IS STOCK PICKING DECLINING AROUND THE WORLD?*

Utpal Bhattacharya
Kelley School of Business
Indiana University
ubhattac@indiana.edu

Neal Galpin
Kelley School of Business
Indiana University
ngalpin@indiana.edu

November 15, 2005

Key words: modern portfolio theory; indexing; stock picking

JEL number: G11, G14, G15

* We are grateful for suggestions from Craig Holden, Chris Lundblad, and seminar participants at Indiana University.

IS STOCK PICKING DECLINING AROUND THE WORLD?

Abstract

We do three things in this paper. We first develop a metric to measure the maximum fraction of volume explained by stock picking in a market. We then use our metric to measure stock picking around the world. We find that though there is more stock picking in emerging markets than in developed countries, it is declining everywhere. In the United States, for example, stock picking has secularly declined from a high of 60% in the 1960s to a low of 24% in the 2000s. Finally, as markets cannot be efficient if everyone believes that they are efficient and, therefore, do no stock picking – the Grossman and Stiglitz (1980) paradox – we ask what is the long-run steady state fraction of stock pickers? We develop a simple theoretical model, and calibrate this model to the United States economy to conclude that stock picking will eventually settle at 11% of trading volume in the United States.

IS STOCK PICKING DECLINING AROUND THE WORLD?

I. INTRODUCTION

"A small gamble in a large number of different companies where I have no information to reach a good judgment, as compared with a substantial stake in a company where one's information is adequate, strikes me as a travesty of investment policy"

John Maynard Keynes, 1883-1946

John Maynard Keynes was a good stock picker. From 1928 to 1945, the fund he managed for King's College, Cambridge, produced positive returns at a time when the U.K. stock market was declining by 0.5% per year.[1] The intellectual foundations of stock picking were laid out in the classic text on valuation by Graham and Dodd (1934), who showed us how to figure out whether a stock was a "buy." Many of today's famous investors like Warren Buffett have been influenced by their theories. Indexing, which is the practice of passively investing in a portfolio containing a large number of stocks, is the philosophical opposite of stock picking. Instead of picking winners and losers, indexing emphasizes diversification. The intellectual foundations of indexing are in Markowitz's (1952) paper on modern portfolio theory and Tobin's (1958) paper on two-fund separation.[2] Indexing also has its fans in the investment world, of which perhaps the most influential is the Vanguard group of mutual funds. Ironically, the index funds that Vanguard popularized as an asset class now face serious competition from Exchange Traded Funds that exchanges have introduced to cash in on the popularity of passive indexing.[3]

The purpose of this paper is to find out which investment philosophy, stock picking or indexing, is dominant in the stock markets around the world. This is an important research question because, though there is much anecdotal evidence that the ideas in Markowitz (1952), a paper which led to the "birth of modern financial economics" (Rubinstein (2002)), have permeated investment practice, there has been, to the best of our

[1] This data and the quote above come from Wikipedia, an online user-contributed encyclopedia (http://www.answers.com/library/Wikipedia)

[2] The practice of diversification existed before Markowitz (1952). See Goetzmann and Ukhov (2005) for an empirical study of British overseas investments in 1870-1913.

[3] Wall Street Journal, November 5-6, 2005.

knowledge, no paper formally measuring this permeation.[4] Our paper makes a modest first attempt to formally measure how the investment community has accepted one of the ideas of modern portfolio theory – indexing.

The prevalence of indexing can only be measured if there exists a measure for indexing. No such measure exists in the literature. So the first part of our paper develops a metric for the opposite of indexing – stock picking. The idea behind this measure is inspired by a theoretical insight in Lo and Wang (2000). They proved that, if the two-fund separation theorem holds, dollar turnover of a stock, which is defined as the dollar volume of shares traded divided by the dollar market capitalization of the stock, should be identical for all stocks.

An empirical implication of the above theoretical insight is that if every person in the world indexes between a risk-free portfolio and the market portfolio (or a value-weighted portfolio that is a proxy for the market portfolio), trading volume in stock i should be explained *completely* by the market capitalization of stock *i*. This would mean that $(1-R^2)$ of the cross-sectional regression of the log of volume against the log of market capitalization would reflect the *deviation* from the indexing investment philosophy. This deviation will occur because some agents pick individual stocks. This deviation will also occur if some agents index to portfolios other than the value-weighted portfolio. This means that the $(1-R^2)$ of the above cross-sectional regression between log dollar volume and log dollar market capitalization will be a measure of the *maximum* proportion of trade that can be explained by stock picking.

We run these cross-sectional regressions every month, for every country for which we have data, for as long as we have the data. We plot the $(1-R^2)$ over time for 43 countries. We get two big results, and many small results.

Our first big result is that, on an average, there is more stock picking in emerging markets than in developed markets. As a matter of fact, the maximum fraction of volume explained by stock picking in emerging markets in the 1995-2004 period is 63%, whereas the maximum fraction of volume explained by stock picking in developed countries in the same period is only 45%. In our sample of 43 countries, the maximum fraction of

[4] Rubinstein (2002), in his retrospective of Markowitz's (1952) paper, states that "Markowitz's approach is now commonplace among institutional portfolio managers who use it both to structure their portfolios and measure their performance. It has been generalized and refined in innumerable ways, and is even being used to manage the portfolios of ordinary investors." The website of Yahoo Finance (http://biz.yahoo.com/edu/bi/ir_bi5.ir.html), a popular site for individual investors, states: "You can divide the history of investing in the United States into two periods: before and after 1952. That was the year that an economics student at the University of Chicago named Harry Markowitz published his doctoral thesis."

volume explained by stock picking in the 1995-2004 period is the least in the United States (29%) and is the most in China (80%).

Our second big result is that, on an average, stock picking is declining around the world. Of the 43 countries under investigation, we record that for 38 countries, the maximum fraction of volume explained by stock picking is lower in the last five years (2000-2004) than in the previous five years (1995-1999). The declines in stock picking are quite dramatic, especially in the emerging markets. The most dramatic decline in the popularity of stock picking is recorded in the United States, but that is probably because we have a longer time-series data for the United States. In the United States, the maximum fraction of volume explained by stock picking has secularly declined from a high of 60% in the 1960s to a low of 24% in the 2000s.

As we have a lot more data on the United States, we are able to get many small cross-sectional results for the United States. We find that though stock picking is less in S&P 500 stocks than in non S&P stocks, the difference seems to have disappeared in recent times. This fact shows that the though the actual mechanics of indexing in S&P 500 stocks is easier, the mechanics do not matter much anymore because indexing seems to be popular even in stocks not in the S&P 500 index. In terms of exchanges, there is more stock picking in AMEX than in NYSE. Nasdaq starts out looking like the AMEX, but it looks like the NYSE today. In terms of size, there is more stock picking in small stocks than in large stocks. In terms of age, there is more stock picking in young firms than in old firms. In terms of industries, stock picking is highest in the telecommunication industries, and lowest in the utilities industries. Our last cross-sectional result is that there is more stock picking in stocks that are covered by fewer analysts than in stocks that are covered by more analysts. As analysts are the quintessential stock pickers, it seems that investors who pick stocks avoid stocks that analysts pick. Finally, whatever the above cross-sectional results, we record that stock picking is declining over time in each and every category.

Our summary of findings from the first two parts of the paper is that stock picking is more pronounced in emerging markets than in developed markets, but it is declining in nearly all stock markets of the world. Though our paper is the first to formally document the declining popularity of stock picking around the world, indications that this may be happening are in a paper by Fernando et al (2003), who document the explosive growth of mutual funds around the world. Also, though our paper is the first paper to formally document the large cross-sectional variation in the popularity of stock picking across countries, indications that this may occur are in a paper by Khorana, Servaes and Tufano (2004), who document that the mutual fund industry is larger in countries with better laws and regulations and more wealth. Interestingly, these are the countries with the lowest stock picking in our sample. The above two arguments implicitly assume that mutual funds emphasize diversification over stock picking. That may be a reasonable assumption, but as Wermers (2000) shows, stock picking is alive and well in some United States mutual funds. Further, the explosive growth of hedge funds in recent years, also means that stock picking may be making a comeback.

Our results show that modern portfolio theory has won. However, it is premature to write the epitaph for stock picking. News of the death of stock picking will be an exaggeration. The reason no stock picking cannot be an equilibrium strategy is because of the Grossman-Stiglitz (1980) paradox: if no one picks stocks, information that stock pickers communicate with their trades cannot be impounded in prices, and so markets become inefficient, and so develops an opportunity to gather information, pick stocks, and make trading profits. This begs the question: how many stock pickers will exist in equilibrium? In other words, what is the long-run steady state fraction of stock pickers?

We develop a simple theoretical model. Our theoretical model is based on a crucial insight that comes from an early theoretical model by Treynor and Black (1973): in a mean-variance optimizing framework, even active stock pickers would like to maximize their Sharpe ratio, which is the ratio of the risk premium of the portfolio they choose divided by the standard deviation of the return of the portfolio (assumed to be the measure of risk). This implies that if we restrict an investor to be either a passive investor in the market portfolio or an active investor in a single stock, this would mean that this investor would be indifferent if the Sharpe ratios are the same. The Sharpe ratio of the market portfolio is simply the market price of risk. The active investor, who we allow to hedge the systematic risk of the single stock by taking an opposite position in the market portfolio, has the following Sharpe ratio. It is his excess profit from his superior information divided by the cost of active investing. The voluminous literature on market microstructure, which begins with Kyle (1985) and Glosten and Milgrom (1985), tells us that the excess profit an insider obtains from his inside information is his profit based on his superior information (Jensen's alpha) minus his adverse selection cost and other transaction costs of trading. We should also subtract his cost of obtaining the superior information from this number. Modern portfolio theory tells us that the cost of active investing is the exposure to idiosyncratic risk. Equating the two Sharpe ratios allows us to express the excess profit of the indifferent investor as a product of the market price of risk and the idiosyncratic risk.

Note that the excess profit that makes an investor indifferent between stock picking and passive indexing *is the product* of the market price of risk and the idiosyncratic risk. We call this value the "indifferent" excess profit. This implies that stock picking would become less popular if the market price of risk is increasing and/or idiosyncratic risk is increasing. The intuition is obvious. If the reward for holding market risk is increasing, investing in the market portfolio is more attractive, and so stock picking is less popular. If idiosyncratic risk is increasing, the cost of non-diversification, which is what stock picking entails, is increasing, and so stock picking is less popular.

The market price of risk is time-varying. See Lettau and Ludvigson (2003) for an excellent survey of this voluminous literature. There are many methods to estimate the market price of risk. We use the Whitelaw (1994, 1997) methodology as our primary method to estimate the market price of risk, though we report our estimates for the other methods as well. Idiosyncratic risk is also time-varying. As a matter of fact, a growing literature has documented that idiosyncratic risk has secularly increased over time in the United States (see Campbell et al. (2001)) and all over the world (see Morck, Yeung and Yu (2000). We use the method of Morck, Yeung and Yu (2000) to back out idiosyncratic risk for the United States. Our estimate of idiosyncratic risk tallies with their estimate. It also tallies with the estimate of Campbell et al. (2001), who use a variance decomposition method of estimating idiosyncratic risk. What is important, however, is that we get the same result: idiosyncratic risk is increasing in the United States.

We then estimate the "indifferent" excess profit every year, which makes an investor indifferent between stock picking and passive indexing, as a product of the estimated market price of risk every year and the estimated idiosyncratic risk every year. We find that this "indifferent" excess profit is *increasing* over time. If we assume that every agent in the economy has a different excess profit from stock picking, which depends on skill and/or luck, it is reasonable to assume that this excess profit is drawn from a distribution that is stable over time. The people who pick stock have a draw that is *higher* than the "indifferent" excess profit. So, if the "indifferent" excess profit is *increasing* over time, this would imply that the proportion of stock pickers is *declining*.

We, therefore, conclude that stock picking is declining in the United States, because the total cost of stock picking is increasing in the United States. This total cost is increasing because the direct cost of stock picking – idiosyncratic risk – is increasing in the United States, though the indirect opportunity cost of stock picking – the forgone market reward for risk – does not seem to have a trend.

As total risk is not changing (see Schwert (1989) and Campbell et al (2001)), idiosyncratic risk cannot increase without bound. This suggests that idiosyncratic risk may asymptote to a steady-state. As the market price of risk seems not to have a trend (see Lettau and Ludvigson (2003)), the product of idiosyncratic risk and the market price of risk – the "indifferent" excess profit curve -- may asymptote as well. Therefore, if we assume that the distribution of excess profits in the economy is stable, then the proportion of investors whose excess profit is above the "indifferent" excess profit curve – the stock pickers – will stabilize where the "indifferent" excess profit curve asymptotes. This is the steady-state proportion of stock pickers. For the United States, our model estimates that stock picking will eventually settle at 11% in the United States.

The paper is organized as follows. We develop a metric for stock picking in section II. Section III describes our data. Section IV is the main result of this paper. It documents that there is more stock picking in emerging markets than in developed countries, but it is declining everywhere. As we have more data on the United States, section V covers the United States in greater detail, and reports results for different categorizations of stocks. Section VI explores the steady state proportion of stock pickers, and due to data availability, we restrict ourselves to the United States Section VII concludes. In this section, we discuss extensions of our paper, of which the most important extension is a deeper exploration of *why* there is so much cross-sectional variation in stock picking.

II. A METRIC FOR STOCK PICKING

Lo and Wang (2000) proved that, if the two-fund separation theorem holds, dollar turnover of a stock, which is defined as the dollar volume of shares traded divided by the market capitalization of the stock, should be identical for all stocks. The intuition behind their result was simple. If an agent just invests in the risk-free asset and the market portfolio, and if the market portfolio has $85 million of stock A and $15 million of stock B, then if the agents buys (sells) $100 of the market portfolio, she buys (sells) $85 of stock A and $15 of stock B. If prices do not change between trades, the share turnover of a stock, which is defined as volume of shares traded divided by the number of shares outstanding of the stock, should also be identical for all stocks if the two-fund separation theorem holds.

We run the following cross-sectional regression model every month for every market

$$\ln(\text{volume of shares})_i = a + b \ln(\text{number of shares outstanding})_i + \epsilon_i \qquad (1)$$

where

Volume of shares_i is the monthly trading volume of stock i,

Number of shares outstanding_i is the number of shares outstanding at the end of the month for stock i, and

ϵ_i is the volume of shares of stock *i* that cannot be explained by the number of shares outstanding at the end of the month for stock *i*.

If every person in the world indexes between a risk-free portfolio and the market portfolio (or a value-weighted portfolio that is a proxy for the market portfolio) with little error, trading volume in stock *i* should be explained *completely* by the market capitalization of stock *i*. So we should obtain the following estimates in our above regression equation:

a = ln (turnover),

b = 1, and

$R^2 = 1$.

If we do not obtain 1 as our estimate for R^2 in the above regression, this would mean that there is *deviation* from the indexing investment philosophy. This deviation will occur because some agents pick individual stocks. This deviation will also occur if some agents index to portfolios other than the market portfolio. As a matter of fact, for K factors, as Lo and Wang (2000) showed, indexing will occur in K funds. This means that the ($1\text{-}R^2$) of the above cross-sectional regression between log volume and log number of shares outstanding will be a measure of the *maximum* proportion of trade that can be explained by stock picking.

So our metric for stock picking every month in a market is the ($1\text{-}R^2$) of a cross-sectional regression of log volume of stock *i* against the number of shares outstanding of stock *i* in that market. To be precise, the estimate of ($1\text{-}R^2$) gives us the *maximum* proportion of trade that can be explained by stock picking that month in a market.

This ($1\text{-}R^2$) metric has the following advantages. First, it is simple to estimate. Second, the data requirement is minimal. Volume and shares outstanding data are publicly available for nearly all stock markets of the world. Third, the definition of a market is flexible. Markets could be the various country stock markets, which would allow us to compare stock picking across the world. Markets could be local markets within a country, like NYSE, AMEX or Nasdaq, which would allow us to compare stock picking across these local markets. Markets could be defined by different types of stock categorization like size, age of firm, industry, etc., which would allow us to compare stock picking across different sizes, different ages, different industries, etc. The fourth and the biggest advantage of this metric is that the cross-sectional regression can be estimated at different points in time, which would allow us to detect time-trends, if any, in the popularity of stock picking.

This ($1\text{-}R^2$) metric has the following disadvantage. It does not give us an estimate of the proportion of trade that can be explained by stock picking; it gives us an estimate of the *maximum* proportion of trade that can be explained by stock picking. Our estimate of stock picking is, therefore, biased upward. We can decrease this bias by introducing additional independent variables in our cross-sectional regression. These variables could be factors that mimic possible factor portfolios that investors also index to, or it could be variables that have appeared in the volume literature (see Lo and Wang (2000) for a description). These additional explanatory variables would reduce $1\text{-}R^2$, thus reducing the bias in our estimate. We, however, refrain from this exercise for two reasons. First, it complicates a simple metric. Second, as these additional independent variables are quite ad hoc and are not motivated by well-established theories, we would not know how to interpret our new results.

III. DATA

Broadly speaking, the data for this paper can be classified into three types. First, we require basic stock-level variables to estimate the maximum fraction of trade explained by stock picking. Second, we require other variables that help us explore the determinants of stock picking. Third, we use macroeconomic variables to estimate the long-run fraction of volume explained by stock picking.

A. Basic variables

Estimating the maximum fraction of volume explained by stock picking requires only two data items: the share volume and the number of shares outstanding. These data items are readily available from well-known sources for most stocks.

For United States firms, we collect share volume and shares outstanding data from CRSP. Share volume is the number of shares traded within a given month. The number of shares outstanding is the number of shares of common equity outstanding as of the end of the month. In order to be included in the final sample, a stock must be listed on the NYSE, AMEX, or Nasdaq. In addition, the share must be an ordinary common share; we exclude, for example, ADRs and REITs from the analysis.[5] Since share volume is not available before July of 1962, this is where we begin our analysis for the United States.

For non-United States firms, we collect share volume and shares outstanding at the individual stock level from Datastream. As we did for United States firms, we calculate the volume of shares traded within a given month. Again, we measure shares outstanding at the end of each month. Datastream provides lists of stocks associated with a country.[6] Within these lists, some stocks have data collected from an exchange in a given country, but have another home country. Other stocks in these lists have a given country listed as their home

[5] Share codes 10 and 11 in CRSP.

[6] These lists are typically named FXX and DEADXX, where XX is an abbreviation for the country name. There are a few exceptions to this rule. The FXX file includes firms maintaining Datastream coverage as of 2004, while the DEADXX file includes firms that have in the past been covered by Datastream, but for whom coverage has ceased.

country, but have data collected on an exchange in another country. We include in a country only those firms both headquartered and listed on an exchange within that country.[7]

The overall coverage by Datastream improved dramatically in the mid 1990s (see, also, Morck, Yeung, and Yu, 2000), though some countries have data available well before this time. Therefore, we focus on the period from 1995 through 2004 for our cross-country analysis.

B. Other variables

For the country cross-sectional analysis, we compare stock picking in emerging markets versus developed markets. We classify a market as developed if MSCI has a developed market index in that country. We classify any markets not on this list as emerging.[8] The total number of countries in our analysis is 43, of which 21 are classified as developed countries, and the remaining 22 are classified as emerging markets. A list of these 43 countries appears in Table 1.

In the United States, we estimate the maximum fraction of volume explained by stock picking for stocks with different characteristics. We first estimate the amount of stock picking based on whether the firm is included in the S&P 500 index. Second, we estimate stock picking based on the exchange on which the firm's equity is listed. Third, we break firms into size quintiles based on market capitalization to estimate the effect of firm size on stock picking. Fourth, we break firms into quintiles according to age, which we proxy for using the number of months for which the firm has had data available in CRSP. Fifth, we consider the industry effect on stock picking. We classify firms into industries based on the ten industry classifications provided by Fama and French (1997).[9] Except for the last variable, the other data items are available directly from CRSP.

In addition to the above variables that are available from CRSP, we classify firms into quintiles according to the number of analysts following the firm. We obtain this by counting the number of analysts making annual EPS forecasts made in the month prior to the firm's fiscal year end. This timing convention allows us to capture

[7] We use the Datastream variable "GEOGC" to determine the home country of a company and "EXNAME" to determine the exchange from which the volume is reported.
[8] We cross-checked this definition of emerging with countries explicitly labeled as emerging by MSCI. The only country that does not explicitly appear is Cyprus, which we classify as emerging.
[9] The industry classifications are available from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

the maximum number of analysts making forecasts about the firm, as more analysts make earnings forecasts as the forecast period end date nears. Data on analyst following comes from I/B/E/S and becomes available in 1976.

C. Macroeconomic variables used for estimating the steady state stock-picking measure

We follow the model of Whitelaw (1994, 1997), and model the expected market risk premium as a function of three macroeconomic variables. First, we use the yield spread between Baa rated companies and Aaa rated companies to proxy for default risk premia. Second, we include the three-month treasury yield. Third, we include the dividend yield on the S&P composite index. Global Insight provides these data items.

As in Whitelaw (1994, 1997), we also model the volatility of market returns based on two macroeconomic variables. We again include the three-month treasury yield. In addition, we include the spread between yields on commercial paper and the three-month treasury yield. We collect this series from the Federal Reserve Board of Governors.

IV. STOCK PICKING IN THE WORLD

Every month, beginning from January of 1995 and ending in July of 2004, for each country, we run a simple OLS regression of log monthly stock volume (measured by number of shares traded in that month for stock *i* in a country) against the log of monthly shares outstanding (measured by number of shares outstanding in the beginning of the month for stock *i*). We run this regression for 43 countries, of which 21 are classified as developed markets and 22 are classified as emerging markets. We begin in January of 1995 because that is when we can obtain data for all of our countries. Since we have data for the United States beginning July 1962, we study the United States in greater detail in the next section.

Each month, we find the median $(1-R^2)$ for the 21 developed countries to obtain a $(1-R^2)$ for the developed markets as a group, and the median $(1-R^2)$ for the 22 emerging markets to obtain a $(1-R^2)$ for the emerging markets as a group. Figure 1 plots the $(1-R^2)$ of this cross-sectional regression over time for developed markets and emerging markets. Figure 1 depicts the two big results of our paper.

The first big result that Figure 1(a) illustrates is that there is more stock picking in emerging markets than in developed markets. Panel A of Table 1 confirms this. The maximum fraction of volume explained by stock picking in emerging markets in the 1995-2004 period is 63%, whereas the maximum fraction of volume explained by stock picking in developed countries is only 45%.

Figure 1(b) depicts stock picking for a few select markets: two developed markets (Japan and the United States) and two emerging markets (Mexico and Taiwan). As can be seen, the maximum fraction of volume explained by stock picking in the two developed markets is less than the maximum fraction of volume explained by stock picking in the two emerging markets in the 1995-2004 period. Panel B of Table I confirms this. The maximum fraction of volume explained by stock picking in United States, Japan, Taiwan and Mexico in the 1995-2004 period is 29%, 41%, 61% and 72% respectively. Panel B of Table I sorts the 43 countries under investigation in ascending order of the popularity of stock picking in 1995-2004. As can be seen, the maximum fraction of volume explained by stock picking in the 1995-2004 period is the least in the United States (29%) and is the most in China (80%). As can also be seen, popularity of stock picking is generally lower in the developed markets (they usually are on the top of the list) than in the emerging markets (they usually are on the bottom of the list). Germany is an important exception because, considering that it is a developed market, it has too much stock picking. Russia is also an important exception because, considering that it is an emerging market, it has too little stock picking.

The second big result that Figure 1(a) illustrates is that, on an average, stock picking is declining around the world. The declines in stock picking are quite dramatic, especially in the emerging markets. Panel A of Table 1 shows that the maximum fraction of volume explained by stock picking fell from 72% in 1995-1999 to 59% in 2000-2004 in emerging markets, whereas the maximum fraction of volume explained by stock picking fell from 51% in 1995-1999 to 41% in 2000-2004 in developed markets.

Figure 1(b) depicts the decline in stock trading for a few select markets: two developed markets (Japan and the United States) and two emerging countries (Mexico and Taiwan). Panel B of Table I confirms this. The maximum fraction of volume explained by stock picking in United States, Japan, Taiwan and Mexico in the earlier 1995-2000 period is 36%, 43%, 66% and 76% respectively, whereas the maximum fraction of volume explained by stock picking in United States, Japan, Taiwan and Mexico in the later 2000-2004 period is 24%, 40%, 59% and 67% respectively.

Panel B of Table 1 shows that, of the 43 countries under investigation, the maximum fraction of volume explained by stock picking is lower in the last five years (2000-2004) than in the previous five years (1995-1999) for 38 countries. The most dramatic decline in the popularity of stock picking is recorded in the United States, but that is probably because we have a longer time-series data for the United States. In the United States, the maximum fraction of volume explained by stock picking has secularly declined from a high of 60% in the 1960s to a low of 24% in the 2000s.

V. STOCK PICKING IN THE UNITED STATES (DETAILS)

Every month, beginning from July of 1962 and ending in December of 2004, for the United States, we run a simple OLS regression of log monthly stock volume (measured by number of shares traded in that month for stock *i*) against the log of monthly shares outstanding (measured by number of shares outstanding in the beginning of the month for stock *i*). Figure 1 plots the $(1\text{-}R^2)$ of this cross-sectional regression over time. As the number of listed stocks change over time, the sample sizes of the cross-sectional regressions change over time. To correct for this, we run the same cross-sectional regression for 400 stocks randomly selected each period. As can be seen in Figure 2, this does not make much difference.

Figure 2 is the main result for the United States. It shows that the maximum fraction of volume explained by stock picking has secularly declined over time. Stock picking could explain a maximum of 60% of volume in the late 1960s, but it could explain only a maximum of 24% of volume in the 2000s. This reflects a huge decline in the popularity of stock picking in the last forty years.

We now run the cross-sectional regression for various sub-samples of the data. Figure 3 plots the $(1\text{-}R^2)$ of these cross-sectional regressions over time for the various sub-samples. These are depicted in various panels in Figure 3. The objective of the various panels of Figure 3 is to graphically explore which markets exhibited the greatest decline in stock picking. Figure 3 is an informal eyeball test. The corresponding panels of Table 2 show formal statistical tests that confirm our informal eyeball tests from Figure 3.

Figure 3(a) shows that there is more stock picking in the non-S&P 500 stocks than in the S&P 500 stocks at nearly all points in time. This is to be expected, because belief in modern portfolio theory exhibits itself through the practical technique of indexing, and indexing is far easier for the S&P 500 stocks. What may be surprising, however, is that indexing is becoming popular over time not only in the S&P 500 stocks, but even in the non-S&P 500 stocks. As a matter of fact, in recent times, it seems that there is more indexing in the non-S&P 500 stocks than in the S&P 500 stocks!

Panel A of Table 2 confirms the above conclusion. The $(1-R^2)$ of the S&P 500 stocks is statistically significantly lower than the $(1-R^2)$ of the non-S&P 500 stocks. The test we use is to sort the monthly $(1-R^2)$ of the OLS regressions of each of the two sub-samples, and test which group has a higher median. We do the same test for the intercept term as well as the coefficient term. Panel A of Table 2 shows some interesting tangential results. As the intercept term is the log of the turnover, it shows that there is more trade in S&P 500 stocks than in non-S&P 500 stocks. As the coefficient term tells us whether volume is linear in shares outstanding – a coefficient of 1 means linearity – it does seems that volume is roughly proportional to shares outstanding for the S&P 500 stocks, but it seems that there is higher turnover for the larger stocks in the non-S&P 500.

Figure 3(b) examines stock picking among the United States exchanges. It shows that there is less stock picking in the NYSE than in the AMEX at all points in time. The case of Nasdaq is fascinating. In terms of the popularity of stock picking, it starts out looking like the AMEX in the 1980s, and ends up looking like the NYSE in the 2000s. The dramatic change occurs in the year 2000. This is the year that the Nasdaq bubble burst and, we believe, day traders and other investors who were stock pickers in Nasdaq became disillusioned. Finally, it should be noticed that stock picking is becoming less popular over time for all the three markets.

Panel B of Table 2 confirms the above conclusion. The $(1-R^2)$ of the AMEX stocks is statistically significantly higher than the $(1-R^2)$ of the Nasdaq stocks, which in turn is statistically significantly higher than the $(1-R^2)$ of the NYSE stocks. The intercept terms in Panel B of Table 2 shows that turnover is highest in NYSE and lowest in AMEX. The coefficient terms in Panel B, as they are all above 1, shows that there is a higher turnover in the larger stocks in all the three markets.

As stock picking was less popular for the larger S&P 500 stocks than for the smaller non-S&P 500 stocks (Figure 3(a) and Panel A of Table 2), and as stock picking was less popular for the larger NYSE stocks than for the smaller AMEX stocks (Figure 3(b) and Panel B of Table 2), we suspect a "size effect." Figure 2(c) examines stock picking among various sizes. We measure size every month by sorting the number of shares outstanding that month into five quintiles. Figure 3(c) shows that there is less stock picking in the largest quintile of stocks than in the smallest quintile of stocks most of the time. The result is, however, not unambiguous. There was more stock picking in the largest quintile stocks in the early 1960s. This finding plus the finding that stock picking in Nasdaq and NYSE are roughly similar in recent times (Figure 1(b) and Panel B of Table 1) hint to us that the "size effect" is not the only story. Finally, it should be noticed that stock picking is becoming less popular over time for all size stocks.

Panel C of Table 2 confirms the above conclusion. The $(1\text{-}R^2)$ of the smallest quintile stocks is statistically significantly higher than the $(1\text{-}R^2)$ of the largest quintile stocks. The intercept terms in Panel C of Table 2 shows that turnover is higher in the largest quintile stocks than in the smallest quintile stocks. The coefficient terms in Panel C shows that, among the smallest quintile stocks, there is a higher turnover in the larger stocks, but this "size effect" is non-existent in the largest quintile stocks.

As information about young firms is difficult to come by, it would seem to suggest that finding out information about young firms and trading on that information would be more profitable than stock picking older firms. We measure age every month by sorting the ages of firms that month into five quintiles. Figure 3(d) shows that there is less stock picking in the oldest quintile of stocks than in the youngest quintile of stocks all the time. Finally, it should be noticed that stock picking is becoming less popular over time for firms of all ages.

Panel D of Table 2 confirms the above conclusion. The $(1\text{-}R^2)$ of the youngest quintile stocks is statistically significantly higher than the $(1\text{-}R^2)$ of the oldest quintile stocks. The intercept terms in Panel D of Table 2 shows that turnover is higher in the youngest quintile stocks than in the oldest quintile stocks. The coefficient terms in Panel D shows that there is a "size effect" in firms of all ages: there is a higher turnover in the larger stocks.

Industry is an important element in the money management industry. Specialization occurs by industry. So we do an industry classification, and ask whether stock picking is more popular in some industries. Our priors are the same as before. As information about new industries is difficult to come by, it would seem to suggest that finding out information about new industries and trading on that information would be more profitable than stock picking in older "boring" industries. We use a 10-industry classification based on Fama and French (1997). Figure 3(e) confirms our intuition. It shows that the "exciting" telecommunication industry has the highest amount of stock picking, whereas the "boring" utilities industry has the least amount of stock picking. Finally, it should be noticed that stock picking is becoming less popular over time for all industries.

Panel E of Table 2 shows a ranking for all industries. The ranking, from the industry where stock picking is least popular to the industry where stock picking is most popular is: utilities, medical, high tech, manufacturing, durables, others, non-durables, retail/wholesale, energy and telecommunications. The intercept terms in Panel E of Table 2 shows that turnover is generally lower in the industries where there is less stock picking. The coefficient terms in Panel E show that though there is a "size effect" in stocks of nearly all industries – there is a higher turnover in the larger stocks in the industry – there seems to be a "reverse size effect" in the telecommunication industry – there is a smaller turnover in the larger stocks in this industry.

The quintessential stock picker in the money management industry is the research analyst. He gets paid a lot to pick stocks. So it would seem natural that we should end this section by classifying stocks by the number of analysts who follow them. Our priors are that analysts pick stocks that others pick later. We measure the number of analysts following a stock every month by sorting the number of analysts following the stock that month into quintiles. Figure 3(f) shows that our priors are wrong: it seems that investors do more stock picking in stocks that analysts do not pick most of the time, and investors do less stock picking in stocks that analysts do pick most of the time. A hypothesis to explain this seemingly counter-intuitive finding is that analysts improve the information environment of the stocks they are picking, thus, paradoxically, reducing the return to information gathering and stock picking for themselves as well as for others. Serious stock pickers, therefore, avoid the stocks that analysts pick. To the best of our knowledge, this is the first time someone has provided empirical evidence in favor or against this hypothesis. Our evidence also suggests that stock picking is becoming less popular over time for all stocks, whether these stocks are followed by a large number of analysts or followed by a small number of analysts.

Panel F of Table 2 confirms the above conclusion. The $(1-R^2)$ of the stocks that are covered by a small number of analysts is statistically significantly higher than the $(1-R^2)$ of the stocks that are covered by a large number of analysts. The intercept terms in Panel F of Table 2 shows that turnover is much higher in the stocks that are covered by a large number of analysts than in the stocks that are covered by a small number of analysts. The coefficient terms in Panel E show that though there is a "size effect" in stocks which are covered by few analysts – there is a higher turnover in the larger stocks in this sub-group – there seems to be no "size effect" in stocks that a large number of analysts follow.

VI. STOCK PICKING IN THE STEADY STATE IN THE UNITED STATES

A. A simple model

An economy consists of a large number of assets whose returns are risky, and one single risk-free asset with a one-period return of r_f. The one-period return of a "representative" risky asset *i* is

$$\tilde{r}_i = E(r_i) + \tilde{\delta} + \tilde{\alpha}_i + \tilde{\varepsilon}_i \qquad (2),$$

and the one-period return of the market portfolio is

$$\tilde{r}_m = E(r_m) + \tilde{\delta} \qquad (3),$$

where

δ = market risk component, with mean = 0 and standard deviation = σ_m ,

α_i and ϵ_i = idiosyncratic risk component, with means equal to 0 and standard deviations equal to σ_α and σ_ϵ respectively.

Investors have a one-period investment horizon. They care only about the mean return (which they like) and the standard deviation of the return (which they dislike) of their portfolios. This implies that every investor tries to maximize the Sharpe ratio of her portfolio. An investor can choose to be a passive investor who invests in the market portfolio and in the risk-free asset, or an active investor who picks the "representative" risky asset *i*.

If the investor chooses to be a passive investor, her Sharpe ratio is

$$S_{passive} = \frac{(E(r_m) - r_f)}{\sigma_m} \qquad (4)$$

If the investor chooses to be an active investor, she obtains α_i, which is a private signal. She then takes a position in the "representative" risky asset and an opposite position in the market portfolio. This allows her to get rid of the market risk component, δ, of the "representative" risky asset. However, she cannot get rid of the firm-specific risk component, ϵ_i. Moreover, she cannot fully exploit her inside information, α_i. This is because the voluminous literature on market microstructure, which begins with Kyle (1985) and Glosten and Milgrom (1985), tells us that the excess profit an insider obtains from her inside information, is her profit based on her superior information minus her adverse selection cost. There are also other transaction costs of trading like commissions. We subtract all these costs from α_i to obtain the profit from obtaining inside information, α_{net}. The Sharpe ratio of the active investor is then

$$S_{active} = \frac{\alpha_{net}}{\sigma_{\varepsilon}} \tag{5}$$

We now obtain the α_{net} that makes (4) equal to (5).

$$\alpha_{net,indifferent} = \frac{(E(r_m) - r_f)}{\sigma_m} \sigma_{\varepsilon} \tag{6}$$

This implies that any investor who obtains the α_{net} that is higher (lower) than $\alpha_{net,\ indifferent}$ will choose to be an active (passive) investor. Now assume that every agent in the economy has a different α_{net}, which depends on skill and/or luck, and this α_{net} is drawn from a distribution that is stable over time. This additional assumption implies that if $\alpha_{net,\ indifferent}$ is *increasing* over time, the proportion of stock pickers is *declining*.

We now go on to check whether $\alpha_{net,\ indifferent}$ is really *increasing* over time. As $\alpha_{net,\ indifferent}$ from (6) is the product of the Sharpe ratio of the market and the idiosyncratic risk of the "representative" stock, we need to estimate these two variables from United States data.

B. Estimation

We first estimate the average idiosyncratic risk of stocks in CRSP. For each stock each month, we fit a market model to monthly returns using the trailing 5 years of data. We then calculate the standard deviation of the residuals from these regressions to estimate the idiosyncratic risk. The overall idiosyncratic risk measure is the equal-weighted average of the individual stock idiosyncratic risk. This is the same method used in Morck, Yeung and Yu (2000) to back out idiosyncratic risk. Our estimate of idiosyncratic risk tallies with their estimate. It also tallies with the estimate of Campbell et al. (2001), who use a variance decomposition method of estimating idiosyncratic risk.

Figure 4 shows that idiosyncratic risk is increasing in the United States. Morck, Yeung and Yu (2000) as well as Campbell et al. (2001) have documented this as well. Although the fact that idiosyncratic risk, or firm-specific risk, is increasing over time is not disputed, the jury is still out on why this is so.[10] Figure 4 also shows the market model R^2. Again, as seen in Morck, Yeung and Yu (2000), this is declining over time. This should not be surprising. If idiosyncratic risk is increasing, the explanatory power of the market factor to explain firm-specific return should be declining. Figure 4 also shows the (1- R^2) of our cross-sectional regression. It shows that stock picking is becoming less popular when idiosyncratic risk is rising.

We next estimate the market price of risk, the market Sharpe ratio. We use the methodology described in Whitelaw (1994, 1997). Whitelaw suggests fitting expected returns and standard deviations as a function of macroeconomic variables using GMM. The moment conditions are

$$\mathrm{E}\begin{bmatrix} (r_{m,t+1} - \mathbf{X}_t\beta)\mathbf{X}_t' \\ \left(\sqrt{\frac{\pi}{2}} \times \left|(r_{m,t+1} - \mathbf{X}_t\beta)\right| - \mathbf{X}_t\theta\right)\mathbf{X}_t' \end{bmatrix} = 0 .$$

We estimate the market Sharpe ratio using the predicted values of expected market excess return and volatility from this two equation system. We include the spread between Baa and Aaa rated firms' yields, the three-month treasury yield, the spread between commercial paper and the treasury yield, and the dividend yield as instruments.

[10] Morck, Yeung and Yu (2000) investigate three hypotheses to explain why idiosyncratic risk is increasing over time. The first hypothesis is the obvious one, and this is the hypothesis they find the least evidence for: correlations among firm-specific cash flows are becoming weaker. The other two hypotheses are based on the dark side of finance. They surmise that as property rights strengthen, risk arbitrage strengthens, and so more firm-specific information is incorporated in stock prices. Their last hypothesis is that as firm-specific information becomes more credible due to more credible accounting statements, more firm-specific information is incorporated in stock prices. Campbell et al. (2001), on the other hand, do not take a particular view as to what is causing the increase in volatility. They lay down a list of reasons that, in their view, are plausible. Here is the list. First, conglomerates are breaking up into single-unit firms. Second, initial public offerings come earlier, when firm-specific information is still very imprecise. Third, stock options may tempt managers from increasing firm-specific risk. Fourth, there are the reasons expounded by Morck, Yeung and Yu (2000). Fifth, it could be day traders.

We then construct a time-series of estimated $\alpha_{net, indifferent}$ as the product of the estimated idiosyncratic risk and the estimated market Sharpe ratio. We are interested in the long-run trend of this series toward an asymptote. This is the reason why. If we assume that the distribution of excess profits in the economy is stable, then the proportion of investors whose excess profit is above the "indifferent" excess profit curve – the stock pickers – will stabilize where the "indifferent" excess profit curve asymptotes. This is the steady-state proportion of stock pickers.

We estimate the asymptote of $\alpha_{net, indifferent}$ by choosing estimates of a_0 and a_1 to minimize the function

$$\sum_t |\alpha_{net,indifferent} - (a_0 - e^{a_1 \times t})|.$$

We find the estimates of a_0 and a_1 to be 1.007 and -5.0e^{-6}, respectively. This tells us that our long-run estimate of $\alpha_{net, indifferent}$ is 1.007.

We now need to estimate the relationship between $\alpha_{net, indifferent}$ and the percent of volume explained by stock picking. The volume explained by stock picking (indexing) represents the percent of investors with α_{net} above (below) $\alpha_{net,indifferent}$. We assume that investors' α_{net} are drawn from a normal distribution with parameters μ and σ. This means that the R^2 from the regression of share volume on shares outstanding each month is the cumulative normal density function evaluated at $\alpha_{net, indifferent}$. We estimate the values of μ and σ to minimize the function

$$\sum_t |R_t{}^2 - N(\alpha_{net,indifferent}, \mu, \sigma)|,$$

where $N(x, \mu, \sigma)$ is the cumulative density function for a normal distribution with mean μ and deviation σ evaluated at x. Our estimates of μ and σ are -0.307 and 1.064, respectively. Since μ is negative, it implies that the average person loses if he picks stocks.

Now that we know all the parameters of the normal distribution that α_{net} is drawn from, we calculate the cumulative density function of this normal distribution at our long-run estimate of $\alpha_{net, indifferent}$, which had turned out to be 1.007. This is our long-run estimate of R^2, which turns out to be 0.892.[11] So (1- R^2) is roughly 11%. This implies that our estimate of the steady-state fraction of stock pickers in the United States is 11%.

Figure 5 shows the actual (1-R^2) of our cross-sectional regression. It also shows the in-sample predicted (1-R^2), the out-of-sample forecasted (1-R^2), and the estimated steady state (1-R^2). The last three variables are 1 minus $N(x,\mu,\sigma)$ (the cumulative density function for a normal distribution with mean μ and standard deviation σ evaluated at x). The value of x to obtain the in-sample predicted (1-R^2), to obtain the out-of-sample forecasted (1-R^2), and to obtain the estimated steady state (1-R^2) is the value of the in-sample $\alpha_{\text{net,indifferent}}$, the value of the out-of-sample forecasted $\alpha_{\text{net,indifferent}}$, and the value of the estimated steady state $\alpha_{\text{net,indifferent}}$, respectively.

As a robustness check, we use three alternative methods to estimate the market price of risk. All of these methods assume that the market price of risk is constant. The first method is the simplest. We calculate the average market excess return, and divide this average by the standard deviation of market returns. The second and third methods are from Harvey (1989), who estimates the market price of risk using GMM. The two methods differ only based on their definition of risk. The moment conditions for the second method, in which the absolute value of residuals measure risk, are:

$$\mathrm{E}\begin{bmatrix} (r_{m,t+1} - \mathbf{X}_t\beta)\mathbf{X}_t' \\ \left(r_{m,t+1} - \lambda\left|r_{m,t+1} - \mathbf{X}_t\beta\right|\right)\mathbf{X}_t' \end{bmatrix} = 0\,.$$

The moment conditions for the third method, in which the squared value of residuals measure risk, are:

$$\mathrm{E}\begin{bmatrix} (r_{m,t+1} - \mathbf{X}_t\beta)\mathbf{X}_t' \\ \left(r_{m,t+1} - \lambda(r_{m,t+1} - \mathbf{X}_t\beta)^2\right)\mathbf{X}_t' \end{bmatrix} = 0\,.$$

We compare the results from these three alternative methods with the results from the Whitelaw (1994, 1997) model in Table 3. The properties of $\alpha_{\text{net,indifferent}}$ are quite similar under all of the measures. The $\alpha_{\text{net,indifferent}}$ asymptotes to just over 1 using all of the four models. The biggest differences come in the estimate of the mean of the normal distribution, which is what causes the difference in the estimates. Since the average of the gross Jensen's alpha for all stocks is likely to be zero, the average of the net Jensen's alpha for all stocks is likely to be

[11] N(1.007, -0.307, 1.064)=0.892

negative. As this average is negative for the Whitelaw (1994, 1997) model, we report primarily the result from this model in our text, though Table 3 reports all results from all our four models.

VII. CONCLUSION

We do three things in this paper. We first develop a metric to measure the maximum fraction of volume explained by stock picking in a market. We then use our metric to measure stock picking around the world. We find that though there is more stock picking in emerging markets than in developed countries, it is declining everywhere. In the United States, for example, stock picking has secularly declined from a high of 60% in the 1960s to a low of 24% in the 2000s. Finally, we ask what is the long-run steady state fraction of stock pickers? We develop a simple theoretical model, and calibrate this model to the United States economy to conclude that stock picking will eventually settle at 11% in the United States.

Though this paper shows cross-sectional variation in stock picking across markets, it does not formally investigate hypotheses that may explain this cross-sectional variation. Neither does it formally investigate hypotheses that explain why stock picking is declining over time. An overall theme that comes through from our descriptive statistics is that stock picking is more in markets where there is less public disclosure of stock-specific information. It may be that as this disclosure is improving over time, stock picking is becoming less profitable and, therefore, declining in popularity. The above are just conjectures. We leave it to future research to formally explore the determinants of stock picking.

Another area of research is further development of our metric. The firm-specific error in our cross-sectional regression is really a measure of abnormal volume at the firm-level. Is abnormal volume at the firm-level linked to some firm-specific variables like size or book-to-market? Can abnormal volume predict returns? We leave these interesting questions to future research.

REFERENCE

Campbell, John, Martin Lettau, Burton Malkiel, Yexiao Xu, 2001, Have individual stocks become more volatile?: An empirical exploration of idiosyncratic risk, *Journal of Finance* 56, 1-43.

Graham, Benjamin and David Dodd, 1934, *Security Analysis: Principles and Technique*, McGraw-Hill, Columbus, OH.

Fama, Eugene and Kenneth French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153-193.

Fernando, Deepthi, Leora Klapper, Victor Sulla and Dimitri Vittas, 2003, The global growth of mutual funds, World Bank Policy Research Working Paper 3055.

Goetzmann, W., and A. Ukhov, 2005, British investment overseas 1870-1913: A modern portfolio theory approach, working paper 11266, National Bureau of Economic Research.

Glosten, Lawrence, and Paul Milgrom, 1985, Bid, ask and transaction prices in a specialist model with heterogeneously informed traders, *Journal of Financial Economics* 14, 71-100.

Grossman, Sanford and Joseph Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393-407.

Harvey, Campbell, 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289-317.

Khorana, Ajay, Henri Servaes and Peter Tufano, 2004, Explaining the size of the mutual fund industry around the world, working paper, Harvard Business School.

Kyle, Albert, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Lettau, Martin and Sydney Ludvigson, 2003, Measuring and modeling variation in the risk-return tradeoff, working paper, New York University.

Lo, Andrew and Jiang Wang, 2000, Trading volume: Definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257-300.

Markowitz, Harry, 1952, Portfolio selection, *Journal of Finance* 7, 77-91.

Morck, Randall, Bernard Yeung and Wayne Yu, 2000, The information content of stock prices: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58, 215-260.

Rubinstein, Mark, 2002, Markowitz's "Portfolio Selection": A fifty-year retrospective, *Journal of Finance* 57, 1041-1045.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Tobin, James, 1958, Liquidity preference as behavior towards risk, *Review of Economic Studies* 25, 65-86.

Treynor, Jack and Fischer Black, 1973, How to use security analysis to improve portfolio selection, *Journal of Business* 46, 66-86.

Wermers, Russ, 2000, Mutual Fund Performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55, 1655-1695.

Whitelaw, Robert, 1994, Time variations and covariations in the expectation and volatility of stock market returns, *Journal of Finance* 49, 515-541.

Whitelaw, Robert, 1997, "Time-varying Sharpe Ratios and market timing", NYU working paper.

Table 1 – Maximum Fraction of Volume Explained by Stock Picking by Country

Country	*1-R^2 1995-2004*	*1-R^2 1995-1999*	*Ranking 1995-1999*	*1-R^2 2000-2004*	*Ranking 2000-2004*
Panel A: Stock Picking in Emerging and Developed Countries					
Emerging	0.63	0.72		0.59	
Developed	0.45	0.51		0.41	
Panel B: World Rankings of Stock Picking					
United States	0.29	0.36	2	0.24	1
Sweden	0.31	0.33	1	0.30	3
Italy	0.34	0.41	4	0.31	4
Belgium	0.34	0.42	5	0.29	2
Russian Federation	0.35	0.39	3	0.33	6
Japan	0.41	0.43	6	0.40	11
Spain	0.41	0.65	21	0.34	8
France	0.42	0.55	15	0.32	5
Netherlands	0.42	0.46	8	0.40	10
Chile	0.44	0.58	16	0.36	9
Finland	0.44	0.50	13	0.41	12
Denmark	0.45	0.50	14	0.34	7
Portugal	0.47	0.44	7	0.47	16
Austria	0.47	0.49	11	0.44	14
United Kingdom	0.49	0.47	9	0.51	21
Pakistan	0.51	0.49	10	0.54	25
Norway	0.53	0.59	17	0.42	13
South Africa	0.55	0.65	20	0.50	19
Argentina	0.56	0.49	12	0.62	32
Ireland	0.57	NA	43	0.57	29
Greece	0.57	0.62	18	0.52	23
Bangladesh	0.59	0.69	24	0.49	18
New Zealand	0.60	0.64	19	0.50	20
Philippines	0.60	0.67	23	0.57	28
Turkey	0.61	0.70	25	0.57	27
Taiwan	0.61	0.66	22	0.59	30
Israel	0.62	0.73	29	0.54	24
Cyprus	0.63	0.72	27	0.56	26
Canada	0.65	0.72	28	0.47	15
Australia	0.66	0.74	34	0.62	31
Korea	0.67	0.76	35	0.48	17
Morocco	0.68	0.92	42	0.52	22
Germany	0.71	0.74	33	0.69	38
Poland	0.71	0.85	39	0.68	35
Malaysia	0.71	0.73	31	0.70	39
Thailand	0.71	0.73	32	0.69	37
Indonesia	0.72	0.73	30	0.71	40
Mexico	0.72	0.76	36	0.67	34
Hong Kong	0.73	0.72	26	0.80	43
Brazil	0.74	0.91	41	0.69	36
India	0.78	0.81	37	0.73	41
Peru	0.79	0.83	38	0.78	42
China	0.80	0.86	40	0.67	33

Notes: This table reports median time-series values of $1 - R^2$ from a regression of the natural log of share volume on the natural log of shares outstanding. The value of $1 - R^2$ serves as an upper bound on the volume explained by stock picking.

Table 2 – Regressions of Ln (Share Volume) on Ln (Shares Outstanding)

		Intercept	*Coefficient*	*1-R^2*
Panel A: S&P 500 Inclusion				
Non-S&P 500		-2.997	1.199	0.473
S&P 500		0.512	0.882	0.365
Group A	**Group B**	**A - B, P-Value**	**A - B, P-Value**	**A - B, P-Value**
Non-S&P 500	S&P 500	0.000	0.000	0.000
Panel B: Exchange Listing				
NYSE		-1.578	1.066	0.319
AMEX		-2.321	1.082	0.509
Nasdaq		-4.040	1.368	0.470
Group A	**Group B**	**A - B, P-Value**	**A - B, P-Value**	**A - B, P-Value**
NYSE	AMEX	0.000	0.000	0.000
NYSE	Nasdaq	0.000	0.000	0.000
AMEX	Nasdaq	0.000	0.000	0.000
Panel C: Size Quintiles				
Smallest		-2.164	1.020	0.670
Largest		0.120	0.921	0.543
Group A	**Group B**	**A - B, P-Value**	**A - B, P-Value**	**A - B, P-Value**
Small	Large	0.000	0.000	0.000
Panel D: Firm Age Quintiles				
Youngest		-1.710	1.076	0.574
Oldest		-2.745	1.175	0.222
Group A	**Group B**	**A - B, P-Value**	**A - B, P-Value**	**A - B, P-Value**
Young	Old	0.000	0.000	0.000
Panel E: Industry				
Utilities		-3.305	1.184	0.144
Manufacturing		-2.990	1.192	0.336
High-Tech		-2.836	1.243	0.340
Durables		-2.181	1.123	0.347
Medical		-1.067	1.026	0.367
Non-Durables		-2.501	1.128	0.370
Other		-3.213	1.249	0.403
Retail/Wholesale		-2.605	1.181	0.415
Energy		-1.682	1.055	0.480
Telecom		0.249	0.898	0.498
Panel F: Analyst Coverage Quintiles				
Lowest		-2.767	1.187	0.573
Highest		1.647	0.818	0.440
Group A	**Group B**	**A - B, P-Value**	**A - B, P-Value**	**A - B, P-Value**
Low	High	0.000	0.000	0.000

Notes: This table reports the results of a regression of ln (share volume) on ln (shares outstanding). The value of $1 - R^2$ from these regressions serves as the maximum fraction of volume that can be explained by stock picking. In each panel, the first set of rows gives the median time-series estimate of the regression intercept, slope, and $1–R^2$. The second set of rows give p-values from a Wilcoxon rank sum test of differences in medians.

Table 3 – Steady State Maximum Fraction of Stock Picking Under Alternative Models of the Market Price of Risk

Methodology	Mean λ	a_0	a_1	μ	σ	R^2	$1 – R^2$
Model 1	0.1145	1.005	$-2.6e^{-5}$	0.011	0.017	1	0
Model 2	0.1864	1.008	$-4.3e^{-5}$	0.019	0.027	1	0
Model 3	2.5515	1.087	$-7.5e^{-4}$	0.256	0.369	0.988	0.012
Model 4	0.0868	1.007	$-5.0e^{-6}$	-0.307	1.064	0.892	0.108

Notes: We estimate the market price of risk, λ, using four different models. For Model 1, we calculate the simple average market risk premium and divide it by the standard deviation of the market return. Models 2 and 3 are based on Harvey (1989), who suggests using GMM to estimate the market price of risk. The moment conditions for Model 2 are:

$$\mathrm{E}\begin{bmatrix} (r_{m,t+1} - \mathbf{X}_t\beta)\mathbf{X}_t' \\ \left(r_{m,t+1} - \lambda\left|r_{m,t+1} - \mathbf{X}_t\beta\right|\right)\mathbf{X}_t' \end{bmatrix} = 0 .$$

The moment conditions for Model 3 are:

$$\mathrm{E}\begin{bmatrix} (r_{m,t+1} - \mathbf{X}_t\beta)\mathbf{X}_t' \\ \left(r_{m,t+1} - \lambda(r_{m,t+1} - \mathbf{X}_t\beta)^2\right)\mathbf{X}_t' \end{bmatrix} = 0 .$$

For Model 4, we follow Whitelaw (1994, 1997), who suggests fitting expected returns and standard deviations as a function of macroeconomic variables using GMM. The moment conditions are:

$$\mathrm{E}\begin{bmatrix} (r_{m,t+1} - \mathbf{X}_t\beta)\mathbf{X}_t' \\ \left(\sqrt{\frac{\pi}{2}} \times \left|(r_{m,t+1} - \mathbf{X}_t\beta)\right| - \mathbf{X}_t\theta\right)\mathbf{X}_t' \end{bmatrix} = 0 .$$

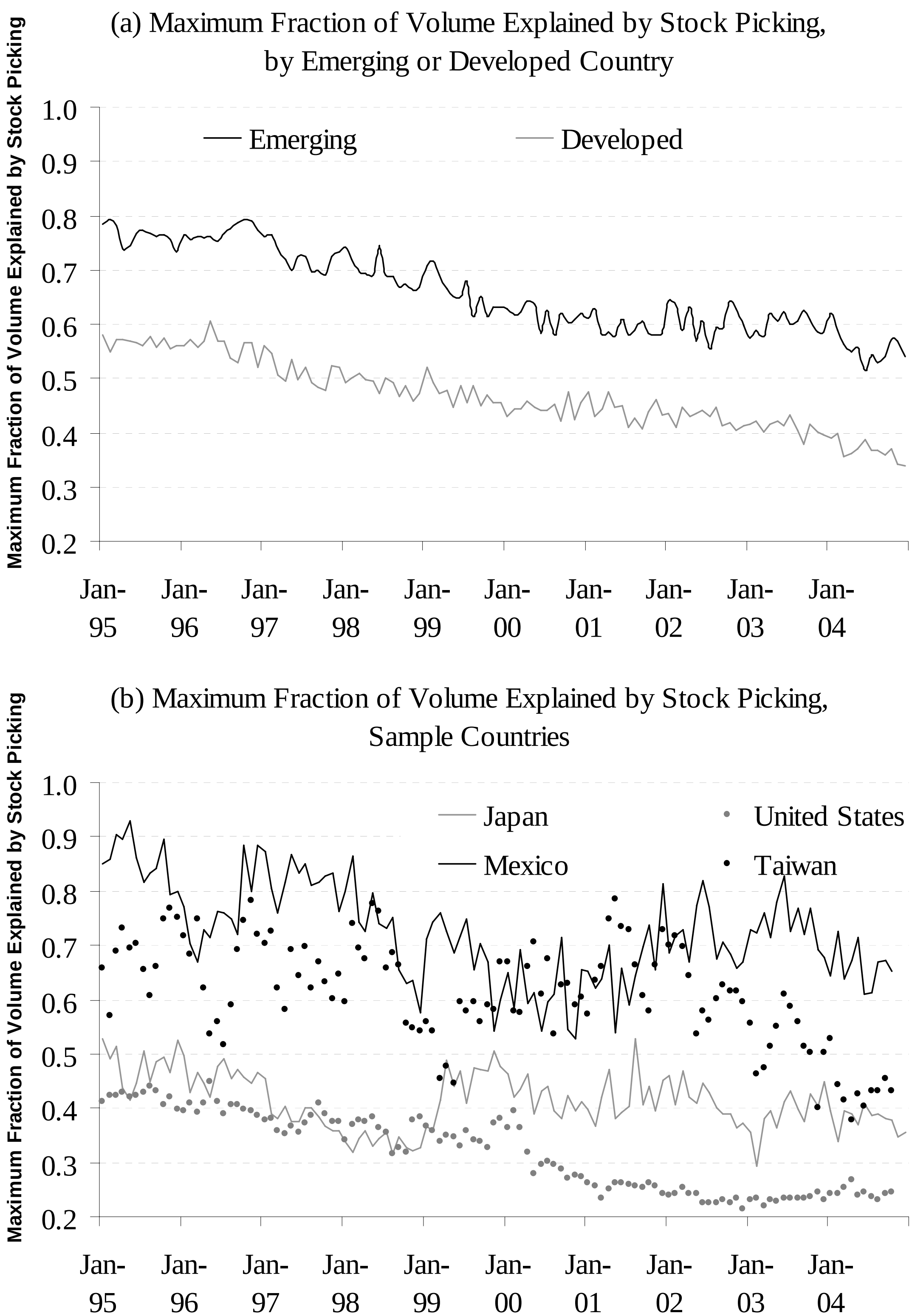


Figure 1: The Maximum Fraction of Volume Explained by Stock Picking, International Evidence. Each month, we fit a cross sectional regression, ln(Share Volume) = a + b*ln(Shares Outstanding) + e to all firms within a given country. Figure 1 (a) shows the monthly median value of $1 - R^2$ for countries categorized by MSCI as emerging versus those categorized as developed. Figure 1 (b) shows the United States against three other markets, two emerging and one developed.

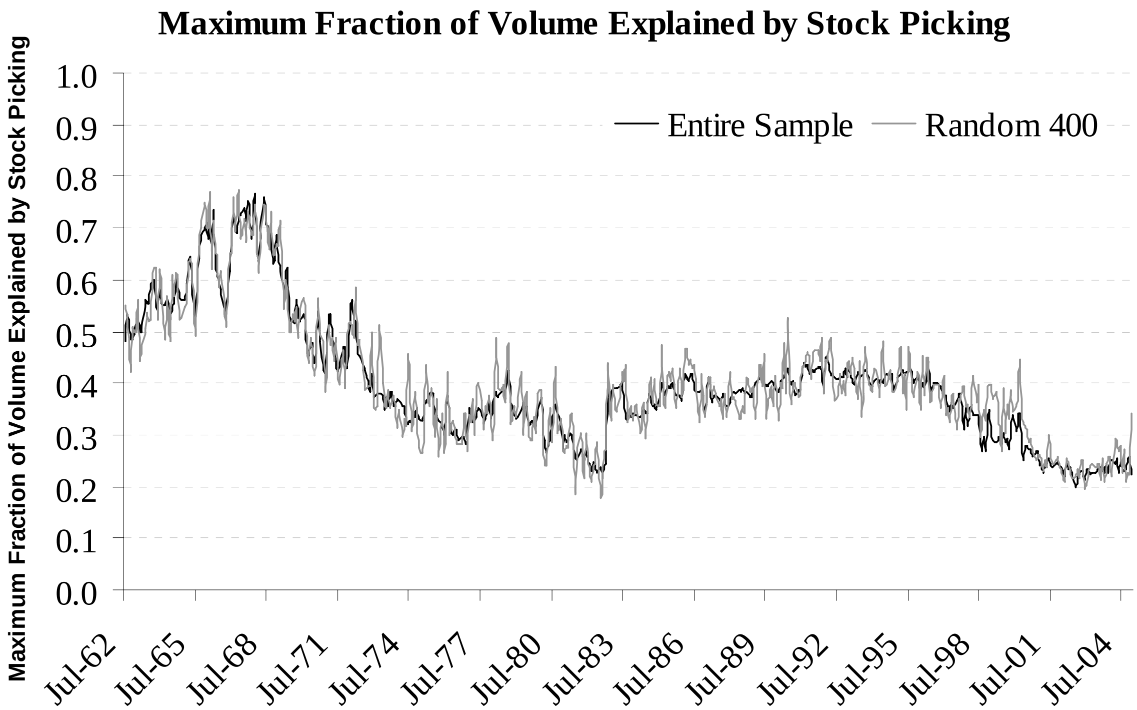
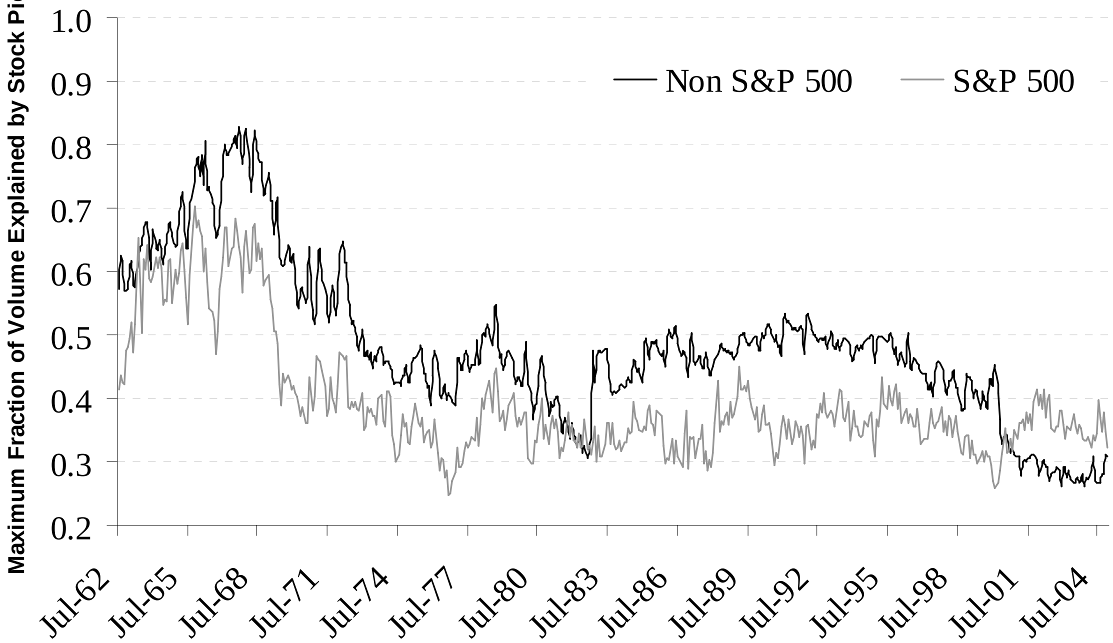
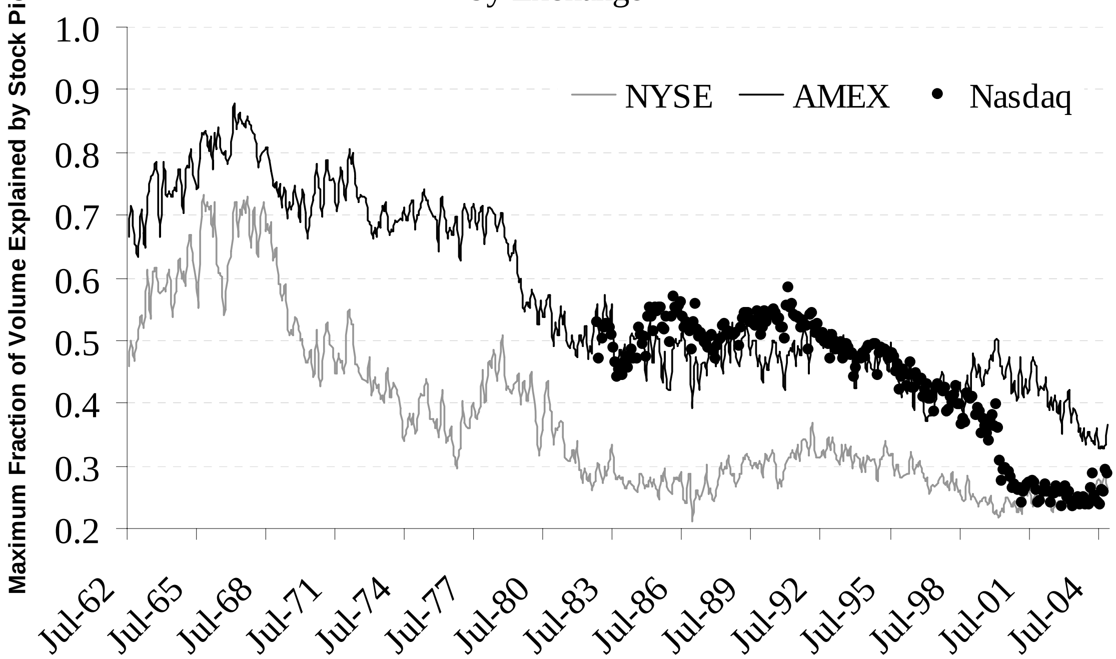


Figure 2: The U.S. History of the Maximum Fraction of Volume Explained by Stock Picking. Each month, we fit a cross sectional regression, ln(Share Volume) = a + b*ln(Shares Outstanding) + e to all firms in the U.S., and to 400 stocks randomly selected each month.

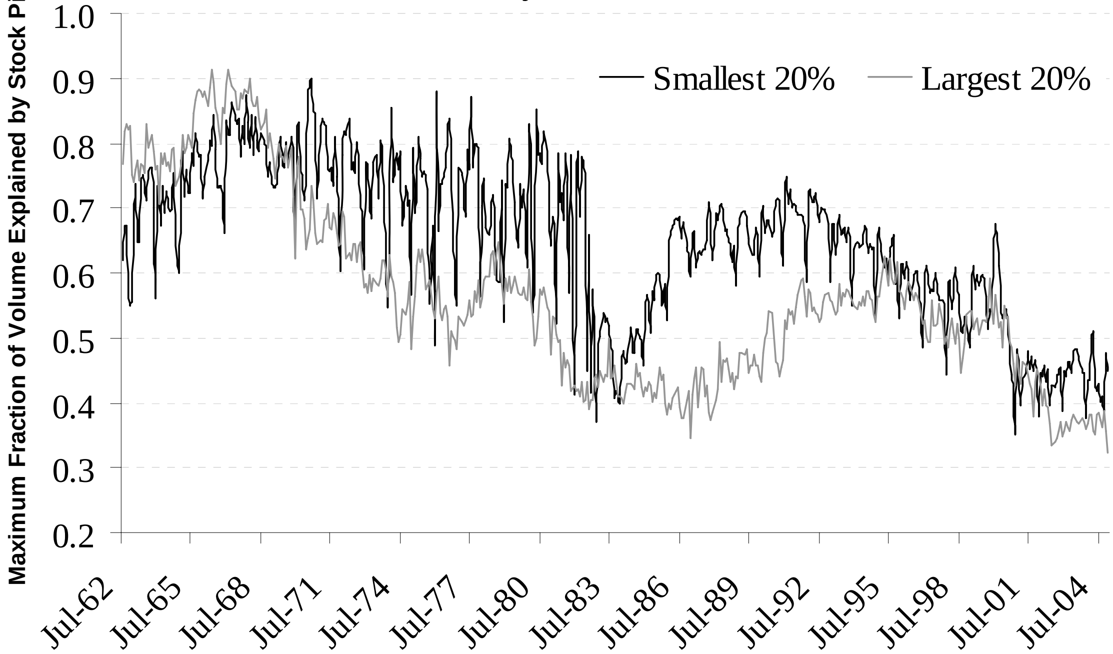
(a) Maximum Fraction of Volume Explained by Stock Picking, by S&P 500



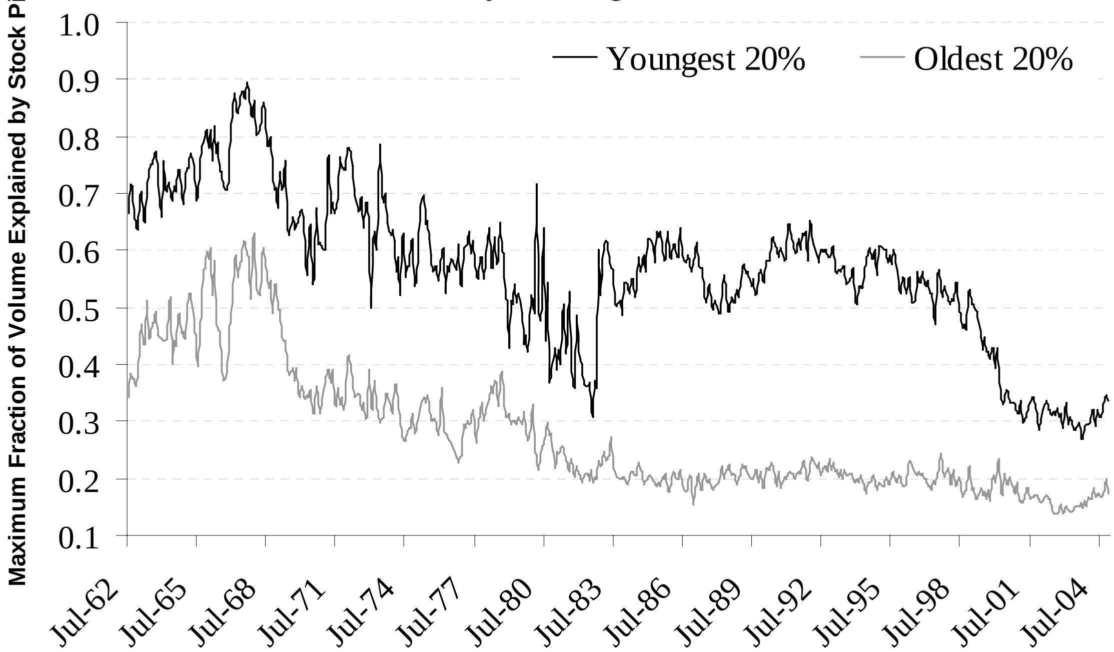
(b) Maximum Fraction of Volume Explained by Stock Picking, by Exchange

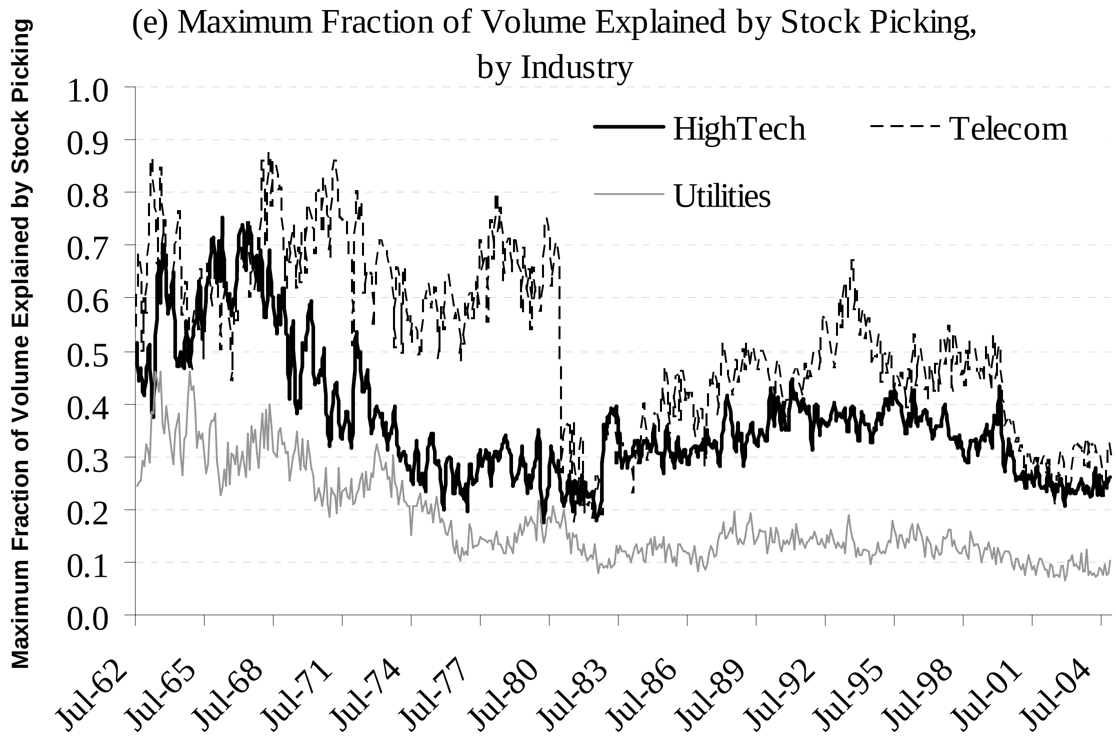


(c) Maximum Fraction of Volume Explained by Stock Picking, by Size



(d) Maximum Fraction of Volume Explained by Stock Picking, by Firm Age





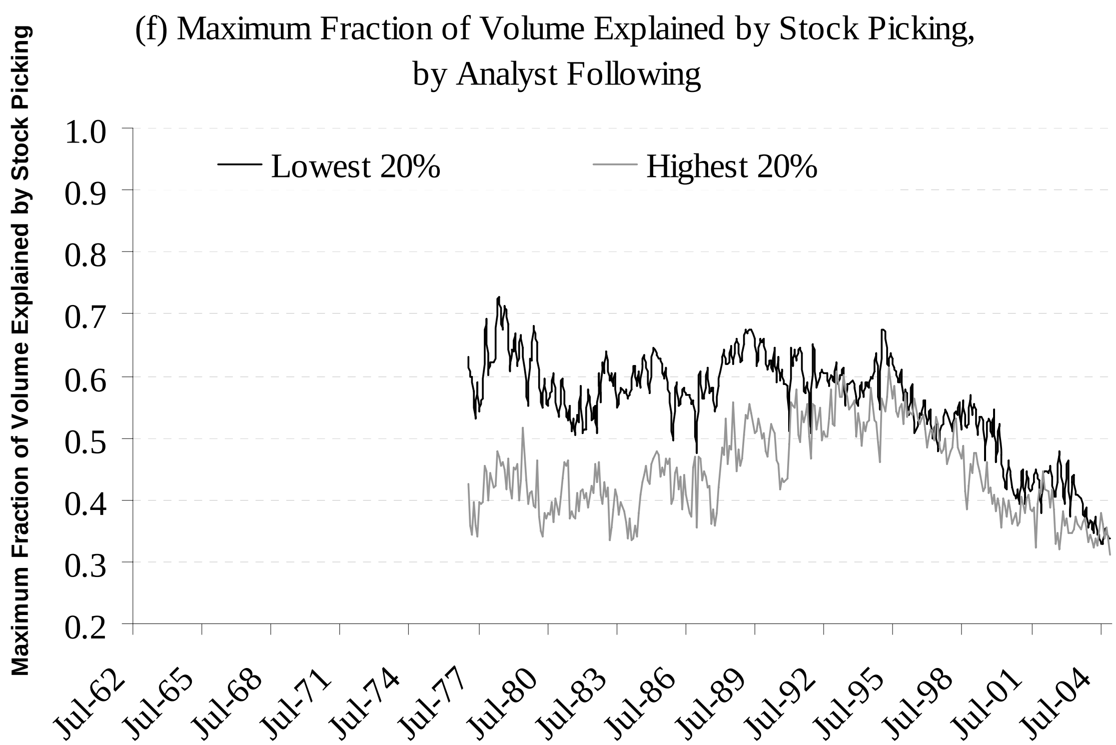


Figure 3: Firm Characteristics and the Maximum Fraction of Volume Explained by Stock Picking. Each month, we fit a cross sectional regression, ln(Share Volume) = a + b*ln(Shares Outstanding) + e to all firms within a specific category. In figures 2a through 2f, we plot 1 – R^2 from these regressions. The plots show (a) S&P 500 vs. not in S&P500, (b) by exchange, (c) by size quintile, (d) by firm age quintile, (e) by Fama-French 10 industry classification, and (f) by intensity of analyst following quintile.

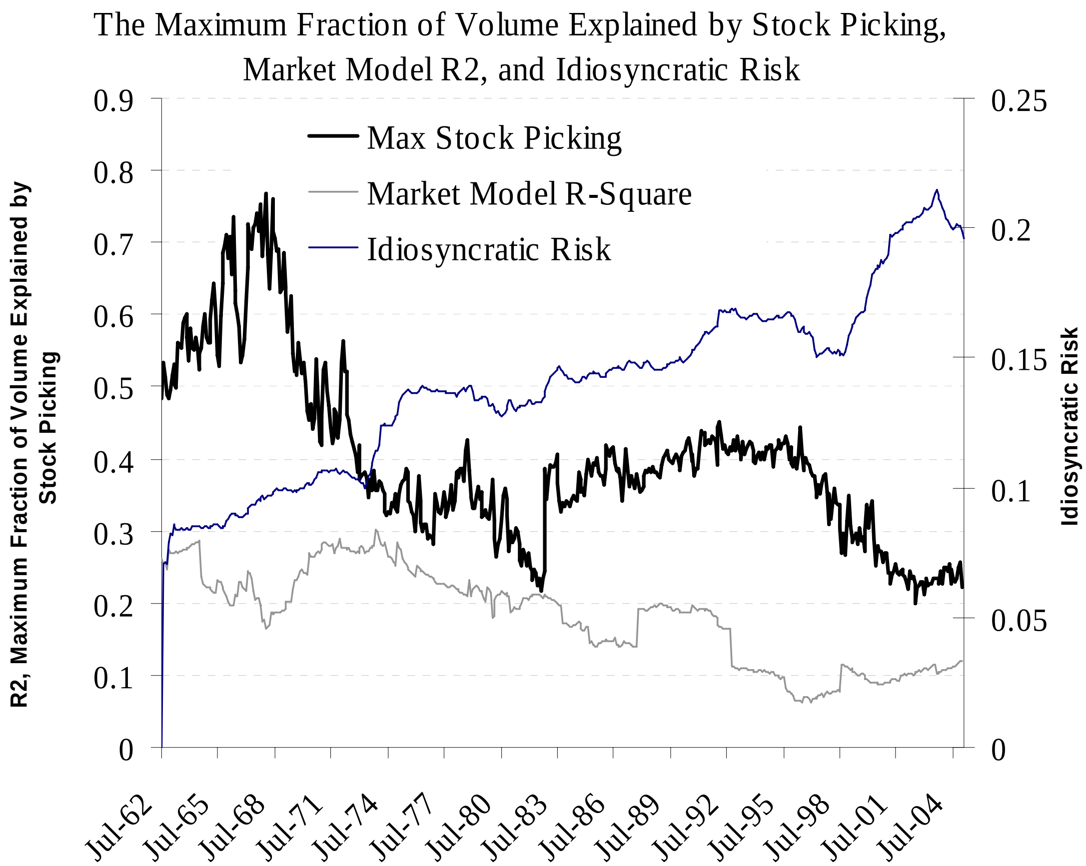


Figure 4: The Maximum Fraction of Volume Explained by Stock Picking, Market Model R^2, and Idiosyncratic Risk. For each stock, each month, we fit a trailing 5-year market model to monthly returns. Idiosyncratic risk is the average standard deviation of residuals from these regressions. The market model R^2 is the average R^2 of these regressions. In each month, we also run a cross sectional regression of ln (share volume) on ln (shares outstanding). The value of $1 – R^2$ from this regression measures the maximum fraction of volume explained by stock picking.

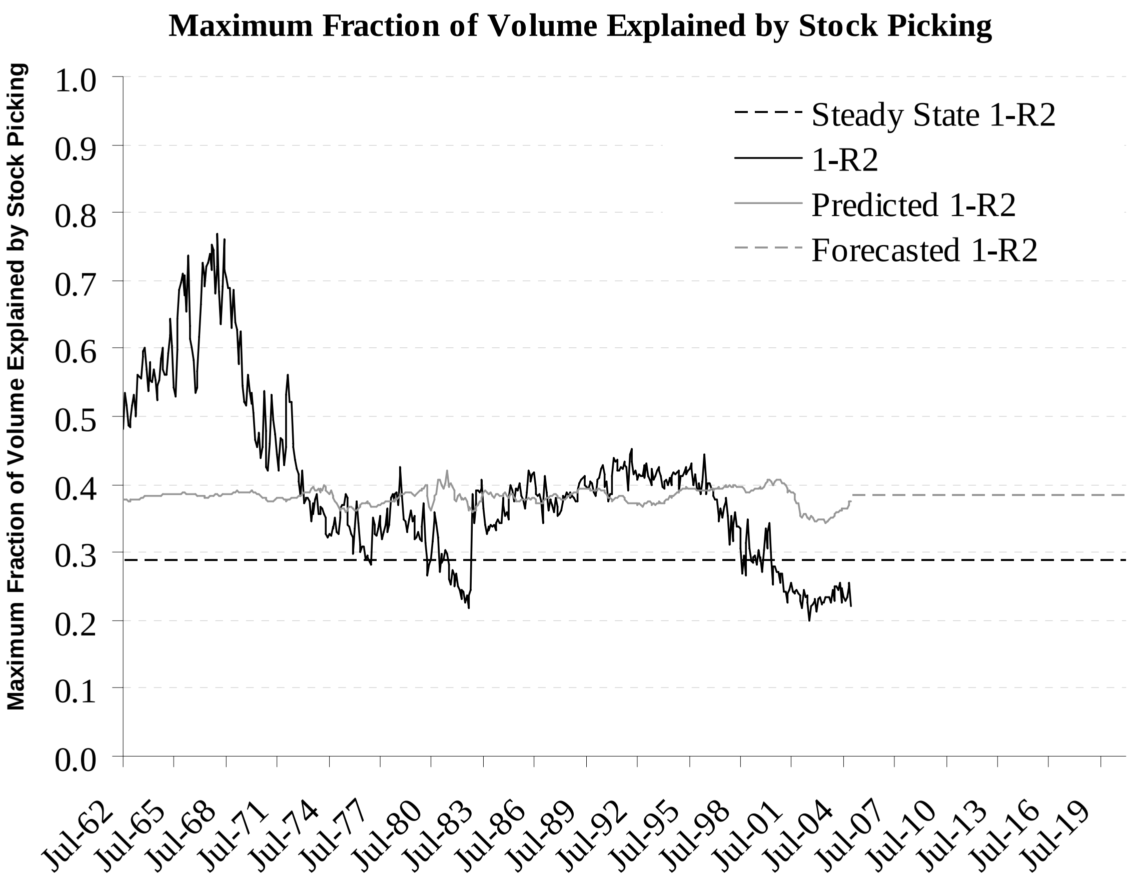


Figure 5: Maximum Fraction of Volume Explained by Stock Picking, Actual Values, Predicted Values, and Steady State Values. Each month, we fit a cross sectional regression ln(*Share Volume*$_i$) = **a** + **b***ln(*Shares Outstanding*$_i$) + u_i to all firms in the U.S. We plot the value of (1-R^2) from these regressions. We next estimate the profits to the marginal stock picker, which we refer as $\alpha_{net,indifferent}$, by multiplying the average idiosyncratic risk of U.S. stocks by the market Sharpe ratio. We find a long-run value for these profits by fitting the model $\alpha_{net,,indifferent,\,t}$ = **c** + **d** e $^{\mathbf{f}\,*\,time}$ + v_t using non-linear least absolute deviations. The estimate of **c** represents the long-run value of $\alpha_{net,indifferent}$. We assume that individual investor *alpha*s are normally distributed, and estimate the parameters **m** and **s** in the model R_t^2 = N($\alpha_{net,,indifferent,\,t}$, **m**, **s**) + w_t via non-linear least absolute deviations. N(x, $\mathbf{m_x}$, $\mathbf{s_x}$) is the cumulative normal distribution with mean $\mathbf{m_x}$ and standard deviation $\mathbf{s_x}$ evaluated at point x. The steady state maximum fraction of volume explained by stock picking is then N(c, m, s).