forecast future returns have been conducted by Fama and French (1988) and Campbell and Shiller (1988). Depending on the forecast horizon involved, as much as 40 percent of the variance of future returns for the stock market as a whole can be predicted on the basis of the initial dividend yield of the market index.

An interesting way of presenting the results is shown in the top panel of Exhibit 1. The exhibit was produced by measuring the dividend yield of the broad U.S. stock market the Standard and Poor’s 500 Stock Index each quarter since 1926 and then calculating the market’s subsequent ten-year total return through the year 2000. The observations were then divided into deciles depending upon the level of the initial dividend yield. In general, the exhibit shows that investors have earned a higher rate of return from the stock market when they purchased a market basket of equities with an initial dividend yield that was relatively high, and relatively low future rates of return when stocks were purchased at low dividend yields.

These findings are not necessarily inconsistent with efficiency. Dividend yields of stocks tend to be high when interest rates are high, and they tend to be low when interest rates are low. Consequently, the ability of initial yields to predict returns may simply reflect the adjustment of the stock market to general economic conditions. Moreover, the use of dividend yields to predict future returns has been ineffective since...
the mid-1980s. Dividend yields have been at the three percent level or below continuously since the mid-1980s, indicating very low forecasted returns. In fact, for all 10 year periods from 1985 through 1992 that ended June 30, 2002, realized annual equity returns from the market index have averaged approximately 15 percent. One possible explanation is that the dividend behavior of U.S. corporations may have changed over time (See Bagwell and Shoven, 1989, and Fama and French, 2001). Companies in the twenty-first century may be more likely to institute a share repurchase program rather than increase their dividends. Thus, dividend yield may not be as meaningful as in the past as a useful predictor of future equity returns.

Finally, it is worth noting that this phenomenon does not work consistently with individual stocks, as has been shown by Fluck, Malkiel and Quandt (1997). Investors who simply purchase a portfolio of individual stocks with the highest dividend yields in the market will not earn a particularly high rate of return. One popular implementation of such a “high dividend” strategy in the United States is the “Dogs of the Dow Strategy,” which involves buying the ten stocks in the Dow Jones Industrial Average with the highest dividend yields. For some past periods this strategy handily outpaced the overall average, and so several “Dogs of the Dow” mutual funds were brought to market and aggressively sold to individual investors. Such funds have generally underperformed the market averages during the 1995-99 period.

Predicting Market Returns from Initial Price-earnings Multiples

The same kind of predictability for the market as a whole, as was demonstrated for dividends, has been shown for price-earnings ratios. The data are shown in the
bottom half of Exhibit 1. The exhibit presents a decile analysis similar to that described for dividend yields above. Investors have tended to earn larger long-horizon returns when purchasing the market basket of stocks at relatively low price-earnings multiples. Campbell and Shiller (1998) report that initial P/E ratios explained as much as 40 percent of the variance of future returns. They conclude that equity returns have been predictable in the past to a considerable extent.

Consider, however, the recent experience of investors who have attempted to undertake investment strategies based either on the level of the price-earnings multiple or the dividend yield to predict future long horizon returns. Price-earnings multiples for the Standard & Poor’s 500 stock index rose into the low 20s on June 30, 1987 (suggesting very low long horizon returns). Dividend yields fell below three percent. The average annual total return from the index over the next 10 years was an extraordinarily generous 16.7 percent. Dividend yields, again, fell to three percent in June of 1992. Price-earnings multiples rose to the mid-twenties. The subsequent return through June 2002 was 11.4 percent. The yield of the index fluctuated between two and three percent from 1993 through 1995 and earnings multiples remained in the mid-twenties, yet long horizon returns through June 30, 2002 fluctuated between 11 and 12 percent. Even from early December 1996, the date of Campbell and Shiller’s presentation to the Federal Reserve suggesting near zero returns for the S&P500, the index provided almost a seven percent annual return through mid-2002. Such results suggest to me a very cautious assessment of the extent to which stock market returns are predictable.

Other Predictable Time Series Patterns
Studies have found some amount of predictability of stock returns based on various financial statistics. For example, Fama and Schwert (1977) found that short-term interest rates were related to future stock returns. Campbell (1987) found that term structure of interest rates spreads contained useful information for forecasting stock returns, and Keim and Stambaugh (1986) found that risk spreads between high-yield corporate bonds and short rates had some predictive power. Again, even if some predictability exists, it may reflect time varying risk premiums and required rates of return for stock investors rather than an inefficiency. And it is far from clear that any of these results can be used to generate profitable trading strategies.

Cross-Sectional Predictable Patterns Based on Firm Characteristics and Valuation Parameters

A large number of patterns that are claimed to be predictable are based on firm characteristics and different valuation parameters.

The Size Effect

One of the strongest effects investigators have found is the tendency over long periods of time for smaller-company stocks to generate larger returns than those of large-company stocks. Since 1926, small-company stocks in the United States have produced rates of return over one percentage point larger than the returns from large stocks (Keim, 1983). Fama and French (1992) examined data from 1963 to 1990 and divided all stocks into deciles according to their size as measured by total capitalization. Decile one
contained the smallest ten percent of all stocks while decile ten contained the largest
stocks. The results, plotted in Exhibit 2, show a clear tendency for the deciles made up of
portfolios of smaller stocks to generate higher average monthly returns than deciles made
up of larger stocks.

The crucial issue here is the extent to which the higher returns of small companies
represents a predictable pattern that will allow investors to generate excess risk-adjusted
returns. According to the capital asset pricing model, the correct measure of risk for a
stock is its “beta” – that is, the extent to which the return of the stock is correlated with
the return for the market as a whole. If the “beta” measure of systematic risk from the
capital asset pricing model is accepted as the correct risk measurement statistic, the size
effect can be interpreted as indicating an anomaly and a market inefficiency, because
using this measure portfolios consisting of smaller stocks have excess risk-adjusted
returns. Fama and French point out, however, that the average relationship between
“beta” and return during the 1963-1990 period was flat – not upward sloping as the
capital asset pricing model predicts. Moreover, if stocks are divided up by beta deciles,
ten portfolios constructed by size display the same kind of positive relationship shown in
Exhibit 2. On the other hand, within size deciles, the relationship between beta and
return continues to be flat. Fama and French suggest that size may be a far better proxy
for risk than beta, and therefore that their findings should not be interpreted as indicating
that markets are inefficient.

Dependability of the size phenomenon is also open to question. From the mid-
1980s through the decade of the 1990s, there has been no gain from holding smaller
stocks. Indeed, in most world markets, larger capitalization stocks produced larger rates
of return. It may be that the growing institutionalization of the market led portfolio
managers to prefer larger companies with more liquidity to smaller companies where it
would be difficult to liquidate significant blocks of stock. Finally, it is also possible that
some studies of the small-firm effect have been affected by survivorship bias. Today’s
computerized databases of companies include only small firms that have survived, not the
ones that later went bankrupt. Thus, a researcher who examined the ten-year
performance of today’s small companies would be measuring the performance of those
companies that survived – not the ones that failed.

“Value” Stocks

There have been several studies that suggest that “value” stocks have higher
returns than so-called “growth” stocks. The most common two methods of identifying
value stocks have been price-earnings ratios and price-to-book-value ratios.

Stocks with low price-earnings multiples (often called “value” stocks) appear to
provide higher rates of return than stocks with high price-to-earnings ratios as first shown
by Nicholson (1960) and later confirmed by Ball (1978) and Basu (1977). This finding is
consistent with the views of behavioralists that investors tend to be overconfident of their
ability to project high earnings growth and thus overpay for “growth” stocks (for
example, Kahneman and Riepe, 1998). The finding is also consistent with the views of
Graham and Dodd (1934), first expounded in their classic book on security analysis and
later championed by the legendary U.S. investor Warren Buffett. Similar results have
been shown for price/cash flow multiples, where cash flow is defined as earnings plus
depreciation and amortization (Hawawini and Keim, 1995).
The ratio of stock price to book value, defined as the value of a firm’s assets minus its liabilities divided by the number of shares outstanding, has also been found to be a useful predictor of future security returns. Low price-to-book is considered to be another hallmark of so-called “value” in equity securities and is also consistent with the view of behavioralists that investors tend to overpay for “growth” stocks that subsequently fail to live up to expectations. Fama and French (1992) concluded that size and price-to-book-value together provide considerable explanatory power for future returns and once they are accounted for, little additional influence can be attributed to P/E multiples. Fama and French (1997) also conclude that the P/BV effect is important in many world stock markets other than the United States.

Such results raise questions about the efficiency of the market if one accepts the capital asset pricing model, as Lakonishok, Schleifer and Vishny (1994) point out. But these findings do not necessarily imply inefficiency. They may simply indicate failure of the CAPM to capture all the dimensions of risk. For example, Fama and French (1993) suggest that the price-to-book value ratio may reflect another risk factor that is priced into the market and not captured by CAPM. Companies in some degree of financial distress, for example, are likely to sell at low prices relative to book values. Fama and French (1993) argue that a three-factor asset-pricing model (including price-to-book-value and size as measures of risk) is the appropriate benchmark against which anomalies should be measured.

We also need to keep in mind that the results of published studies – even those done over decades – may still be time-dependent and ask whether the return patterns of academic studies can actually be generated with real money. Exhibit 3 presents average
actual returns generated by mutual funds classified by either their “growth” or “value”
objectives. “Value” funds are so classified if they buy stocks with price-to-earnings and
price-to-book-value multiples that are below the averages for the whole stock market.

Over a period running back to the 1930s, it does not appear that investors could actually
have realized higher rates of return from mutual funds specializing in “value” stocks.
Indeed, the exhibit suggests that the Fama-French period from the early 1960s through
1990 may have been a unique period in which value stocks rather consistently produced
higher rates of return.

Schwert (2001) points out that the investment firm of Dimensional Fund Advisors
actually began a mutual fund that selected value stocks quantitatively according to the
Fama and French (1993) criteria. The abnormal return of such a portfolio (adjusting for
beta, the capital asset pricing model measure of risk) was a negative 0.2 percent per
month over the 1993-1998 period. The absence during that period of an excess return to
the “value” stocks is consistent with the results from “actively managed” value mutual
funds shown in Exhibit 3.

The Equity Risk Premium Puzzle

Another puzzle that is often used to suggest that markets are less than fully
rational is the existence of a very large historical equity risk premium that seems
inconsistent with the actual riskiness of common stocks as can be measured statistically.
For example, using the Ibbotson data from 1926 through 2001, common stocks have
produced rates of return of approximately 10½ percent while high grade bonds have
returned only about 5½ percent. I believe that this finding is simply the result of a
combination of perceived equity risk being considerably higher during the early years of the period and of average equity returns being much higher than had been forecast by investors.

It is easy to say 50 to 75 years later that common stocks were underpriced during the 1930s and 1940s. But it is well to remember that the annual average almost six percent growth in corporate earnings and dividends that we have experienced since 1926 was hardly a foregone conclusion during a period of severe depression and world war. Indeed, the U.S. stock market is almost unique in that it is one of the few world markets that remained in continuous operation during the entire period and the measured risk premium results, in part, from survivorship bias. One must be very careful to distinguish between \textit{ex ante} expected risk premiums and \textit{ex post} measured ones. Eugene Fama and Kenneth French (2002) argue that the high average realized returns result in part from large \textit{unexpected} capital gains. Economists such as Shiller have suggested that during the early 2000s, the \textit{ex ante} equity risk premium was, if anything, irrationally too low.

**Summarizing the “Anomalies” and Predictable Patterns**

As the preceding sections have pointed out, many “anomalies” and statistically significant predictable patterns in the stock returns have been uncovered in the literature. However, these patterns are not robust and dependable in different sample periods, and some of the patterns based on fundamental valuation measures of individual stocks may simply reflect better proxies for measuring risk.
Moreover, many of these patterns, even if they did exist, could self-destruct in the future, as many of them have already done. Indeed, this is the logical reason why one should be cautious not to overemphasize these anomalies and predictable patterns.

Suppose, for example, one of the anomalies or predictable patterns appears to be robust. Suppose there is a truly dependable and exploitable January effect, that the stock market – especially stocks of small companies – will generate extraordinary returns during the first five days of January. What will investors do? They will buy on the last day of December, and sell on January 5. But then investors find that the market rallied on the last day of December and so they will need to begin to buy on the next-to-last day of December; and because there is so much “profit taking” on January 5, investors will have to sell on January 4 to take advantage of this effect. Thus, to beat the gun, investors will have to be buying earlier and earlier in December and selling earlier and earlier in January so that eventually the pattern will self-destruct. Any truly repetitive and exploitable pattern that can be discovered in the stock market and can be arbitraged away will self-destruct. Indeed, the January effect became undependable after it received considerable publicity.

Similarly, suppose there is a general tendency for stock prices to underreact to certain new events, leading to abnormal returns to investors who exploit the lack of full immediate adjustment (DeBondt and Thaler, 1995; Campbell, Lo and MacKinlay, 1977). “Quantitative” investment managers will then develop trading strategies to exploit the pattern. Indeed, the more potentially profitable a discoverable pattern is, the less likely it is to survive.
Many of the predictable patterns that have been discovered may simply be the result of data mining. The ease of experimenting with financial databanks of almost every conceivable dimension makes it quite likely that investigators will find some seemingly significant but wholly spurious correlation between financial variables or among financial and nonfinancial datasets. Given enough time and massaging of data series, it is possible to tease almost any pattern out of most datasets. Moreover, the published literature is likely to be biased in favor of reporting such results. Significant effects are likely to be published in professional journals while negative results, or boring confirmations of previous findings, are relegated to the file drawer or discarded. Data-mining problems are unique to nonexperimental sciences, such as economics, which rely on statistical analysis for their insights and cannot test hypotheses by running repeated controlled experiments.

An exchange at a symposium about a decade ago between Robert Shiller, an economist who is sympathetic to the argument that stock prices are partially predictable and skeptical about market efficiency, and Richard Roll, an academic financial economist who also is a portfolio manager, is quite revealing (Roll and Shiller, 1992). After Shiller stressed the importance of inefficiencies in the pricing of stocks, Roll responded as follows:

I have personally tried to invest money, my client’s money and my own, in every single anomaly and predictive device that academics have dreamed up. … I have attempted to exploit the so-called year-end anomalies and a whole variety of strategies supposedly documented by
academic research. *And I have yet to make a nickel on any of these* supposed market inefficiencies … a true market inefficiency ought to be an exploitable opportunity. If there’s nothing investors can exploit in a systematic way, time in and time out, then it’s very hard to say that information is not being properly incorporated into stock prices.

Seemingly Irrefutable Cases of Inefficiency

Critics of efficiency argue that there are several instances of recent market history where there is overwhelming evidence that market prices could not have been set by rational investors and that psychological considerations must have played the dominant role. It is alleged, for example, that the stock market lost about one-third of its value from early to mid-October 1987 with essentially no change in the general economic environment. How could market prices be efficient both at the start of October and during the middle of the month? Similarly, it is widely believed that the pricing of Internet stocks in early 2000 could only be explained by the behavior of irrational investors. Do such events make a belief in efficient markets untenable?

The Market Crash of October 1987

Can the October 1987 market crash be explained by rational considerations, or does such a rapid and significant change in market valuations prove the dominance of psychological rather than logical factors in understanding the stock market? Behaviorists
would say that the one-third drop in market prices, which occurred early in October 1987, can only be explained by relying on psychological considerations since the basic elements of the valuation equation did not change rapidly over that period. It is, of course, impossible to rule out the existence of behavioral or psychological influences on stock market pricing. But logical considerations can explain a sharp change in market valuations such as occurred during the first weeks of October 1987.

A number of factors could rationally have changed investors’ views about the proper value of the stock market in October 1987. For one thing, yields on long-term Treasury bonds increased from about 9 percent to almost 10 ½ percent in the two months prior to mid-October. Moreover, a number of events may rationally have increased risk perceptions during the first two weeks of October. Early in the month, Congress threatened to impose a “merger tax” that would have made merger activity prohibitively expensive and could well have ended the merger boom. The risk that merger activity might be curtailed increased risks throughout the stock market by weakening the discipline over corporate management that potential takeovers provide. Also, in early October 1987, then Secretary of the Treasury James Baker had threatened to encourage a further fall in the exchange value of the dollar, increasing risks for foreign investors and frightening domestic investors as well. While it is impossible to correlate each day’s movement in stock prices to specific news events, it is not unreasonable to ascribe the sharp decline in mid-October to the cumulative effect of a number of unfavorable “fundamental” events. As Merton Miller (1991) has written, “… on October 19, some weeks of external events, minor in themselves… cumulatively signaled a possible change in what had been up to then a very favorable political and economic climate for
equities… and … many investors simultaneously came to believe they were holding too
large a share of their wealth in risky equities.”

Share prices can be highly sensitive as a result of rational responses to small
changes in interest rates and risk perceptions. Suppose stocks are priced as the present
value of the expected future stream of dividends. For a long-term holder of stocks, this
rational principle of valuation translates to a formula:

\[ r = \frac{D}{P} + g, \]

where \( r \) is the rate of return, \( \frac{D}{P} \) is the (expected) dividend yield, and \( g \) is the long-term
growth rate. For present purposes, consider \( r \) to be the required rate of return for the
market as a whole. Suppose initially that the “riskless” rate of interest on government
bonds is 9 percent and that the required additional risk premium for equity investors is 2
percentage points. In this case \( r \) will be 11 percent (0.09 + 0.02 = 0.11). If a typical
stock’s expected growth rate, \( g \), is 7 percent and if the dividend is $4 per share, we can
solve for the appropriate price of the stock index \( (P) \), obtaining

\[ 0.11 = \frac{0.04}{P} + 0.07 \]

\[ P = \$100. \]

Now assume that yields on government bonds rise from 9 to 10 ½ percent, with
no increase in expected inflation, and that risk perceptions increase so that stock-market
investors now demand a premium of 2 ½ percentage points instead of the 2 points in the
previous example. The appropriate rate of return or discount rate for stocks, \( r \), rises then
from 11 percent to 13 percent (0.105 + 0.025), and the price of our stock index falls from
$100 to $66.67:
\[
0.13 = \frac{4\%}{P} + 0.07
\]

\[
P = $66.67
\]

The price must fall to raise the dividend yield from 4 to 6 percent so as to raise the total return by the required 2 percentage points. Clearly, no irrationality is required for share prices to suffer quite dramatic declines with the sorts of changes in interest rates and risk perceptions that occurred in October 1987. Of course, even a very small decline in anticipated growth would have magnified these declines in warranted share valuations.

This is not to say that psychological factors were irrelevant in explaining the sharp drop in prices during October 1987—they undoubtedly played a role. But it would be a mistake to dismiss the significant change in the external environment, which can provide an entirely rational explanation for a significant decline in the appropriate values for common stocks.

The Internet Bubble of the Late 1990s

Another stock market event often cited by behavioralists as clear evidence of the irrationality of markets is the Internet “bubble” of the late 1990s. Surely, the remarkable market values assigned to internet and related high-tech companies seem inconsistent with rational valuation. I have some sympathy with behavioralists in this instance, and in reviewing Robert Shiller’s (2000) *Irrational Exuberance* I agreed that it was in the high-tech sector of the market that his thesis could be supported. But even here, when we know after the fact that major errors were made, there were certainly no arbitrage opportunities available to rational investors before the bubble popped.
Equity valuations rest on uncertain future forecasts. Even if all market participants rationally price common stocks as the present value of all future cash flows expected, it is still possible for clear excesses to develop. We know now, with the benefit of hindsight, that outlandish and unsupportable claims that being made regarding the growth of the Internet (and the related telecommunications structure needed to support it). We know now that projections for the rates and duration of growth of these for “new economy” companies were unsustainable. But remember, it was the sharp-pencilled professional investors who argued that the valuations of high-tech companies were proper. Many of Wall Street’s most respected security analysts, including those independent of investment banking firms, were recommending Internet stocks to the firm’s institutional and individual clients as being fairly valued. Professional pension-fund and mutual fund managers over-weighted their portfolios with high-tech stocks.

While it is now clear in retrospect that such professionals were egregiously wrong, there was certainly no obvious arbitrage opportunity available. One could disagree with the projected growth rates of security analysts. But who could be sure, with the use of the Internet for a time doubling every several months that the extraordinary growth rates that could justify stock valuations were impossible? After all, even Alan Greenspan was singing the praises of the new economy. Nothing is ever as clear in prospect as it is in retrospect. Certainly, the extent of the bubble was only clear in retrospect.

Not only is it almost impossible to judge with confidence what the proper fundamental value is for any security, but also potential arbitrageurs face
additional risks. Shleifer (2000) has argued that noise trader risk limits the extent
to which one should expect arbitrage to bring prices quickly back to rational
values even in the presence of an apparent bubble. Professional arbitrageurs will
be loath to sell short a stock they believe is trading at two times its “fundamental”
value when it is always possible that some greater fools may be willing to pay
three times the stock’s value. Arbitrageurs are quite likely to have short horizons
since even temporary losses may induce their clients to withdraw their money.

While there were no arbitrage opportunities available during the Internet
bubble that adjusted returns, and while stock prices eventually did adjust to levels
that more reasonably reflected the likely present value of their cash flows, an
argument can be maintained the asset prices did remain “incorrect” for a period of
time. The result was that too much new capital flowed to Internet and related
telecommunications companies. Thus, the stock market may well have
temporarily failed in its role as an efficient allocator of equity capital.
Fortunately, “bubble” periods are the exception rather than the rule and
acceptance of such occasional mistakes is the necessary price of a flexible market
system that usually does a very effective job of allocating capital to its most
productive uses.

Other Illustrations of Irrational Pricing

Are there not some illustrations of irrational pricing that can be clearly
ascertained as they arise, not simply after a bubble has burst? My favorite
illustration concerns the spin off of Palm Pilot from its parent 3-Com Corporation
during the height of the Internet boom in early 2000. Initially, only 5 percent of the Palm Pilot shares were distributed to the public; the other 95 percent remained on 3-Com’s balance sheet. As Palm Pilot began trading, enthusiasm for the shares was so great that the 95 percent of its shares still owed by 3-Com had a market value considerably more than the entire market capitalization of 3-Com, implying that all the rest of its business had a negative value. Other illustrations involve ticker symbol confusion. Rasches (2001) finds clear evidence of co-movement of stocks with similar ticker symbols; for example, the stock of MCI Corporation (ticker symbol MCIC) moves in tandem with an unrelated closed-end bond investment fund Mass Mutual Corporate Investors (ticker symbol MCI).. In a charming article entitled “A Rose.com by Any Other Name,” Cooper, Dimitrov, and Rau (2001) found positive stock price reactions during 1998 and 1999 on corporate name changes when dot com was added to the corporate title. Finally, it has been argued that closed-end funds sell at irrational discounts from their net asset values (for example, Shleifer, 2000).

But none of these illustrations should shake our faith that exploitable arbitrage opportunities should not exist in an efficient market. The apparent arbitrage in the Palm Pilot case (sell Palm Pilot short and buy 3-Com) could not be undertaken because not enough Palm stock was outstanding to make borrowing the stock possible to effectuate a short sale. The “anomaly” disappeared once 3-Com spun off more of Palm stock. Moreover, the potential profits from name or ticker symbol confusion are extremely small relative to the transactions costs that would be required to exploit them. Finally, the “closed-end
fund puzzle” is not really a puzzle today. Discounts have narrowed from historical averages for funds with assets traded in liquid markets and researchers such as Ross (2001) have suggested that they can largely be explained by fund management fees. Perhaps the more important puzzle today is why so many investors buy high expense, actively managed mutual funds instead of low cost index funds.

The Performance of Professional Investors

For me, the most direct and most convincing tests of market efficiency are direct tests of the ability of professional fund managers to outperform the market as a whole. Surely, if market prices were determined by irrational investors and systematically deviated from rational estimates of the present value of corporations, and if it was easy to spot predictable patterns in security returns or anomalous security prices, then professional fund managers should be able to beat the market. Direct tests of the actual performance of professionals, who often are compensated with strong incentives to outperform the market, should represent the most compelling evidence of market efficiency.

A remarkably large body of evidence suggesting that professional investment managers are not able to outperform index funds that simply buy and hold the broad stock market portfolio. The first study of mutual fund performance was undertaken by Jensen (1969). He found that active mutual fund managers were unable to add value and, in fact, tended to underperform the market by approximately the amount of their added
expenses. I repeated Jensen’s study with data from a subsequent period and confirmed
the earlier results (Malkiel, 1995). Moreover, I found that the degree of “survivorship
bias” in the data was substantial; that is, poorly performing funds tend to be merged into
other funds in the mutual fund’s family complex thus burying the records of many of the
underperformers. Exhibit 4 updates the study I performed through mid-2002.
Survivorship bias makes the interpretation of long-run mutual fund data sets very
difficult. But even using data sets with some degree of survivorship bias, one cannot
sustain the argument that professional investors can beat the market.

Exhibit 5 presents the percentage of actively managed mutual funds that have been outperformed by the Standard & Poor’s 500 and the Wilshire stock indexes. Throughout the past decade about three-quarters of actively managed funds have failed to
beat the index. Similar results obtain for earlier decades. Exhibit 6 shows that the
median large capitalization professionally managed equity fund has underperformed the
S&P 500 index by almost two percentage points over the past 10, 15, and 20-year
periods. Exhibit 7 shows similar results in different markets and against different
benchmarks.

Managed funds are regularly outperformed by broad index funds, with equivalent
risk. Moreover, those funds that produce excess returns in one period are not likely to do
so in the next. There is no dependable persistence in performance. During the 1970s, the
top 20 mutual funds enjoyed almost double the performance of the index. During the
1980s, those same funds underperformed the index. The best performing funds of the
1980s similarly underperformed during the 1990s. And a more dramatic example of the
lack of persistence in performance is shown in Exhibit 8. The top 20 mutual funds during
1998 and 1999 enjoyed three times the performance of the index. During 2000 and 2001 they did three times worse than the index. Over the long run, the results are even more devastating to active managers. One can count on the fingers of one hand the number of professional portfolio managers who have managed to beat the market by any significant amount. Exhibit 9 shows the distribution of returns over a 30-year period. Of the original 355 funds, only five of them outperformed the market by two percentage points per year or more.

The record of professionals does not suggest that sufficient predictability exists in the stock market or that there are recognizable and exploitable irrationalities sufficient to produce excess returns.

Conclusion

As long as stock markets exist, the collective judgment of investors will sometimes make mistakes. Undoubtedly, some market participants are demonstrably less than rational. As a result, pricing irregularities and predictable patterns in stock returns can appear over time and even persist for short periods. Moreover, the market cannot be perfectly efficient or there would be no incentive for professionals to uncover the information that gets so quickly reflected in market prices, a point stressed by Grossman and Stiglitz (1980). Undoubtedly, with the passage of time and with the increasing sophistication of our databases and empirical techniques, we will document further apparent departures from efficiency and further patterns in the development of stock returns.
But I suspect that the end result will not be an abandonment of the belief of many in the profession that the stock market is remarkably efficient in its utilization of information. Periods such as 1999 where “bubbles” seem to have existed, at least in certain sectors of the market, are fortunately the exception rather than the rule. Moreover, whatever patterns or irrationalities in the pricing of individual stocks that have been discovered in a search of historical experience are unlikely to persist and will not provide investors with a method to obtain extraordinary returns. If any $100 bills are lying around the stock exchanges of the world, they will not be there for long.
An Introduction to Portfolio Theory

Paul J. Atzberger

Any comments or errors please e-mail: paulatz@cims.nyu.edu
Introduction

Portfolio theory deals with the problem of constructing for a given collection of assets an investment with desirable features. A variety of different asset characteristics can be taken into consideration, such as the amount of value, on average, an asset returns on over a period of time and the riskiness of reaping returns comparable to the average. The financial objectives of the investor and tolerance of risk determine what types of portfolios are to be considered desirable. In these notes we shall discuss a quantitative approach to constructing portfolios. In particular, we shall use the methods of constrained optimization to construct portfolios for a given collection of assets with desirable features as quantitated by an appropriate utility function and constraints. The materials presented here are taken from the following sources: Theory of Finance - Mean Variance Analysis by Simon Hubbert, and Investments by Bodie, Kane, and Marcus.

Characterizing the Rates of Return of Assets and Portfolios

We shall concern ourselves with primarily two basic features of an asset. The first is the average return of an asset over a period of time. The second characteristic is how risky it is to obtain similar returns comparable to the average over the investment period.

For an asset with value $S(0)$ at time 0 and value $S(T)$ at time $T$, the rate of return $\rho$ is defined by:

$$S(T) = (1 + \rho)S(0).$$

(1)

The rate of return can be thought of as an “effective interest rate” which would be required for a deposit of $S(0)$ into a savings account at a bank to obtain the same change in value as the asset over the period $[0, T]$. For example, if $S(0) = $4 and after one year $S(1) = $6, the rate of return of the asset is $\rho = S(1) - S(0)/S(0) = 50\%$. The rate of return of an asset is also sometimes referred to as the “yield” of the asset.

Since the outcome of an investment in an asset has some level of uncertainty, the value $S(T)$ is unknown exactly at time 0. To model the uncertainty we shall consider the value of the asset at time $T$ as a random variable. Correspondingly, the rate of return $\rho$ defined by equation 1 is also a random variable. To characterize the asset we shall consider the average rate of return defined by:

$$\mu = E(\rho)$$

(2)

where $E(\cdot)$ denotes the expectation of a random variable. This is also sometimes referred to as the “expected rate of return”. While the expected rate of return is a useful way to characterize an asset and gives us some indication of how
large the returns may be, it does not capture the uncertainty in obtaining a comparable return rate to the average.

To quantify how much the rate of return deviates from the expected return and in order to capture the riskiness of the asset, we shall use the variance defined by:

$$\sigma^2 = \text{Var}(\rho) = E(\left|\rho - \mu\right|^2).$$  \hspace{1cm} (3)

For a given collection of \(n\) assets \(\{S_1, S_2, \ldots, S_n\}\), for the \(i^{th}\) asset we denote the rate of return by \(\rho_i\) and the variance by \(\sigma_i^2\). For a collection of assets we shall find it useful to consider, in addition, how the random rates of return are coupled.

For example, in choosing investments it is important to take into account not only the individual returns of the assets but also how the returns are coupled among the assets. A natural investment strategy to reduce risk in which an asset loses value should a given event occur, is to try to find another asset which increases in value should this event happen. This requires that the assets exhibit a coupling in which values move in opposite directions should the event occur. To quantify this for random rates of return we shall use the covariance for the returns defined by:

$$\sigma_{i,j} = E((\rho_i - \mu_i)(\rho_j - \mu_j)).$$  \hspace{1cm} (4)

We remark that \(\sigma_{i,j} = \sigma_{j,i}\) and that when \(i = j\) we have \(\sigma_{i,i} = \sigma_i^2\).

To describe the coupling of all \(n\) assets we defined the covariance matrix by:

$$V = \begin{bmatrix}
\sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,n} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{n,1} & \sigma_{n,2} & \cdots & \sigma_{n,n}
\end{bmatrix}.\hspace{1cm} (5)$$

We remark that \(V\) is a symmetric matrix and can also be shown to be positive definite.

A portfolio is an investment made in \(n\) assets using some amount of wealth \(W\). Let \(W_i\) denote the amount of money invested in the \(i^{th}\) asset. We shall allow negative values of \(W_i\), which for example can be interpreted as short selling an asset. In other words, we assume a liability and must deliver the asset at a future time. Since the total wealth invested is \(W\) we have:

$$\sum_{i=1}^{n} W_i = W.\hspace{1cm} (6)$$

To avoid working with absolute magnitudes of the assets and portfolios, we shall find it convenient to instead describe the investments in terms of relative values such as the rates of return of the assets and the relative portion of wealth invested in a given asset. For the fraction of wealth invested in the \(i^{th}\) asset we make the definition:

$$w_i = \frac{W_i}{W}.\hspace{1cm} (7)$$
Equation 6 then implies:

$$\sum_{i=1}^{n} w_i = 1.$$  \hspace{1cm} (8)

The value $Q_p$ of the portfolio at time $t$ can be expressed as:

$$Q_p(t) = \sum_{i=1}^{n} \frac{W_i}{S_i(0)} S_i(t).$$  \hspace{1cm} (9)

where we use that the portfolio has the value $Q_p(0) = W$ from equation 6.

The rate of return of the portfolio $\rho_p$ at time $t$ is given by:

$$\rho_p(t) = \frac{Q_p(t) - Q_p(0)}{Q_p(0)} = \sum_{i=1}^{n} \frac{W_i}{S_i(0)} S_i(t) - W$$

$$= \sum_{i=1}^{n} w_i \left( \frac{S_i(t) - S_i(0)}{S_i(0)} \right) - W$$

$$= \sum_{i=1}^{n} w_i \rho_i.$$  \hspace{1cm} (10)

In other words, the rate of return for a portfolio is the weighted average of the rates of return of the assets where the weights are determined by the fraction of wealth invested in each asset.

The expected rate of return $\mu_p$ of the portfolio is given by:

$$\mu_p = E \left( \sum_{i=1}^{n} w_i \rho_i \right)$$

$$= \sum_{i=1}^{n} w_i \mu_i.$$  \hspace{1cm} (11)

where the linearity property of expectations has been used.
The variance $\sigma_p^2$ for the rate of return of the portfolio is given by:

$$\sigma_p^2 = E \left( |\rho_p - \mu_p|^2 \right)$$

$$= E \left( \left( \sum_{i=1}^{n} w_i (\rho_i - \mu_i) \right)^2 \right)$$

$$= \left( \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{i,j} \right)$$

$$= w^T V w.$$  

where $w^T = [w_1, \ldots, w_N]$ and $V$ is defined in equation 5.

For a portfolio $a$ and a portfolio $b$ we shall quantitatively describe the coupling between two portfolios by using the covariance of the random rates of returns of the two portfolios $\rho_{p}^{(a)}$ and $\rho_{p}^{(b)}$. This is given by:

$$\sigma^{(a,b)} = E \left( (\rho_{p}^{(a)} - \mu_{p}^{(a)}) (\rho_{p}^{(b)} - \mu_{p}^{(b)}) \right)$$

$$= \left( \sum_{i=1}^{n} \sum_{j=1}^{n} w_i^{(a)} w_j^{(b)} \sigma_{i,j} \right)$$

$$= \left( w_p^{(a)} \right)^T V w_p^{(b)}.$$ 

To summarize, we shall characterize the average rate of return, riskiness in obtaining comparable returns to the average, and coupling among the returns for both portfolios and individual assets. This will be done quantitatively by using, respectively, the expected rate of return, variance of the return, and covariance of the returns.

**Desirable Portfolios (Markowitz Theory)**

Determining what constitutes a desirable portfolio depends on many factors. The primary factors we shall consider are the financial objectives of the investor and his or her tolerance for risk in achieving these objectives.
For example, imagine two assets such as stock of a biotechnology start-up company and blue chip stock in a company such as General Electric. The biotechnology company is a rather unproven enterprise that could go bankrupt should some important technical obstacle arise on their way to producing a product, while General Electric is a proven enterprise with a solid track record. Given these circumstances a wise investor would require some form of compensation for the riskiness of choosing the biotechnology stock over the blue chip stock. The biotechnology stock while risky likely has the potential to skyrocket in value should the enterprise carry through on its ambitious objectives, while the General Electric stock is more likely to have a more modest rate of return somewhere around its stated performance objectives.

To cast these considerations in terms of our quantitative theory, the scenario above suggests that if an investor is presented with two assets, the investor would choose the one having the larger variance only if this also entails having a larger expected return. This larger expected return acts as an incentive for the investor, we shall refer to this as a “risk premium” which compensates the investor for taking the larger risk. Another way to think about these preferences in our theory, is that if the two assets had the same expected rate of return, in other words there was no “risk premium”, then an investor would choose the one with the smallest variance (least risk).

With this understanding about the preferences of the investor, we shall consider a portfolio to be desirable if for a given expected rate of return $\mu_p$, the portfolio has the least variance $\sigma_p^2$. Finding such a portfolio is referred to as the Markowitz problem and can be stated mathematically as the constrained optimization problem:

$$\text{minimize } f(w_1, \ldots, w_n) = \frac{1}{2} \sum_{i,j=1}^{n} w_i w_j \sigma_{i,j} \quad (15)$$

subject:

$$g_1(w_1, \ldots, w_n) = \sum_{i=1}^{n} w_i \mu_i - \mu_p = 0$$

$$g_2(w_1, \ldots, w_n) = \sum_{i=1}^{n} w_i - 1 = 0 \quad (16)$$

The objective function is the variance of the portfolio, as computed in equation 13. The first constraint specifies that the constructed portfolio is to have expected rate of return $\mu_p$ while the second constraint arises from equation 8 defining the portfolio. To solve the constrained optimization problem 15 analytically, we shall use the Method of Lagrange Multipliers. Many numerical methods also exist for these types of optimization problems.
The Method of Lagrange Multipliers

In the Method of Lagrange Multipliers the constraints are taken into account by solving an unconstrained optimization problem for the following utility function, referred to as the Lagrangian:

\[ L(w_1, \ldots, w_n | \lambda_1, \lambda_2) = f(w_1, \ldots, w_n) - \lambda_1 g_1(w_1, \ldots, w_n) - \lambda_2 g_2(w_1, \ldots, w_n). \]  

(17)

By design a critical point of the Lagrangian is also a critical point of \( f \) and satisfies the constraints \( g_1 = 0 = g_2 \). This follows since the condition that \( (\mathbf{w}, \lambda_1, \lambda_2) \) be a critical point of \( L \) is given by:

\[
\nabla_{\mathbf{w}} L = \begin{bmatrix}
\frac{\partial L}{\partial w_1} \\
\vdots \\
\frac{\partial L}{\partial w_n}
\end{bmatrix} = 0
\]  

(18)

and

\[
\frac{\partial L}{\partial \lambda_1} = 0, \quad \frac{\partial L}{\partial \lambda_2} = 0.
\]  

(19)

In particular, we have that

\[
\frac{\partial L}{\partial \lambda_1} = g_1 = 0 \quad \text{and} \quad \frac{\partial L}{\partial \lambda_2} = g_2 = 0.
\]  

(20)

(21)

From this and equation 17, 18 we have \( \mathbf{w} \) is a critical point of \( f \).

Geometrically, the condition is equivalent to the level surfaces determined by constant values of the utility function \( f \) meeting the level surfaces of the constraints \( g_1 = 0 \) and \( g_2 = 0 \) so that the normals of the surface align. In other words, moving in the permissible directions determined by the intersection of the tangent planes of \( g_1 = 0 \) and \( g_2 = 0 \) the value of \( f \) to first order remains constant. We shall demonstrate more concretely how this method can be used in practice in the sections below.

Optimal Portfolios of \( n \) Risky Assets

We now discuss the portfolios which are optimal in the sense that for a specified expected return the portfolio minimizes the variance of the return. To ensure a well-defined solution exists, we shall make two assumptions about the collection of assets: (i) the random returns are linearly independent, in the sense that any on return can not be expressed as a linear combination of the others, (ii) the expected return rates \( \mu_i \) of the assets are not all equal. These assumptions
ensure there is a unique solution of the Markowitz problem given in equation 15.

We shall now use the Method of Lagrange Multipliers to solve the Markowitz problem analytically. The Lagrangian is given by:

$$ L(w, \lambda_1, \lambda_2) = \frac{1}{2} w^T V w + \lambda_1 (\mu_p - w^T \mu) + \lambda_2 (1 - w^T 1) $$

(22)

To find the critical points of the Lagrangian, and hence the optimal portfolio, we must solve the first order equations:

$$ \nabla_w L = V w_p - \lambda_1 \mu - \lambda_2 1 = 0 $$

(23)

and

$$ \frac{\partial L}{\partial \lambda_1} = \mu_p - w_p^T \mu = 0 $$

(24)

$$ \frac{\partial L}{\partial \lambda_2} = 1 - w_p^T 1 = 0. $$

(25)

From 23 we have:

$$ w_p = \lambda_1 (V^{-1} \mu) + \lambda_2 (V^{-1} 1) $$

(26)

and from 24 and 25 we have:

$$ (\mu^T V^{-1} \mu) \lambda_1 + (\mu^T V^{-1} 1) \lambda_2 = \mu_p $$

(27)

$$ (1^T V^{-1} \mu) \lambda_1 + (1^T V^{-1} 1) \lambda_2 = 1 $$

(28)

Using that

$$ (\mu^T V^{-1} 1) = (\mu^T V^{-1} 1)^T = 1^T (V^{-1})^T \mu = 1^T V^{-1} \mu $$

(29)

we can express 27 and 28 as:

$$ \begin{bmatrix} B & A \\ A & C \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix} $$

(30)

where

$$ \begin{bmatrix} B & A \\ A & C \end{bmatrix} = \begin{bmatrix} \mu^T V^{-1} \mu & \mu^T V^{-1} 1 \\ 1^T V^{-1} \mu & 1^T V^{-1} 1 \end{bmatrix}. $$

(31)

To ensure there is a solution of 31 requires that the determinant be non-zero:

$$ D = BC - A^2 \neq 0. $$

(32)

We shall now show this is indeed the case, under our assumptions. A matrix $V$ is called positive definite if:

$$ w^T V w > 0 \text{ for any } w \neq 0. $$

(33)
The inverse $V^{-1}$ of a positive definite matrix is also positive definite. We shall show $D \neq 0$ as a consequence of the positive definiteness of $V$.

Let us consider the vector:

$$A\mu - B1.$$  \hfill (34)

The vector vanishes only if

$$0 = A\mu - B1 = \mu^TV^{-1}1\mu - \mu^TV^{-1}\mu1.$$  \hfill (35)

This is ruled out, since the only solution of this equation is $\mu = 1$, which is forbidden by the assumption (ii) that the $\mu_i$ are not all equal. Thus under assumption (ii) we can assume that $A\mu - B1 \neq 0$.

From the positive definiteness of $V$ we have that $V^{-1}$ is positive definite and that:

$$0 < (A\mu - B1)^T V^{-1} (A\mu - B1)$$  \hfill (36)

$$= A^2\mu^TV^{-1}\mu - AB\mu^TV^{-1}1 - BA1^TV^{-1}\mu + B^21^TV^{-1}1$$  \hfill (37)

$$= B^2C - A^2B = B(BC - A^2).$$  \hfill (38)

From the fact that $B = \mu^TV^{-1}\mu > 0$ by positive definiteness of $V$ we have that $D = BC - A^2 > 0$.

In particular, $D \neq 0$, and we can invert the matrix in equation 31 to obtain:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \frac{1}{D} \begin{bmatrix} A & -C \\ -C & B \end{bmatrix} \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}.$$  \hfill (39)

More explicitly, this gives the solution for $\lambda_1$, $\lambda_2$:

$$\lambda_1 = \frac{C\mu_p - A}{D}$$  \hfill (40)

$$\lambda_2 = \frac{B - A\mu_p}{D}.$$  \hfill (41)

The weights of the portfolio can be found by substitution into equation 26:

$$w_p = \left( \frac{C\mu_p - A}{D} \right) V^{-1} \mu + \left( \frac{B - A\mu_p}{D} \right) V^{-1}1$$  \hfill (42)

$$= \frac{1}{D} (BV^{-1}1 - AV^{-1}\mu) + \frac{1}{D} (CV^{-1}\mu - AV^{-1}1) \mu_p$$  \hfill (43)

For the definitions:

$$g = \frac{1}{D} (BV^{-1}1 - AV^{-1}\mu)$$  \hfill (44)

$$h = \frac{1}{D} (CV^{-1}\mu - AV^{-1}1)$$  \hfill (45)

we can express this more succinctly as:

$$w_p = g + \mu_p h.$$  \hfill (46)
We shall refer to the portfolio which minimizes the variance for a specified expected rate of return $\mu_p$ as a \textit{frontier portfolio}. We shall refer to the portfolio which has the minimum variance out of all frontier portfolios as the \textit{minimum variance portfolio} and denote this by $p^\star$.

The expression 46 for the solution shows explicitly how the the weights of the portfolio with minimum variance depend on the desired expected rate of return $\mu_p$. In particular, it follows that all frontier portfolios can be attained by a linear combination of the two portfolios $g$ and $h$. In other words, instead of buying the individual assets, in theory one could obtain an equivalent rate of return of any frontier portfolio by investing in only two mutual funds (frontier portfolios) of the market. This is known as the \textit{two fund theorem}. In practice, however, things are more complicated since this presupposes that investors only care about the expected returns, variances, and covariances of the assets and that investors agree on the expected returns, variances, and covariances to associated with the assets.

Some interesting features of investing in portfolios, however, can be obtained from the theory. The covariance between the random rates of return for two frontier portfolios $a$ and $b$ can be expressed in terms of $\mu_p^{(a)}$ and $\mu_p^{(b)}$ as:

$$
\text{cov}(\rho_p^{(a)}, \rho_p^{(b)}) = \left( \mathbf{w}_p^{(a)} \right)^T \mathbf{V} \mathbf{w}_p^{(b)} = \left( \mathbf{w}_p^{(a)} \right)^T \mathbf{V} \left( \mathbf{g} + \mathbf{h}_p^{(b)} \right)
$$

$$
= \left( \mathbf{w}_p^{(a)} \right)^T \mathbf{V} \left( \frac{1}{D} (B\mathbf{1} - A\mathbf{1}) + \frac{1}{D} (C\mathbf{1} - A\mathbf{1}) \mu_p^{(b)} \right)
$$

$$
= \frac{1}{D} \left( \mathbf{w}_p^{(a)} \right)^T \left( (B\mathbf{1} - A\mathbf{1}) + (C\mathbf{1} - A\mathbf{1}) \mu_p^{(b)} \right)
$$

$$
= \frac{C}{D} \left( \mu_p^{(a)} - \frac{A}{C} \right) \left( \mu_p^{(b)} - \frac{A}{C} \right) + \frac{B}{C} - \left( \frac{A}{C} \right)^2
$$

$$
= \frac{C}{D} \left( \mu_p^{(a)} - \frac{A}{C} \right) \left( \mu_p^{(b)} - \frac{A}{C} \right) + \frac{1}{C}.
$$

When the frontier portfolios are identical $a = b$ we obtain the variance:

$$
\sigma_p^2 = \text{cov}(\rho_p, \rho_p) = \frac{C}{D} \left( \mu_p - \frac{A}{C} \right)^2 + \frac{1}{C}.
$$

This can equivalently be expressed as:

$$
\frac{\sigma_p^2}{1/C} - \frac{(\mu_p - A/C)^2}{D/C^2} = 1.
$$

From this, we notice that the relationship of the expected rate of return $\mu_p$ and the variance of the frontier portfolio $\sigma_p^2$ can be summarized as a plot in the
Figure 1: Frontier Portfolios for $n$ Risky Assets. (curve) frontier portfolios, $(\times)$ denotes the minimum variance portfolio, $(\ast)$ denotes an inefficient frontier portfolio for given expected return $\mu_p$, $(\ast)$ denotes an efficient frontier portfolio for given expected return $\mu_p$, and $(\cdot)$ denotes a portfolio that does not minimize the variance for a given return.

$(\sigma, \mu)$-plane by a hyperbola with center $(0, A/C)$, asymptotes

$$\mu = \pm \sqrt{\frac{D}{C}} \sigma + \frac{A}{C},$$

and vertex

$$v_0 = \left(\sqrt{\frac{1}{C}}, \frac{A}{C}\right).$$

From this we have that the minimum variance portfolio is given by:

$$w_{p^*} = g + h\mu_{p^*} = g + \frac{A}{C}h.$$
The frontier portfolios and the minimum variance portfolio $\times$ are plotted in Figure 1. From Figure 1 we see that it is possible for two portfolios with different rates of return to have the same variance $\sigma_p^2$. Since both portfolios offer the same amount of risk, as quantitated by the variance, but one gives a better return, the portfolio with the greater return would be more desirable to an investor.

We shall refer to the portfolio with the lesser expected return
\[ \mu_p < \mu_{p^*} = \frac{A}{C} \]
as inefficient. We shall refer to the portfolio with the greater expected return
\[ \mu_p > \mu_{p^*} = \frac{A}{C} \]
as an efficient frontier portfolio. In Figure 1 we have for the variance $\sigma_p^2 \approx 0.45$ two portfolios, the one denoted by $(\ast)$ is inefficient while the one denoted by $(\oplus)$ is efficient.

We now discuss a feature of the minimum variance portfolio. It can be shown that the covariance of the minimum variance portfolio $p^*$ with any other frontier portfolio is the same. This follows from:
\[ \text{cov}(\rho_{p^*}, \rho_p) = \frac{C}{D} \left( \left( \mu_{p^*} - \frac{A}{C} \right) \left( \mu_p - \frac{A}{C} \right) + \frac{1}{C} \right) = \frac{1}{C} = \sigma_{p^*}^2 \]
where we use that $\mu_{p^*} = A/C$. As a consequence, we see that it is not possible to find any portfolio which is completely independent of the minimum variance portfolio.

However, for any other frontier portfolio $\alpha$ it can be shown that there exists a portfolio $z(\alpha)$ having zero covariance:
\[ \text{cov}(\rho_{p^*}, \rho_p) = 0. \]
This is given by solving:
\[ \text{cov}(\rho_{p^*}^{(\alpha)}, \rho_p^{(z(\alpha))}) = \frac{C}{D} \left( \mu_{p^*}^{(\alpha)} - \frac{A}{C} \right) \left( \mu_p^{(z(\alpha))} - \frac{A}{C} \right) + \frac{1}{C} = 0 \]
which has solution:
\[ \mu_p^{(z(\alpha))} = \frac{A}{C} - \frac{D/C^2}{\left( \mu_p^{(\alpha)} - A/C \right)}. \]

**Optimal Portfolios for $n$ Risky Assets + a Risk-Free Asset**

We shall now consider the case in which in addition to the the $n$ risky assets there is one risk-free asset with return $\rho_0$. By risk-free we mean that the rate
of return of the asset has zero variance. This will make our model somewhat more realistic since in actual markets an investor always has the opportunity to invest in an essentially risk-free treasury bond or put their money in a savings account. In this case the portion of wealth invested in the \( n \) risky assets no longer satisfies \( \sum_{i=1}^{n} w_i = 1 \) since we can always invest the remaining fraction of wealth in the risk-free asset or borrow funds. In the case that

\[
\sum_{i=1}^{n} w_i < 1
\]  

we say that we are \textit{under budget} and invest the remaining portion of wealth in the risk-free asset. In the case that

\[
\sum_{i=1}^{n} w_i < 1
\]  

we say that we are \textit{over budget} and borrow at the risk-free rate the excessive portion of wealth invested in the \( n \) risky assets.

This leads to a reformulation of what is meant by an optimal portfolio. In this case, an investor is no longer constrained to invest all of his or her wealth in the \( n \) risky assets. The objective then becomes for a specified expected rate of return to find the portfolio with the minimum variance subject only to the constraint that the expected return is attained. This gives:

\[
\begin{align*}
\text{minimize } f(w_1, \ldots, w_n) &= \frac{1}{2} w^T V w \\
\text{subject: } g(w_1, \ldots, w_n) &= \rho_0 + w^T (\mu - \rho_0 1) - \mu_{p+} = 0.
\end{align*}
\]  

The Lagrangian in this case becomes:

\[
L(w|\lambda) = \frac{1}{2} w^T V w - \lambda (\rho_0 - \mu_{p+} + w^T (\mu - \rho_0 1)).
\]  

The condition that \( (w, \lambda) \) be a critical point becomes:

\[
\begin{align*}
\nabla_w L &= V w_{p+} - \lambda (\mu - \rho_0 1) = 0 \\
\frac{\partial L}{\partial \lambda} &= \rho_0 - \mu_{p+} + w_{p+}^T (\mu - \rho_0 1) = 0.
\end{align*}
\]  

Using equation 64 we have:

\[
w_{p+} = \lambda V^{-1} (\mu - \rho_0 1).
\]  

From equation 65 we have:

\[
\lambda = \frac{\mu_{p+} - \rho_0}{(\mu - \rho_0 1)^T V^{-1} (\mu - \rho_0 1)}.
\]  

Letting

\[
H := (\mu - \rho_0 1)^T V^{-1} (\mu - \rho_0 1)
\]  

we obtain:

\[ w_{p^+} = \left( \frac{\mu_{p^+} - \rho_0}{H} \right) V^{-1} (\mu - \rho_0 1). \]  

(69)

Now it can be shown that in fact \( H > 0 \). Multiplying out the terms in equation 68 gives:

\[ H = (1^T V^{-1} 1) \rho_0^2 - 2 (1^T V^{-1} \mu) \rho_0 - 2 (1^T V^{-1} 1) . \]  

(70)

Using the definitions of \( A, B, C \) made in equation 30 we obtain the expression:

\[ H = C \rho_0^2 - 2A \rho_0 + B \]  

(71)

which is a quadratic in \( \rho_0 \). The value of \( H \) has the same sign for all values of \( \rho_0 \) if and only if there are no real roots. No real roots occur for the quadratic only if the discriminate is negative:

\[ (2A)^2 - 4CB = 4(A^2 - BC) < 0. \]  

(72)

This can be shown to hold by the positive definiteness of \( V^{-1} \) and follows immediately from equation 36.

The portfolio with the smallest variance for the specified expected return \( \mu_{p^+} \) is then given by:

\[
\begin{align*}
\sigma^2_{p^+} &= w_{p^+}^T V w_{p^+} \\
&= \left( \frac{\mu_{p^+} - \rho_0}{H} \right) V^{-1} (\mu - \rho_0 1)^T V \left( \frac{\mu_{p^+} - \rho_0}{H} \right) V^{-1} (\mu - \rho_0 1) \\
&= \left( \frac{\mu_{p^+} - \rho_0}{H} \right)^2 (\mu - \rho_0 1)^T (\mu - \rho_0 1) \\
&= \frac{(\mu_{p^+} - \rho_0)^2}{H}.
\end{align*}
\]

(73)

This can be expressed by:

\[ \sigma_{p^+} = \left| \frac{\mu_{p^+} - \rho_0}{\sqrt{H}} \right|. \]  

(74)

The relationship between the variance and expected returns of the portfolio can be summarized in the \((\sigma, \mu)\)-plane as two half-lines with slopes \( \pm \sqrt{H} \) intersecting the point \((0, \rho_0)\).

For any frontier portfolio \( a \), other than the one with \( \mu_{p^+} = A/C \), a portfolio \( z(a) \) having zero covariance with \( a \) can be found. The expected return of this portfolio is:

\[ \mu_{p^+}^{z(a)} = \frac{A}{C} - \frac{D/C^2}{\mu_{p^+}^{(a)} - A/C}. \]  

(75)
and has investment weights:
\[ w_{p^+}^{z(a)} = g + h\mu_{p^+}^{z(a)}. \]  

Geometrically, the return \( \mu_{p^+}^{z(a)} \) of the zero-covariance portfolio \( z(a) \) corresponds to the \( \mu \)-intercept of the tangent line at the point \( (\mu_{p^+}^{a}, \sigma_{p^+}^{a}) \) of the hyperbola. In other words, to find the expected return of the zero-covariance portfolio of \( a \), we draw, starting from the point corresponding to the portfolio \( a \), the tangent line and determine where this line crosses the \( \mu \) axis. The corresponding variance \( \sigma_{p^+}^{z(a)} \) to the portfolio \( z(a) \) is then found by drawing a horizontal line in the \( \sigma \) direction and determining the intersection with the frontier hyperbola, see Figure 2.

We now show that there is an interesting geometric relationship between
the frontier portfolios consisting entirely of the $n$ risky assets and the frontier portfolios which include the risk-free asset. In particular, when the risk-free rate is smaller than the expected return of the minimum variance portfolio $p^*$, $\rho_0 < A/C$, the half-line of efficient frontier portfolios defined by equation 8 intersects as a tangent line the hyperbolic curve of frontier portfolios. This has as an important financial consequence that all efficient frontier portfolios made of $n$ assets and a risk-free asset can be attained by a linear combination of the risk-free asset along and some frontier portfolio consisting only of the $n$ risky assets. This is referred to as the \textit{one fund theorem}. We remark that the condition $\rho_0 < A/C$ can be interpreted financially as asserting that the rate of return of the risk-free asset be less than the least risky portfolio consisting purely of the $n$ risk assets. This agrees with our intuition that an investor should be compensated for taking risks. We shall now show that the statements above indeed hold for the frontier portfolios constructed from the $n$ risky assets and the risk-free asset.

To find the frontier portfolio $a$ which gives the point of intersection with the hyperbolic frontier curve, we shall consider the zero-covariance portfolio $z(a)$ which has expected return $\mu_a = \rho_0$. This requires:

$$\mu^{(z(a))}_{p^+} = \frac{A}{C} - \frac{D/C^2}{\mu^{(a)}_{p^+} - A/C} = \rho_0. \hspace{1cm} (77)$$

Solving for $\mu^{(a)}_{p^+}$ gives:

$$\mu^{(a)}_{p^+} = \frac{A}{C} - \frac{D/C^2}{\rho_0 - A/C}. \hspace{1cm} (78)$$

The variance corresponding to this portfolio is then given by equation 73, which gives:

$$\sigma^2_{p^+} = \frac{D}{C^2} \left( \frac{1}{C(\rho_0 - A/C)^2} + \frac{C}{D} \right) \hspace{1cm} (79)$$

$$= \frac{D}{C^2} \left( \frac{D + C^2(\rho_0 - A/C)^2}{CD(\rho_0 - A/C)^2} \right) \hspace{1cm} (79)$$

$$= \frac{1}{C^2} \left( \frac{C^2(\rho_0^2 - 2A\rho_0 + A^2) + D}{C(\rho_0 - A/C)^2} \right) \hspace{1cm} (79)$$

$$= \frac{1}{C^2} \left( \frac{C\rho_0^2 - 2A\rho_0 + B}{(\rho_0 - A/C)^2} \right). \hspace{1cm} (79)$$

From equation 32 we have that $D = BC - A^2 \neq 0$. Dividing the numerator by one of the factors of $C$ we obtain

$$C\rho_0^2 - 2A\rho_0 + B = H \hspace{1cm} (80)$$

16
where $H$ was defined in equation 68.

Since $\rho_0 < A/C$ we have that

$$\sigma_{p^+}^{(a)} = \pm \frac{\sqrt{H}}{C|\rho_0 - \frac{A}{C}|}.$$  

(81)

We have now determined a point on the hyperbolic frontier curve which has a tangent line intersecting $(0, \rho_0)$. To validate the claims made above, we must show that this tangent line in fact coincides with the half-line of frontier portfolios which include the risk-free asset. In particular, we must check that the tangent line has slope $\sqrt{H}$.

The slope of the line passing through $(0, \rho_0)$ and $(\sigma_{p^+}^{(a)}, \mu_{p^+}^{(a)})$ is given by:

$$\frac{\mu_{p^+}^{(a)} - \rho_0}{\sigma_{p^+}^{(a)}}.$$  

(82)

We shall first compute:

$$\mu_{p^+}^{(a)} - \rho_0 = \frac{A}{C} - \frac{D/C^2}{(\rho_0 - \frac{A}{C})} - \rho_0$$  

(83)

$$= \frac{-1}{C(\rho_0 - \frac{A}{C})} \left( C\rho_0^2 - 2A\rho_0 + \frac{A^2 + D}{C} \right)$$  

$$= \frac{-1}{C(\rho_0 - \frac{A}{C})} (C\rho_0^2 - 2A\rho_0 + B)$$  

$$= \frac{-H}{C(\rho_0 - \frac{A}{C})}$$

where we have used that:

$$B = \frac{A^2 + D}{C}.$$  

(84)

and

$$H = C\rho_0^2 - 2A\rho_0 + B.$$  

(85)

We can now compute the slope using this and equation 79 to obtain:

$$\frac{\mu_{p^+}^{(a)} - \rho_0}{\sigma_{p^+}^{(a)}} = \frac{-H}{C(\rho_0 - \frac{A}{C})} = \sqrt{H}.$$  

(86)

This shows that the tangent line is the frontier curve for the portfolios constructed from the $n$ risky assets and risk-free asset. Thus the collection of all frontier portfolios consisting of $n$ risky assets and a risk-free asset can be obtained by investing in the risk-free asset along and the portfolio $a$, confirming the **one fund theorem**.
Conclusions

In these notes we have shown how a quantitative theory for portfolio management can be developed using the expected return to model an investor's financial objectives and variance to quantify riskiness associated with investments. To use this theory in practice requires a significant amount of information about the assets. For example, one must somehow decide what expected return to assign to a given asset and still more challenging how to estimate the variances and covariances of the assets. A natural approach would be to use the past history of an asset, but changing economic conditions may make this a poor indicator of future returns for many types of assets. For further discussion of these issues and how portfolio theory may be applied in practice see the reference *Investments* by Bodie, Kane, and Marcus and the other references.