

Do Industries Lead Stock Markets?

Harrison Hong
Princeton University

Walter Torous
UCLA

Rossen Valkanov
UCLA

First Draft: July 31, 2002
This Draft: November 24, 2004

Abstract: We investigate whether the returns of industry portfolios are able to predict the movements of stock markets. In the U.S., we find that a significant number of industry returns, including retail, services, commercial real estate, metal and petroleum, can forecast the stock market by up to two months. Moreover, the propensity of an industry to predict the market is correlated with its propensity to forecast various indicators of economic activity such as industrial production growth. When we extend our analysis to the eight largest stock markets outside of the U.S., we find remarkably similar patterns. These findings suggest that stock markets react with a delay to information contained in industry returns about their fundamentals and that information diffuses only gradually across markets.

JEL: G12, G14, E44, G15

Keywords: Asset Pricing, Information and Market Efficiency, Financial Markets and Macroeconomy, International Financial Markets

We are grateful to Jeremy Stein and a referee for many insightful comments. We also thank John Campbell, Kent Daniel, Ken French, Owen Lamont, Toby Moskowitz, Sheridan Titman, Rob Engle, Matthew Richardson, David Hirshleifer, Matthew Slaughter, Will Goetzmann, Mark Grinblatt, our discussant at the 2004 AFA meetings in San Diego and participants at the Berkeley-MIT-Texas Real Estate Research Conference, Columbia, Dartmouth, Harvard Business School, NYU, Rice, and Yale for helpful comments.

Do Industries Lead Stock Markets?

Abstract: We investigate whether the returns of industry portfolios are able to predict the movements of stock markets. In the U.S., we find that a significant number of industry returns, including retail, services, commercial real estate, metal and petroleum, can forecast the stock market by up to two months. Moreover, the propensity of an industry to predict the market is correlated with its propensity to forecast various indicators of economic activity such as industrial production growth. When we extend our analysis to the eight largest stock markets outside of the U.S., we find remarkably similar patterns. These findings suggest that stock markets react with a delay to information contained in industry returns about their fundamentals and that information diffuses only gradually across markets.

JEL: G12, G14, E44, G15

Keywords: Asset Pricing, Information and Market Efficiency, Financial Markets and Macroeconomy, International Financial Markets

In this paper, we investigate whether the returns of industry portfolios are able to predict the movements of stock markets. We begin our analysis with the U.S. stock market. Over the period of 1946 to 2002, we find that fourteen out of thirty-four industries, including commercial real estate, petroleum, metal, retail, financial and services, can predict market movements by one month. A number of others such as petroleum, metal and financial can forecast the market even two months ahead. After adding a variety of well-known proxies for risk and liquidity in our regressions as well as lagged market returns, the predictability of the market by these fourteen industry portfolios remains statistically significant. We have also done numerical simulations to gauge just how many industries will (with statistical significance) forecast the market simply by chance, and found that in only 0.04% of the simulations fourteen or more industries forecast the market and that on average, only five (in contrast to the fourteen we find) are able to do so at the 10% level of significance.

In addition, we provide a few metrics regarding the statistical and economic significance of the documented predictability. First, we compare the ability of these industries to predict the market to well-known predictors such as inflation, the default spread or the dividend yield and find comparable forecasting power. Second, we show that a portfolio incorporating information in past industry returns can lead under certain circumstances to a higher Sharpe ratio than simply holding the market. And third, we then extend our analysis to each of the eight largest stock markets outside of the U.S., including Japan, Canada, Australia, U.K., Netherlands, Switzerland, France, and Germany. In contrast to the U.S., these time series are limited to the period of 1973 to 2002 and we are unable to obtain the same set of controls (e.g. market dividend yield, default spread). With these caveats in mind, we find that the U.S. results hold up remarkably well for the rest of the world---the only exception is Japan.

Our investigation is motivated by recent theories that explore the implications of limited information-processing capacity for asset prices. Many economists have recognized for some time that investors, rather than possessing unlimited-processing capacity, are better characterized as being only boundedly rational (see Shiller (2000), Sims (2001)). Even from casual observation, few traders can pay attention to all sources of information much less understand their impact on the prices of the assets that they

trade. Indeed, a large literature in psychology documents the extent to which even attention is a precious cognitive resource (see, e.g., Kahneman (1973), Nisbett and Ross (1980), Fiske and Taylor (1991)).

More specifically, our investigation builds on recent work by Merton (1987) and Hong and Stein (1999). Merton (1987) develops a static model of multiple stocks in which investors only have information about a limited number of stocks and only trade those that they have information about. As a result, stocks that are less recognized by investors have a smaller investor base (neglected stocks) and trade at a greater discount because of limited risk-sharing. Hong and Stein (1999) develop a dynamic model of a single asset in which information gradually diffuses across the investment public and investors are unable to perform the rational expectations trick of extracting information from prices. As a result, price under-reacts to the information and there is stock return predictability.¹

The hypothesis that guides our analysis of whether industries lead stock markets is that the gradual diffusion of information across asset markets leads to cross-asset return predictability. The basic idea is that certain investors, such as those that specialize in trading the broad market index, receive information originating from particular industries such as commercial real estate or commodities like metals only with a lag. As a result, the returns of industry portfolios that are informative about macroeconomic fundamentals will lead the aggregate market.

This hypothesis relies on two key assumptions. The first is that valuable information that originates in one asset market reaches investors in other markets only with a lag, i.e. news travels slowly across markets. The second assumption is that because of limited information-processing capacity, many (though not necessarily all) investors may not pay attention or be able to extract the information from the asset prices of markets that they do not specialize in. These two assumptions taken together lead to cross-asset return predictability.

Our hypothesis appears to be a plausible one for a few reasons. To begin with, even among equity money managers, there is specialization along industries such as

¹ For other models of limited market participation, see Brennan (1975) and Allen and Gale (1994). For related models of limited attention, see also Gabaix, Laibson, Moloche and Weinberg (2003), Peng and Xiong (2002) and Hirshleifer, Lim and Teoh (2002).

sector or market timing funds. Most plausibly, investors have their hands full trying to understand the markets that they specialize in. As a result, they are unable to devote the attention needed to process potentially valuable information from other markets in a timely manner. Moreover, as pointed out by Merton (1987) and the subsequent literature on segmented markets and limited market participation, few investors trade all assets. Put another way, limited participation is a pervasive feature of financial markets and may be another rationale for why investors in one market may be slow to adjust to information emanating from another.²

We recognize that this gradual-information-diffusion hypothesis may not be the only explanation for our findings. However, it does provide a key auxiliary prediction that we can take to the data---namely, the ability of an industry to predict the market ought to be strongly correlated with its propensity to forecast market fundamentals such as industrial production growth or other measures of economic activity. We test this prediction by first using individual industry returns to separately forecast industrial production growth and the growth rate of the Stock and Watson (1989) coincident index of economic activity. Many of the same sectors that lead the market can also forecast these two proxies of market fundamentals. Indeed, industry returns that are positively (negatively) cross-serially correlated with the market are also positively (negatively) cross-serially correlated with future economic activity. For instance, high returns for some industries like retail mean good news for future economic activity and the market, while high returns for other industries such as petroleum mean just the opposite.

Beyond the standard statistical inference techniques, we have also performed various numerical simulation exercises to rule out that such a relationship is due purely to chance. This finding strongly supports our hypothesis that the documented cross-predictability is due to the market reacting with a delay to information contained in industry returns about its fundamentals. It also distinguishes our gradual-information-diffusion hypothesis from other behavioral explanations of stock return predictability (see Section I.C). Importantly, we have also extended this analysis to the eight largest stock markets outside of the U.S. The key relationship, that the propensity of industries to

² Individual investors also participate in a limited number of markets as they hold very un-diversified portfolios (see, Blume and Friend (1978), King and Leape (1984)).

forecast the market is correlated with their propensity to forecast industrial production growth, is also present in seven of the eight countries. Not surprisingly, the only country in which this relationship does not emerge is Japan, where we find that industry returns did not forecast the aggregate market.

A few other papers have also found that some industry portfolios may be able to lead the stock market (see Eleswarapu and Tiwari (1996) and Pollet (2002)). In particular, Pollet (2002) finds that oil can predict stock returns and most interestingly that the Norwegian stock market (which is dominated by oil) leads the world stock market. His finding regarding the Norwegian market fits especially nicely with our gradual-information-diffusion hypothesis since the Norwegian market is likely to be off the radar screen of investors who trade the world market index.

Our paper is related to the literature on lead-lag relationships among stocks, epitomized by the finding of Lo and MacKinlay (1990) that large stocks lead small stocks. A number of other papers followed trying to rationalize this finding (see, e.g., Brennan, Jegadeesh and Swaminathan (1993), Badrinath, Kale and Noe (1995), Jegadeesh and Titman (1995)). These studies typically find that stocks that are in some sense more liquid (e.g. have more analysts following or have institutional ownership) tend to lead less liquid stocks. However, Boudoukh, Richardson and Whitelaw (1994) argue that many of these lead-lag relationships are due to the own-autocorrelation of portfolios and a high contemporaneous correlation among portfolios. In other words, a large stock portfolio does not significantly lead a small stock portfolio once the lagged returns of the small stock portfolio are included in a multiple regression.

Our findings differ from those in the above papers in a few respects. First, the lead-lag relationships we document occur at a much longer horizon of a couple of months as opposed to the those in the literature which play out over a couple of days. Second, we focus on predicting the aggregate stock market whereas the papers in this literature have focused on stocks of different characteristics leading or lagging each other. As a result, we are able to link this predictability to information about fundamentals such as industrial production growth. Third, the industry portfolios are not more liquid than the market portfolio, and yet they seem to lead the market, contrary to the conventional wisdom, which suggests that the more liquid security, the broad market index, ought to react to

news in a more timely fashion than less liquid securities such as, for example, the real estate index. Fourth, our findings are not due to the own-autocorrelation of portfolios since we control for lagged market returns in our predictive regressions and each industry portfolio comprises only a small fraction of the market portfolio.³

Our paper proceeds as follows. In Section I, we develop a simple model to make clear the assumptions behind our hypothesis and generate some testable predictions. We describe the data in Section II. We present our empirical findings for the United States in Section III. We extend our analysis to the eight largest stock markets outside of the U.S. in Section IV. We conclude in Section V.

I. Model

A. Basic Setup

Our model considers the pricing of two assets (stocks) in a three-date economy, $t=0, 1, 2$. We assume for simplicity that the risk-free rate is zero. The two assets, X and Y , have terminal values at $t=2$ given by D_X and D_Y , which are jointly normal with means of zero and variances of $\sigma_{X,D}^2$ and $\sigma_{Y,D}^2$ and covariance $\sigma_{XY,D}$.

Investors either participate in market X or market Y . This limited market participation assumption may be motivated by exogenous reasons such as taxes or regulations. Alternatively, we can motivate it by introducing a fixed cost of participation in each market so that investors will only want to participate in one.⁴

At $t=1$, investors in market X receive signal $S_X = D_X + \varepsilon_{X,S}$ about the terminal value of X , investors in market Y receive signal $S_Y = D_Y + \varepsilon_{Y,S}$ about the terminal value of Y , and these signals are known to all participants at $t=2$. This is our gradual-information-diffusion assumption. The noise in the signals $\varepsilon_{X,S}$ and $\varepsilon_{Y,S}$ are normally distributed with means of zero and variances of $\sigma_{X,S}^2$ and $\sigma_{Y,S}^2$ respectively. We assume that $\varepsilon_{X,S}$ and $\varepsilon_{Y,S}$

³ Our work is also related to the recent work on stock price momentum (see, e.g., Jegadeesh and Titman (1993), Hong, Lim and Stein (2000), Grinblatt and Moskowitz (1999)). While a number of these papers document momentum among industry portfolios, our focus is on whether and why industries lead the aggregate market index.

⁴ As long as the fixed costs to participating in X and Y are not too different, there will be some investors participating in each market. However, this assumption is not crucial to our results. Investors may participate in both markets as long as there are some investors who are slow to optimally combine signals from both markets in forming expectations.

are independent of each other and all other shocks in the economy. The supply of assets are assumed to be Q_X and Q_Y shares outstanding for assets X and Y , respectively.

Investors in asset X cannot process information pertaining to asset Y , and vice versa---this is our limited information-processing capacity assumption. This assumption is a simple way of capturing the idea that investors, due to limited cognitive capabilities, have a hard time processing information from asset markets that they do not participate in. This may be because information from other markets is less salient. Alternatively, investors may be too busy trying to figure out the market that they are in to process this information in a timely fashion.

We assume that investors have CARA preferences with a risk aversion coefficient of a . Given the price function $P_{k,t}$, the investor in asset market k ($k=X,Y$) solves the following optimization problem:

$$\begin{aligned} & \text{Max } E_{k,0} [- \exp (-aW_{k,2})] && k=X,Y \\ & \{ \theta_k \} && (1) \\ & \text{s.t. } W_{k,t} = W_{k,t-1} + \theta_{k,t-1} (P_{k,t} - P_{k,t-1}), \end{aligned}$$

where $W_{k,t}$ and $\theta_{k,t}$ are the wealth and share holding of a representative investor in asset market k at time t (we do not index different investors in the same asset market for simplicity) and $P_{k,2} = D_k$. The solution to this problem and the equilibrium prices are obtained using standard techniques.

The equilibrium price in market k is given by:

$$P_{k,t} = E_{k,t} [D_k] - b_{k,t} Q_k \quad k=X,Y \quad (2)$$

where $E_{k,t} [D_k]$ is the conditional expectation of the terminal payoff of asset k at time t , $b_{k,t} > 0$ is the standard risk discount at time t and Q_k is the supply of the asset.

B. Serial and Cross-Serial Correlations

Given the equilibrium prices described in equation (2), it is straightforward to calculate the serial and cross-serial correlations for assets X and Y . Let $R_{k,t} = P_{k,t} - P_{k,t-1}$

be the date t return for asset k . The two propositions that follow are self-explanatory and are given without proof.

Proposition 1: The own serial return correlations are zero, i.e. $\text{Corr}(R_{k,2}, R_{k,1}) = 0$ for $k=X,Y$. The cross-serial return correlations, $\text{Corr}(R_{Y,2}, R_{X,1})$ and $\text{Corr}(R_{X,2}, R_{Y,1})$, are non-zero and may be positive or negative depending on the sign of the covariance of asset payoffs, $\sigma_{XY,D}$.

Intuitively, investors in market k rationally condition on all information associated with market k . As a result, the price is efficient with respect to own asset information. Hence the own serial correlation is zero. However, investors in asset market Y ignore or cannot process the information from X , including past returns. As a result, the time-1 return in market X predicts the time-2 return in market Y . If investors in market Y (X) condition on the time-1 return in market X (Y), then the cross-serial correlations would be zero.

Moreover, the results in Proposition 1 would remain even if we enrich the model to include the following sets of traders. First, even if some fraction of the investors in each market paid attention to and can process information from the other market, there will still be cross-predictability, though it will be smaller in magnitude. Second, if there are limits of arbitrage (Shleifer and Vishny (1997)), then cross-predictability will remain in equilibrium even if there are arbitrageurs who try to profit from the cross-asset return predictability. We state this more formally in Proposition 2.

Proposition 2: Even if there are arbitrageurs who trade in both markets to exploit the cross-predictability, as long as there are limits to arbitrage, some cross-predictability will remain in equilibrium.

While our model is designed to generate positive cross-serial correlations even if own-serial correlations are zero, it is important to note that the model can be easily augmented to simultaneously generate own- and cross-serial correlations. If we additionally assume that some investors in asset k do not pay attention to or cannot process S_k , then along with cross-serial correlation, there is positive serial correlation, i.e.

$Corr(R_{k,2}, R_{k,1}) > 0$ for $k=X,Y$. Intuitively, if investors in the same market pay attention to (or wake up to) information at different points in time, then information gradually diffuses across investors in the same market (Hong and Stein (1999)), resulting in positive serial correlation as well as non-zero cross-serial correlation in asset returns.

C. Testable Predictions

In our empirical work, we test two specific predictions that are implied by our model. In the context of our model, think of the broad market index as asset Y and an industry portfolio that is informative of market fundamentals as asset X . Proposition 1 implies the following prediction.

Prediction 1: The broad market index can be predicted by the returns of industry portfolios, controlling for lagged market returns and well-known predictors such as inflation, the default spread, and the dividend yield.

Note that our model only implies that an industry will lead the market if it has information about market fundamentals. In other words, an industry with little information about economic activity will not forecast the market whether or not investors are paying attention to it. Indeed, it follows from this logic that an industry's ability to predict the market is correlated with the information that it has about market fundamentals. As a result, we have Prediction 2:

Prediction 2: The ability of an industry to forecast the market is related to its ability to forecast changes in market fundamentals such as industrial production growth or changes in other indicators of economic activity.

This prediction also distinguishes our gradual-information-diffusion hypothesis from other behavioral explanations of stock return predictability due to biased inferences on the part of a representative investor (see, e.g., Daniel, Hirshleifer and Subrahmanyam (1998) and Barberis, Shleifer and Vishny (1998)). For instance, Daniel, Hirshleifer and Subrahmanyam would attribute our cross-asset return predictability finding to continuing

overreaction to industry returns on the part of overconfident investors trading the aggregate market index. Their model, however, is silent on why this cross-predictability is strongly related to the informativeness of an industry about market fundamentals.

By the same token, Barberis, Shleifer and Vishny would argue that the documented cross-predictability is due to a conservatism bias, i.e. a representative investor updates on industry news about market fundamentals a bit slower than a Bayesian would. The degree to which prices under-react depends on how slowly the investor updates information. As such, their model would also not be able to generate this prediction unless there is an additional assumption that investors are slower to adjust to news from certain industries than others.

Moreover, anecdotal evidence is more consistent with our gradual-information-diffusion hypothesis than alternative explanations. For instance, there is plentiful evidence that limited market participation is pervasive and that such segmentation meant that equity investors paid little attention to many industries like the real estate market (see, e.g., Decker (1998)).

II. Data

Our data on industry portfolios from the U.S. stock market come from two sources. From Ken French's website, we obtain monthly returns to thirty-eight value-weighted industry portfolios for the years of 1946-2002 (see Fama and French (1997)). We have to drop five of these industries from our analysis because they have missing observations.⁵ Since commercial real estate is not a separate portfolio and is likely to provide a good setting to test our hypothesis, we augment this sample by constructing a real estate industry portfolio from an index of REIT returns obtained from the NAREIT website (www.nareit.com). We use the comprehensive, value-weighted REIT index of equity, mortgage, and hybrid REITs. The REIT data only goes back to January 1972. So, counting real estate, we consider thirty-four industry portfolios in all.

In addition to these indices, we also utilize the following variables. We use the returns of the CRSP value-weighted portfolio in excess of the risk-free rate (denoted by

⁵ The five industries that we exclude from our analysis are GARBAGE (sanitary services), STEAM (steam supply), WATER (irrigation systems), GOVT (public administration) and OTHER (everything else).

RM) as the proxy for the broad market index. Inflation (INF), measured as the growth rate of the Consumer Price Index, is obtained from the DRI database. Also obtained from the DRI database is the default spread (DSPR), defined as the difference between the yield of BAA-rated and AAA-rated bonds. The market dividend yield (MDY) is the one-year dividend from the CRSP value-weighted market portfolio divided by the current price. We also calculate a time series of monthly market volatility from daily returns to the CRSP value-weighted portfolio as in French, Schwert and Stambaugh (1987), which is denoted by MVOL.

We will also use the following two macroeconomic variables. From the DRI database, we obtain a time series of the level of industrial production, which is available at a monthly frequency. From Mark Watson's web page, we obtain a time series of the Stock and Watson (1989) coincident index of economic activity, which is also available at a monthly frequency. Their experimental coincident index is a weighted average of four broad monthly measures of U.S. economic activity. These measures are industrial production, real personal income, real manufacturing and trade sales, and total employee hours in non-agricultural establishments. The detailed descriptions of this indicator are also available on Mark Watson's web page. The data for industrial production is available back to 1946, but the data for the Stock and Watson index only goes back to March of 1959. We denote the monthly growth rates of industrial production and the Stock and Watson coincident index of economic activity by IPG and SWG, respectively. We will be interested in seeing how past industry returns forecast these two growth rates.

Table 1 provides summary statistics of these variables. The means and standard deviations are in monthly percentage points. Panel A of Table 1 lists the thirty-four industry portfolios (by their abbreviated names) along with their means and standard deviations of their returns. All returns are in excess of the one-month T-Bill rate. Panel B lists the statistics for the remaining variables. The acronyms of the industry portfolios are taken from Fama and French (1997). In most cases, the acronyms are self-explanatory. Precise definitions of these indices are available on Ken French's website (see the Appendix).

Notice that some of the industries are very related. For instance, OIL and PTRLM (petroleum) are treated as two different industries. The main difference between

them is that OIL covers oil and gas extraction, while PTRLM covers petroleum refining and petroleum products. Two other industries that are also related are MINE and STONE, with the difference being that STONE covers non-metallic minerals except fuels. MTLPR or metal products is treated differently from METAL, which covers primary metal industries. Importantly, MONEY includes financial, insurance and real estate. However, real estate comprises a miniscule part of MONEY. As such, we create a separate real estate portfolio (RLEST) by using the REIT index as a proxy.⁶

We obtain from the Datastream database monthly industry returns for each of the eight largest non-U.S. stock markets for the period of 1973 to 2002. The summary statistics for industry returns are summarized in Panel C, where the returns are calculated in each country's local currency. These returns are raw or unadjusted since data on risk-free rates for these countries are difficult to obtain. There are a total of thirty-five industries. Notice that a number of the stock markets outside of the U.S. may not have companies in particular industries. For instance, Australia does not have any companies that are categorized under AERSP or Aerospace. Observations that are not available are denoted appropriately.

The industry categorizations from the Datastream database are somewhat different from the Fama-French industry portfolios. (See the Appendix for a description of the composition of each of the industry portfolios.) While Datastream provides descriptions of their industry portfolios, it is not possible to convert the Fama-French industry portfolios into Datastream portfolios and vice versa. We could have featured the Datastream portfolios for the US stock market, but we decided on the Fama-French portfolios because they are more readily accessible, are available over a much longer period of time series (Datastream data for U.S. also only goes back to 1973) and are probably less prone to measurement error. The results for the U.S. stock market hold regardless of the database that we use (we do not report these results below for brevity). More generally, we think that our results are more convincing if we can get similar results using various types of industry classifications.

⁶ Interestingly, we have replicated our findings using an alternative real estate industry portfolio from Ken French's website. This portfolio consists of small stocks such as realty companies and real estate brokers but excludes REITs, spanning 1970-2002. It is correlated with our real estate index but may not be as informative as RLEST since REITs are required to invest most of their resources in properties.

We also obtain from the Datastream database monthly market returns and monthly industrial production growth figures for these countries from 1973 to 2002. In Panel D of Table 1, we report summary statistics for these variables by country.

III. Evidence from the U.S. Stock Market

A. Predictive Regressions Involving Industry and Market Returns

We begin by exploring the ability of industry returns to lead the market using U.S. data. To see whether industries can forecast the market (Prediction 1), we estimate the following specification separately for each of the thirty-four portfolios:

$$RM_t = \alpha_i + \lambda_i R_{i,t-1} + \mathbf{A}_i \mathbf{Z}_{t-1} + e_{i,t} \quad (3)$$

where RM_t is the excess return of the market in month t , $R_{i,t-1}$ is the excess return of industry portfolio i lagged one month and \mathbf{Z}_{t-1} is a vector of additional market predictors. For each of these thirty-four time-series regressions, there are a total of 684 monthly observations.

We include a number of well-known market predictors in \mathbf{Z}_{t-1} to address alternative explanations for why industry returns might forecast the market. Among them is the lagged excess market (RM_{t-1}), inflation (Fama and Schwert (1977)), the default spread (Fama and French (1989)), and the market dividend yield (Campbell and Shiller (1988), Fama and French (1988)). These variables are typically thought to proxy for time varying risk. To the extent that our results hold even with these predictors in the regressions, it suggests that our findings are not due to time varying risk. Additionally, we worry that industry returns may be forecasting market volatility, so we also include lagged market volatility in our set of control variables.

In an earlier draft, Hong, Torous and Valkanov (2002), we used a shorter sample beginning in 1973 and included as additional controls the term spread and changes in the Fed Funds rate. The term spread has been documented to predict the market (Fama and French (1989)). And since some industries such as financials may proxy for changes in the liquidity of the aggregate market or be especially sensitive to monetary policy variables, we also include lagged changes in the Fed Funds rate for good measure. The

data for these two variables do not go back to 1946 and as a result, we are unable to include them in this draft. However, these controls had little effect on our findings when using the shorter sample and as such are not likely to change our results here.⁷

The coefficients of interest are the thirty-four λ_i 's, which measure the ability of each of the industry portfolios to lead the market. Since many of these industries are likely to contain valuable information about market payoffs, we expect a significant number of these coefficients to be non-zero to the extent that our gradual-information-diffusion hypothesis holds.

Alternatively, rather than investigating whether different industries lead the market separately, we can augment specification (3) by simultaneously including all thirty-four industry returns. The cost of doing this is that the standard errors on our estimates will be larger since we only have a limited number of observations and so we cannot estimate the effect of each industry on future market returns very precisely. The benefit of doing this is that since industry returns are contemporaneously correlated, we worry about issues related to omitted variables. In other words, some of our results may be biased by not simultaneously including all other industries.

It turns out that our results are not significantly affected by whether we run the forecasting regressions separately or by pooling all the lagged industry returns. So, for the sake of precision, we present the results using specification (3). We discuss the results when we pool all the industries into one regression below.

We first present the results for the case of the metal industry portfolio. This allows us to thoroughly describe all the regression specifications used in our analysis without reporting the results of all specifications for every industry. Much of the discussion for this case also applies to the other industries.

In Table 2, we report the results of various regressions that establish the predictive ability of the metal industry portfolio. In column (1), we run a forecasting regression of market return on a constant, lagged values of the metal portfolio and RM.⁸ The coefficient on lagged metal is -0.096 and is statistically significant. Surprisingly, this

⁷ Similar results hold when we augment the specification to include the returns of the Fama and French (1993) small stock minus large stock portfolio and high book-to-market minus low book-to-market portfolio.

coefficient is still statistically significant even after we control for other predictors such as INF, DSPR and MDY in column (2). In column (3), we augment the specification in column (2) by adding in lagged market volatility, but the coefficient of interest remains statistically significant. Across columns (1) to (3), the usual market predictors such as INF, DSPR and MDY have predictive power in this sample.⁹

We now see how many of these industries lead the market in Table 3. Our regression specification includes a constant, lagged one-month industry return, lagged one-month market return, inflation, default spread, market dividend yield and market volatility (all lagged by one month). Note that we are running this regression separately for each industry. Rather than report the coefficient of each of the independent variables for every one of the thirty-four regressions, we report just the coefficient of the particular lagged one-month industry return along with the R^2 of the regression.

The dependent variable in column (1) of Table 3 is next month's market return. The industries that have significant coefficients in this column are denoted by asterisks. There are twelve industries including commercial real estate (RLEST), mines (MINES), apparel (APPRL), print (PRINT), petroleum (PTRLM), leather (LETHR), metal (METAL), transportation (TRANS), utilities (UTILS), retail (RTAIL), money or financial (MONEY), and services (SRVC) that have t-statistics of the corresponding lagged industry return that are greater than 1.96 in absolute value (or significant at the 5% level). Two additional industries, non-metallic minerals (STONE) and television (TV), have t-statistics of about 1.7. So at the 10% level of significance (or t-statistics greater than 1.65 in absolute value), there are a total of fourteen industries that can significantly predict the market return.

The signs on the predictability coefficients for these fourteen industries also seem to make economic sense. For instance, the lagged returns of petroleum and metal industry portfolios are negatively related to next period's market return as one might suspect since these are commodity (input) prices whose shocks have historically led the economy into a downturn. In contrast, retail and apparel are sectors that, when they are

⁸ Indeed, we have also experimented with adding in multiple lags of the market: past month, two months previous and three months previous. Our results are unchanged.

booming, are generally thought to be signs of a thriving economy. The fact that the signs of these predictive relations are consistent with conventional wisdom on the relation of these industries to the macro-economy reassures us that these predictive regressions are indeed capturing the slow diffusion of sector information into the broad market index as opposed to being the result of chance (see also Section III.B below).

Finally, note that our findings are not simply an artifact of industry returns being serially correlated. First, most of the industries represent a small fraction of the market. So it is not likely that they forecast the market simply because their returns are serially correlated and part of the market portfolio. Second, the time series of most of the industry portfolios that can lead the market such as commercial real estate, apparel, petroleum, metal, and utilities are not (statistically significantly) serially correlated at a monthly frequency.¹⁰ They are also not an artifact of the market portfolio being auto-correlated since we control for lagged market returns in our predictive regressions.

An interesting empirical question is whether industries lead the stock market by more than one month? Our model only predicts that there is such cross-predictability but is silent on by how many months industries ought to lead the stock market. However, we know from the literature on stock market predictability that being able to predict next month's return is already quite an achievement as it is notoriously difficult to predict the market at long horizons. Indeed, Valkanov (2003) and Torous, Valkanov, and Yan (2004) argue that previous findings on long horizon predictability are an artifact of not properly adjusting standard errors for the near random walk behavior of various predictors. As such, we would not expect our industry portfolios to predict the market at very long horizons.

In columns (2) and (3) of Table 3, we investigate whether these industry portfolios are able to lead the market by more than one month. In column (2), the dependent variable is the market return in month $t+1$. Now the coefficient on the lagged industry return is statistically significant at the 10% level for nine industries and at the

⁹ Results for regressors such lagged RM, INF, DSPR and MDY are the same when we use other industry returns to forecast the market. Again, this is why we only present detailed results for only one of the industries.

¹⁰ However, a number of other industry portfolios such as construction, smoke, textiles, retail and money exhibit positive serial correlation (see also Grinblatt and Moskowitz (1999)). We omit these results for brevity.

5% level for no industries. RLEST, APPRL, LETHR, TRANS, UTILS, and RTAIL are no longer statistically significant. In column (3), the dependent variable is the market return in month t+2. Of the fourteen industries that are significant in column (1), only MINES and METAL are still statistically significant. Importantly, notice that at the 10% level of significance, only four industries in all have a statistically significant coefficient on lagged industry return and none have such a coefficient at the 5% level. In other words, the evidence is consistent with information taking about two months to be completely incorporated from industries into the broad market index.

Finally, in Table 3, we present a joint test of the null hypothesis that all the λ_i 's are equal to zero ($\lambda_1 = \lambda_2 = \dots = \lambda_{34} = 0$) for each of the three dependent variables (month t, month t+1 and month t+2 market returns). To do this, we stack the regressions given by (3) into a GMM system and calculate the Wald test of the joint significance of the coefficients. The Wald tests (and all other tests reported in the paper) are computed using a Newey-West heteroskedasticity and serial correlation robust variance covariance matrix with 3 monthly lags. Values of the Wald tests and corresponding p-values are reported in the final row of Table 3. For month-t returns, we can reject the null hypothesis that all the λ_i 's are equal to zero at the 1% level of significance. For month t+1 returns, we can only reject the null at the 10% level of significance. And for month t+2 returns, we cannot reject the null.

We have also looked at cross-predictability at horizons of up to 6 months and find that there is virtually no predictability at longer horizons. This is a comforting finding since it suggests that our predictive regressions are informative and not subject to some bias that mechanically yields significant results.

In Table 4, we take a more careful look at the ability of these industries to forecast the next month market return. The goal is to compare the forecasting power of industry returns to those of well known predictors such as inflation, dividend yield, term spread and default spread as a gauge of the economic relevance of our findings. Towards this end, we calculate the effect of a two-standard deviation shock to an industry's lagged monthly return on the next month's market return (using the coefficient estimates from column (1) of Table 3). We report right below a lower and upper bound on the estimate obtained by using the standard errors of the coefficient estimates in column (1) of Table

3. In addition, we report the absolute value of this magnitude as a fraction of market volatility. For instance, the coefficient in front of METAL from column (1) of Table 3 is -0.085 and plus or minus two-standard errors gives us -0.159 as the lower bound and -0.011 as the upper bound. A two standard deviation shock in the monthly return of the metal industry is ($2 \times 6.1\%$). So it leads to a change in next month's market return of 1.04% ($-0.085 \times 2 \times 6.1\%$). The lower and upper bounds on this estimate are obtained by using -0.159 and -0.011 instead of -0.085. The estimate of 1.04% is roughly 25% of market volatility.

The industries are listed in descending order, by the most economically significant industry first. As one might expect, the fourteen industries that have a statistically significant ability to predict the market are also among the leaders in terms of economic significance. MONEY is very significant, with a two standard deviation shock in its returns resulting in a movement of market returns that is forty percent of market volatility. The next most economically significant is real estate (RLEST). Print (PRINT), apparel (APPRL), and services (SRVC) round out the top five.

We perform the same exercise for the other market predictors (INF, DSPR, MDY, and MVOL) and report the results in Panel B. The coefficients on these other market predictors change very little across the separate industry-by-industry regressions. To save on space, we only discuss the coefficient estimates on these other market predictors from the METAL regression, which are already reported in Table 2, specification (3). INF is by far economically the strongest of the usual market predictors. A two-standard deviation shock in inflation (INF) leads to a 1.13% ($-0.624 \times 2 \times 0.905$) movement in the market, which is roughly 27% of market volatility. Our industry portfolios such as METAL do a comparable job of forecasting the market. These results must be interpreted with caution as the economic predictors have been shown to forecast expected returns over longer horizons. They are slower moving and may not capture short-term fluctuations as do the industry returns. Nonetheless, comparing Panel A to Panel B indicates that the industry returns are economically significant leading indicators of the market returns.

Finally, we investigate in Panel C if our industry portfolios help us to better allocate between the market and the risk-free rate. That is, we form a portfolio comprising of the market and the risk-free rate based on information from lagged industry returns and the other economic predictors and we see whether this portfolio does better than just passively owning the market or, alternatively, timing the market using only information from the economic predictors. This exercise provides additional evidence about the usefulness of the industry returns as predictors, though the comparison is not perfect since there are transactions costs associated with moving in and out of the market into the risk-free and we do not take into account.

Using data from the period of January 1946 to June 1974, which is exactly the half-way point of our sample, we estimate two predictive regressions. The first is of RM_t on Z_{t-1} (INF, DSPR, MDY, and MVOL). The second is of RM_t on Z_{t-1} and simultaneously all the industry returns. Then, using these estimated coefficients, we form one period ahead forecasts of RM_t without re-estimating the coefficients for the second-half of the sample. The mean squared error (MSE) of these forecasts are computed. We report the relative MSE (Relative MSE) for the case in which we only use the other market predictors (Z_{t-1}) and the case in which we use both Z_{t-1} and all the industry returns as predictors. We observe an improvement in MSE of about 1.5%.

Then we consider the following portfolio strategies, which we implement using data from the second-half of the sample. Again, the forecasts are generated using estimates obtained from the first-half of the sample. For each month, if the forecasted RM is higher than the risk-free rate (in the last period), we invest everything in the market for a return RM. Otherwise, we invest everything in the risk-free asset. Thus, we obtain returns from two investment strategies, one using only Z_{t-1} and the other using Z_{t-1} as well as all the lagged industry returns. The mean, standard deviation, Sharpe Ratio, and certainty equivalent returns are reported in Table 4 for these two strategies along with the strategy of passively holding the market. The certainly equivalent returns are computed as follows. We compute the utility $U(1+R_p)$, where R_p is the return from the strategy and $U(x)$ is either $\log(x)$ or $x^{1-RRA}/(1-RRA)$, where RRA is the coefficient of relative risk aversion. Then, we compute the average utility over the period and invert it to obtain the certainly equivalent return.

Notice that the means of these two strategies are lower than simply just holding the market. The standard deviations of these two strategies, however, are also lower, leading to higher Sharpe Ratios than just simply holding the market. The Sharpe Ratio of the market during the latter half of the sample is 0.46, while the strategy incorporating only Z_{t-1} information yields a Sharpe Ratio of 0.49 and the one using both Z_{t-1} and all the industry returns yields a Sharpe Ratio of 0.52.

Moreover, it also appears that industry information also leads to higher certainty equivalent returns. The interpretation of the certainty equivalent return is that an investor with a RRA parameter of 1 will be willing to pay up to 1.5% on an annual basis to use the lagged industry returns in forming a portfolio. Naturally, as the RRA parameter increases, the investor becomes more risk averse and the advantage of the strategy for RRA=5 is only about 0.3%.

The point of this comparison is to show that, compared to other predictors, industry returns do have some additional information about the market. However, we caution against jumping to the conclusion that this is a more profitable strategy. Using information on lagged industry returns requires more frequent trading. For instance, using the strategy with Z_{t-1} only, we were rebalancing between the risk-free and risky assets 49 times and were in the risk-free asset 148 months out of a total of 342 months. In contrast, we were re-allocating between the two assets 107 times and were in the risk-free 195 months out of a total of 342 months using the strategy that also incorporated industry information. This suggests that the latter strategy may involve substantially more transaction costs.

B. Industry Returns and Market Fundamentals

In this section, we attempt to test Prediction 2, an industry's ability to predict the market ought to be correlated with its ability to forecast indicators of economic activity (i.e. market fundamentals). We begin by specifying the regression for forecasting market fundamentals:

$$X_t = \eta_i + \gamma_i R_{i,t-1} + C_i Z X_{t-1} + v_{i,t} \quad (4)$$

where X_t is the month t realization of the indicator of economic activity, $R_{i,t-1}$ is the previous month's return of industry i and the ZX_{t-1} is the same as Z_{t-1} in equation (3) except that we also include three monthly lags of the indicators of economic activity. The coefficients of interest are the γ_i 's, which measure the ability of the various industry returns to predict the economic activity indicator of interest.

To the extent that Prediction 2 holds, we expect that the relationship between the λ_i 's and γ_i 's to be positive. In other words, the industries that can strongly forecast the market ought to also forecast market fundamentals. For instance, industries such as commercial real estate that have a positive λ_i ought to also have a positive γ_i . And industries such as metals or petroleum that have a negative λ_i ought to also have a negative γ_i .

To implement the regression specified in equation (4), we need to identify proxies for economic activity. We use two well-known measures that have been previously studied in the literature. The first is industrial production growth, IPG. We use this measure because it is one of the few measures of economic activity that is available at a monthly frequency. Industrial production growth is contemporaneously correlated with the aggregate market. Over the period of 1946-2002, IPG and RM have a contemporaneous correlation of 0.147.

The second measure of economic activity that we use is SWG, the monthly growth rate of the Stock and Watson (1989) coincident index of economic activity. SWG is also contemporaneously correlated with the aggregate stock market. Over the period of March 1959 to December 2002, SWG and RM have a (monthly) contemporaneous correlation of 0.030.

In Panel A of Table 5, we determine which of the thirty-four industries can forecast industrial production growth. The regression specification is equation (4) but we only report the coefficient on lagged industry return. Twelve of the thirty-four industries are statistically significant at the 10% level and nine are significant at the 5% level. Our finding that industries contain valuable information about future economic fundamentals is consistent with Lamont (2001), who finds that portfolios formed from industry returns can track various economic variables like industrial production growth, inflation and consumption growth.

Finally, at the bottom of Panel A, we present a joint Wald test of the null hypothesis that all the γ_i 's are equal to zero ($\gamma_1 = \gamma_2 = \dots = \gamma_{34} = 0$). To do this (as we did at the end of Table 3) we stack the regressions given by (4) into a GMM system and calculate the Wald statistic of the joint significance of the coefficients. Values of the Wald tests and corresponding p-values are reported in the final row of Panel A. We can reject the null hypothesis that all the λ_i 's are equal to zero at the 5% level of significance.

More importantly, it appears that the industries that forecast the market (from Table 3) also forecast industrial production growth. Recall from Table 3 that MINES, PTRLM and METAL negatively forecast the market: higher returns in these industries in month t lead to lower returns in the market the next month. Interestingly, these three industries also forecast industrial production growth with a negative coefficient: higher returns in these industries in month t lead to lower industrial production growth the next month. This is exactly what we would expect with the slow incorporation of information into the broad market index. Moreover, RETAIL, MONEY, and RLEST, which are positively cross-serially correlated with the market, also forecast industrial production growth with a positive coefficient.

To formally see that an industry's ability to forecast the market is indeed correlated with its ability to forecast industrial production growth, we plot λ_i on the y-axis and γ_i on the x-axis in Figure 1(a). As a benchmark, recall that in an efficient market, we would expect to see all the λ_i 's be around zero, i.e. the slope of the scatter plot ought to be zero. In contrast, we see a distinctly positive relationship between the λ_i 's and the γ_i 's. In Figure 1 (a), we also plot the fitted values from a linear regression of λ_i 's on γ_i 's. The slope coefficient is 1.40 with a t-statistic of 2.60. In other words, there is a strong positive correlation between the ability of industries to forecast the market and their ability to forecast industrial production growth. (We have also conducted other statistical inference exercises to check that our results are not due purely to chance. See Section III.C below.)

We obtain similar results when we use SWG, the percentage change in the Stock and Watson coincident index of economic activity. The results are presented in Panel B of Table 5. Using this measure, we find that eight industries now are able to forecast market fundamentals at the 10% level of significance and five industries at the 5% level

of significance. The null hypothesis that all the γ_i 's are equal to zero ($\gamma_1 = \gamma_2 = \dots = \gamma_{34} = 0$) is rejected using the Wald test. In Figure 1(b), we plot the fitted values from a linear regression of the λ_i 's on γ_i 's. The slope coefficient is 3.37 with a t-statistic of 5.50. In other words, there is a strong positive correlation between the ability of industries to forecast the market and their ability to forecast economic activity when we use an alternative measure of the change in market fundamentals.

One potential worry is that the results in Figure 1 may be due to coincidence/measurement error. Suppose that, for whatever accidental reason, in our sample period, high returns for say the retail industry just happened to be followed by increases in industrial production. Since increases in industrial production will be naturally contemporaneously correlated with positive returns in the stock market, it is likely that retail will then also lead the market. In other words, in a given sample, the measurement error in our estimate of the coefficient of industry x returns on future industrial production is going to be correlated with the measurement error in our estimate of the coefficient of industry x returns on future market returns. This can generate a pattern like that in Figure 1 even if the underlying true coefficients are not related.

One conservative way to deal with this issue is to split the sample period in half, and for every industry, estimate one of the coefficients in the first half, and the other in the second half. (This is very conservative because there might be genuine time variation in the parameters.) We can then do similar exercises to Figure 1. We use the 1946-1973 sub-sample to estimate the λ_i 's for RM and the 1974-2002 sub-sample to estimate the γ_i 's for IPG. (We get similar results if we use the 1946-1973 sub-sample to estimate γ_i 's and the 1974-2002 sub-sample to estimate the λ_i 's.) We then re-do the analysis in Figure 1(a). The coefficient is 2.3 with a t-stat of 1.97. We repeat the same exercise for SWG. We use the March 1959 to December 1981 sub-sample to estimate λ_i 's for RM and the 1982-2002 sub-sample to estimate the γ_i 's for SWG. We then re-do the analysis in Figure 1(b). The coefficient is 2.9 with a t-stat of 2.3. These findings are similar to those obtained in Figure 1. For brevity, we omit the figures. Hence, we can conclude that our findings in Figure 1 are not due to coincidence/measurement error.

C. Additional Analysis

C.1 Numerical Simulations

In this section, we report the numerical simulations that we have conducted to check that the results of regression (3) are not due to chance. To do so, we randomize the returns of the time-series of the market portfolio, RM. In other words, rather than preserving the time series ordering of the market returns, we shuffle them and pretend as if the shuffled series is the true time series of the market. We then run regression (3) for all thirty-four industries and save the coefficients using this randomized market series. We keep repeating this procedure for 10,000 times. Each time, we also keep track of the number of coefficients that are significant at the 5 and 10 percent levels of significance. The number of simulations in which we have 14 significant λ_i 's is 3 out of 10,000. The number of simulations in which we have 15 significant λ_i 's is 1. There are no simulations with 16 or more significant λ_i 's. Hence, in 4 out of 10,000 simulations, or 0.04% we find 14 or more significant results. In addition, we find that on average (across the 10,000 trials) about 5.4 industries have significant coefficients at the 10% level of significance and 2.6 industries at the 5% level of significance. These small numbers indicate that our findings are not due to chance.

We have repeated a similar analysis for regression (4) as well as the regressions presented in Figure 1. In all instances, when we conduct these randomized experiments, we conclude that our findings are not spurious. For brevity, we do not report these additional calculations here.

C.2 Sub-Periods

As a robustness check, we arbitrarily divide our sample into two equal sub-samples (1946-1973 and 1974-2002) to see whether the ability of these industries to lead the market differs across these sub-periods. We implement this sample split by augmenting the specification in equation (3) by adding in a dummy variable, $1(\text{Year} > 1973)$, that equals one if the observation is after 1973 and zero otherwise and this dummy interacted with the lagged industry return. The coefficient of interest is the interaction term involving lagged industry return and the indicator $1(\text{Year} > 1973)$. We expect that the coefficient in front of the interaction term to be zero for these industries. Among the

thirty-four industries, not one industry has a coefficient in front of the interaction term that is statistically significant at the 10% level. This evidence strongly suggests that information diffuses only gradually even in today's financial markets. This finding is robust to different splits of the sample. We have re-done this analysis for cut-off points ranging from 1973 to 1978 and it does not make a big difference. We omit these results for brevity and they are available upon request.

C.3 Forecasting the Market and Indicators of Economic Activity Using All Industries Simultaneously

Up to this point, we have looked at how industries separately forecast the market and the various indicators of economic activity. The reason we do this is to improve the precision of our estimates since we have a limited number of monthly observations and 34 industries along with a host of other control variables. However, a cost of decreasing the standard errors of our estimates is the omitted variable bias due to industry returns being contemporaneously correlated. Certain industries may look like they forecast the market or economic activity but do not once we consider all industries jointly. So, a natural follow-up question is what happens when we simultaneously include all industries in our forecasting specifications?

In Panel A of Table 6, we answer this question by forecasting the market and the indicators of economic activity using all industries simultaneously along with the control variables specified in Tables 3 and 5, respectively. (Note that since RLEST only goes back to the early seventies, this industry portfolio is excluded from the current analysis.) The values of the F-tests (with thirty-three restrictions) under the various null hypotheses are reported. The first row reports the F-tests and p-values for three separate null hypotheses.¹¹ The first is whether the industry returns jointly do not forecast the market (RM). The p-value is less than 0.01, which means that we can strongly reject this null at the 5% level. The other two hypotheses are whether the industry returns can jointly forecast the various indicators of economic activity (IPG and SWG). Again, in the cases of IPG and SWG, we can reject the null at the 5% level.

¹¹ The degrees of freedom of the F-tests are determined as in Hayashi (2000, p. 43).

The second row reports the F-tests and p-values for the null hypotheses that the control variables jointly do not forecast the market and the indicators of economic activity. In each case, we can reject these null hypotheses at the 5% level of significance. The third row reports the p-values for the null hypotheses that all industry returns and all control variables do not jointly forecast the market or indicators of economy activity. The results are similar to those in the first and second rows in that we can reject that these variables do not jointly have forecasting power.

We do not report the coefficients from the regressions in Table 6 for brevity. However, we want to point out that the coefficients from these forecasting regressions are similar to those coefficients obtained in Tables 3 and 5. Moreover, we have also re-done the analysis in Figure 1 using the coefficients from these alternative regression specifications and obtain similar results. So even though the coefficients are estimated much more imprecisely when we include all industries simultaneously, the economic messages that we obtained from Table 3, Table 5, and Figure 1 are confirmed.

C.4 Alternative Methodology for Testing Prediction 2

Finally, we attempt to test Prediction 2 in a more parameterized manner. Rather than estimating the regression specifications in (3) and (4) separately and then analyzing the relationship between the estimated coefficients, λ_i 's and γ_i 's, as in Figure 1, we attempt to impose the restriction that these coefficients are proportional to each other, $\gamma_i = \kappa \lambda_i$, and estimate the coefficient κ and λ_i 's from the following specification:

$$\begin{aligned}
 RM_t &= \alpha + \lambda_1 R_{1,t-1} + \dots + \lambda_N R_{N,t-1} + \mathbf{AZ}_{t-1} + e_t \\
 X_t &= \eta + \kappa(\lambda_1 R_{1,t-1} + \dots + \lambda_N R_{N,t-1}) + \mathbf{CZX}_{t-1} + v_t
 \end{aligned}
 \tag{5}$$

where N is the number of industry portfolios in the estimation, \mathbf{Z}_{t-1} and \mathbf{ZX}_{t-1} , are as in equations (3) and (4), and e_t and v_t are random errors. We can estimate the system in (5) using GMM. Note that this estimation excludes the real estate portfolio since we can only obtain data back to 1973. The null hypothesis of interest is that $\kappa=0$, whereas if Prediction 2 hold true, we expect to find that $\kappa>0$. The results are reported in Panel B of

Table 6. We find that the coefficient $\kappa=2.142$ with a t-statistic of 2.644 for IPG. The comparable numbers of SWG are, respectively, 3.620 and 2.488. These numbers are similar to those reported in Figure 1. So the results of this estimation strongly confirm that there is a statistically significant positive relationship between the predictive content of an industry for the market and the information it has for economic activity.

C.5 Alternative Measures of Economic Activity

We have also experimented with how past industry returns forecast the deviations of these two macroeconomic variables from a potentially stochastic trend. Band-pass filters are a popular method used by economists to de-trend these time series. Such filters de-trend a time series, say of industrial production, by subtracting industrial production in month t from a weighted average of the levels of industrial production surrounding month t (say from month $t+k$ to month $t-k$), where the weights are optimally chosen. Note that a simple first-difference is a special case of this filter, which highly weighs high-frequency fluctuations in industrial production. Band-pass filters allow us to put less weight on high-frequency fluctuations in industrial production and other macroeconomic time series and indeed retain the lower frequency fluctuations of our choice.

We use the popular band-pass filter developed by Baxter and King (1999), the codes for which are available on Marianne Baxter's website. In practice, different industries may forecast fluctuations in economic activity at different frequencies. Rather than filtering industrial production and the Stock and Watson index differently for different industries, for the sake of parsimony, we set the parameters of this filter to capture fluctuations in industrial production and the Stock and Watson index at frequencies between two and twelve months.¹² We call the deviations from trend for industrial production and the Stock and Watson index calculated using the Baxter-King band-pass filter, IPD and SWD, respectively. We have re-done all of the same calculations involving for IPG and SWG for IPD and SWD and obtained similar results--these results are reported in an earlier draft (see Hong, Torous and Valkanov (2002)).

¹² More specifically, the parameter k for the weighted average of industrial production around month t is 19 months.

IV. Evidence from the Rest of the World

In this section, we extend the empirical analysis documented in Section III (for the U.S. stock market) for each of the eight largest stock markets outside of the U.S. In Table 7, we present the results of estimating the regression specifications given in (3) and (4) for the eight countries outside of the U.S. For regression specification (3), we now include as a control the lagged monthly market return. We are unable to control for other market predictors used in the U.S. stock market sample due to lack of data. For regression specification (4), we include as controls one lag of the monthly market return and three monthly lags of IPG.

In Panel A of Table 7, rather than listing all the coefficients for all countries, we report the number of industries that can significantly predict the market and IPG at the 10% level of significance for each country. Across the eight countries, the UK and Japan have the smallest proportion of significant industries (8 out of 34). Australia and the Netherlands have the most (18 out of 31 and 16 out of 29, respectively). The numbers for IPG are comparable. These findings are very similar to those obtained for the U.S. stock market and they far exceed the threshold figures that one would expect just from chance (see discussion in Section III.C). Hence, we can safely conclude that the basic findings that industry returns can predict both aggregate market returns and IPG extend beyond the U.S.

Moreover, the remarkable consistency of these findings across the eight countries outside of the U.S. should calm any lingering doubts that our results are due to chance that our numerical simulations did not alleviate. They should allay any concerns that our U.S. findings are purely a by-product of data mining. Indeed, the consistency of these patterns strongly indicates that we are capturing genuine economic phenomena.

In addition to estimating regression specifications (3) and (4) for each country, we have also attempted to forecast the market and IPG using all industries simultaneously just as we did for the U.S. (see Section III.C). The results are quite similar to those presented in Table 6 for the U.S. Across all eight countries, the p-values from F-tests under the null hypothesis that all the coefficients of all industry returns are jointly equal to zero are less than 0.01. The F-tests are reported in Panel B of Table 3.

While it is comforting to know that industries do indeed have information about future aggregate stock market movements and IPG, the crux of our paper lies in whether the propensity of an industry to predict the stock market is related to its propensity to predict economic activity. So we take the estimates from regression specifications (3) and (4) for each country and estimate the relationship between these two sets of coefficients for each country, just as we did for the U.S. in Figure 1. We present these findings in Figure 2. Notice that for every country except Japan, there is a positive and statistically significant relationship between the estimated λ_i 's and γ_i 's.

Finally, in Panel C of Table 7 we present the results of the restricted regressions given in (5) for each of the countries. Not surprisingly, we find that the results are consistent with those in Figure 2. For every country except for Japan, the coefficient of proportionality κ is positive and statistically significant, thereby strongly indicating that propensity of industries to predict returns is correlated with their informativeness about fundamentals.

V. Conclusion

We investigate whether the returns of industry portfolios are able to predict the movements of stock markets. We find that out of thirty-four industry portfolios in the U.S., fourteen including commercial real estate, petroleum, metal, transportation, utilities, retail and financial can predict the market by up to two months. Importantly, the ability of an industry to lead the market is strongly correlated with its propensity to forecast indicators of economic activity such as industrial production growth. When we extend our analysis to the eight largest stock markets outside of the U.S., we find remarkably similar patterns. These findings indicate that markets incorporate information contained in industry returns about their fundamentals only with a lag because information diffuses gradually across asset markets.

The logic of our hypothesis suggests that the gradual diffusion of information across asset markets ought to be pervasive. As a result, we would expect to find cross-asset return predictability in many contexts beyond industry portfolios and the broad market index. The key to finding such cross-predictability is to first identify sets of assets whose payoffs are likely correlated. As such, other contexts for interesting

empirical work include looking at whether returns of stocks from one industry predict those in another or at stocks and the options listed on them.

Indeed, a number of papers following ours have taken up this task and found confirming results. For instance, Menzly and Ozbas (2004) find that industry returns do lead and lag each other according to their place in the supply chain. And Pan and Poteshman (2004) find that information may diffuse slowly from option markets to stock markets as option volume seems to be able to predict stock price movements. But much more work remains to be done on this topic.

References

Allen, Franklin and Douglas Gale, 1994, "Limited market participation and volatility of asset prices," *American Economic Review* 84, 933-955.

Badrinath, S.G., Jayant R. Kale and Thomas H. Noe, 1995, "Of sheperds, sheep and the cross-autocorrelations in equity returns," *Review of Financial Studies* 8, 401-430.

Barberis, Nicholas, Andrei Shleifer and Robert Vishny, 1998, "A model of investor sentiment," *Journal of Financial Economics* 49, 307-343.

Baxter, Marianne and Robert G. King, 1999, "Measuring business cycles: Approximate band-pass filters for economic time series," *Review of Economics and Statistics* 81, 575-593.

Blume, Marshall E., and Irwin friend, 1978, *The Changing Role of the Individual Investors: A Twentieth Century Fund Report*, (New York: Wiley).

Boudoukh, Jacob, Matthew P. Richardson, and Robert F. Whitelaw, 1994, "A tale of three schools: Insights on autocorrelations of short-horizon stock returns," *Review of Financial Studies* 7, 539-573.

Brennan, Michael J., 1975, "The optimal number of securities in a risky asset portfolio when there are fixed costs of transacting: Theory and some empirical results," *Journal of Financial Quantitative Analysis* 10, 483-96.

Brennan, Michael J., Narasimhan Jegadeesh and Bhaskaran Swaminathan, 1993, "Investment analysis and the adjustment of stock prices to common information," *Review of Financial Studies* 6, 799-824.

Campbell, John Y. and Robert J. Shiller, 1988, "The dividend-price ratio and expectations of future dividends and discount factors," *Review of Financial Studies* 1, 195-228.

Daniel, Kent D., David Hirshleifer and Avanidhar Subrahmanyam, 1998, "Investor psychology and security market under- and over-reactions," *Journal of Finance* 53.

Decker, Mark, 1998, "The modern real estate investment trust industry: An overview," in *Real Estate Investment Trusts: Structure, Analysis, and Strategy*, (Eds, Richard T. Garrigan and John F.C. Parsons).

Eleswarapu, Venkat and Ashish Tiwari, 1996, "Business cycles and stock market returns: Evidence using industry-based portfolios," *Journal of Financial Research* 19, 121-134.

Fama, Eugene and Kenneth French, 1988, "Dividend yields and expected stock returns," *Journal of Financial Economics* 22, 3-25.

Fama, Eugene and Kenneth French, 1989, "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics* 25, 23-49.

Fama, Eugene and Kenneth French, 1993, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics* 33, 3-56.

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

Fama, Eugene and G. William Schwert, 1977, "Asset returns and inflation," *Journal of Financial Economics* 5, 115-146.

French, Kenneth R., G. William Schwert and Robert F. Stambaugh, 1987, "Expected stock returns and volatility," *Journal of Financial Economics* 19, 3-29.

Fiske, Susan and Shelley Taylor, 1991, *Social Cognition* 2nd ed., McGraw-Hill, New York.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg, 2003, "The allocation of attention: Theory of evidence," MIT Economics Working Paper.

Grinblatt, Mark and Tobias Moskowitz, 1999, "Do industries explain momentum?" *Journal of Finance* 54, 1249-1290.

Hayashi, Fumio, *Econometrics*, Princeton University Press: Princeton, NJ, USA.

Hirshleifer, David, Seongyeon Lim and Siew Hong Teoh, 2002, "Disclosure to a credulous audience: The role of limited attention," Ohio State University Working Paper.

Hong, Harrison, Lim, Terence, and Stein, Jeremy C., 2000, "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies," *Journal of Finance* 55, 265-295.

Hong, Harrison and Jeremy C. Stein, 1999, "A unified theory of underreaction, momentum trading and overreaction in asset markets," *Journal of Finance* 54, 2143-2184.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2002, "Do industries lead the stock market? Gradual diffusion of information and cross-asset return predictability," FEN Working Paper (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=326422).

Jegadeesh, Narasimhan and Titman, Sheridan, 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency," *Journal of Finance* 48, 93-130.

Jegadeesh, Narasimhan and Sheridan Titman, 1995, "Overreaction, delayed reaction and contrarian profits," *Review of Financial Studies* 8, 973-993.

Kahneman, Daniel, 1973, *Attention and Effort* (Prentice-Hall, Englewood Cliffs, New Jersey).

King, Mervyn and Jonathan I. Leape, 1984, "Wealth and portfolio composition: Theory and Evidence," NBER Working Paper No. 1468.

Lamont, Owen, 2001, "Economic tracking portfolios," *Journal of Econometrics* 105, 161-184.

Lo, Andrew and Craig MacKinlay, 1990, "When are contrarian profits due to stock market overreaction?" *Review of Financial Studies* 3, 175-206.

Menzly, Lior and Oguzhan Ozbas, 2004, Cross-industry momentum, USC Working Paper.

Merton, Robert C., 1987, "A simple model of capital market equilibrium with incomplete information," *Journal of Finance* 42, 483-510.

Newey, Whitney and Kenneth West, 1987, "A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica* 55, 703-708.

Nisbett, Richard and Lee Ross, 1980, *Human Inference: Strategies and Shortcomings of Social Judgment* (Prentice-Hall, Englewood Cliffs, NJ).

Pan, Jun and Allen Poteshman, 2004, The Information in Option Volume for Future Stock Prices, MIT Working Paper.

Peng, Lin and Xiong, Wei, 2002, "Capacity constrained learning and asset price comovements," Princeton University Working Paper.

Pollet, Joshua, 2002, "Predicting asset returns with expected oil price changes," Harvard University Working Paper.

Shiller, Robert J., 2000, *Irrational Exuberance*, (Broadway Books: New York).

Shleifer, Andrei and Robert Vishny, 1997, "The limits of arbitrage," *Journal of Finance* 52, 35-55.

Sims, Christopher, 2001, "Rational Inattention," Princeton University Working Paper.

Stock, James and Mark W. Watson, 1989, “New indexes of coincident and leading economic indicators,” *NBER Macroeconomics Annual 1989*, 351-394.

Torous, Walter, Ross Valkanov, and Shu Yan, 2004, “On predicting stock returns with nearly integrated explanatory variables,” *Journal of Business* 77(4), forthcoming.

Valkanov, Ross, 2003, “Long-horizon regressions: Theoretical results and applications,” *Journal of Financial Economics* 68, 201-232.

Table 1: Summary Statistics

The table presents summary statistics of the variables of interest. Panels A and B contain data about the U.S. stock market, while Panels C and D contain data for the eight largest equity markets outside of the U.S. In Panel A, the variables are the returns of the thirty-four industry portfolios in excess of the risk-free rate. In Panel B, RM is the CRSP value-weighted market portfolio return in excess of the risk-free rate. INF is the CPI inflation rate. DSPR is the default spread between BAA-rated and AAA-rated bonds. MDY is the dividend yield of the market portfolio. MVOL is the market volatility computed from daily return data. IPG is industrial production growth and SWG is the growth rate of the Stock and Watson (1989) coincident index of economic activity. All variables in panels A and B are from January 1946 to December 2002 with the exception of RLEST, which is from January 1972 to December 2002, and SWG, which is from March 1959 to December 2002. In Panel C, the variables are the raw (unadjusted) returns of the thirty-five industry portfolios (if available) for the United Kingdom, Australia, Canada, France, Germany, Japan, the Netherlands, and Switzerland. In Panel D, RM is the market portfolio return and IPG is industrial production growth for these same countries. All variables in Panels C and D are from January 1973 to December 2002. The data are at monthly frequency and in monthly percentage points.

Panel A: Industry Portfolio Returns (U.S.)					
Industry	Mean	Std. Dev.	Industry	Mean	Std. Dev.
RLEST	0.435	4.373	LETHR	0.691	6.232
AGRIC	0.433	7.239	GLASS	0.515	5.866
MINES	0.416	6.259	METAL	0.434	6.109
OIL	0.691	6.671	MTLPR	0.587	4.876
STONE	0.843	7.654	MACHN	0.616	5.821
CNSTR	0.523	6.934	ELCTR	0.650	6.236
FOOD	0.669	4.247	CARS	0.628	5.406
SMOKE	0.872	5.664	INSTR	0.704	5.367
TXTLS	0.488	5.934	MANUF	0.595	6.344
APPRL	0.402	6.573	TRANS	0.504	5.728
WOOD	0.661	7.268	PHONE	0.402	4.696
CHAIR	0.496	5.522	TV	0.897	6.714
PAPER	0.699	5.341	UTILS	0.471	3.868
PRINT	0.647	5.342	WHLSL	0.619	5.542
CHEMS	0.660	4.593	RTAIL	0.634	5.120
PTRLM	0.765	4.952	MONEY	0.673	4.835
RUBBR	0.610	6.019	SRVC	0.657	6.517

Panel B: Other Variables (U.S.)

Variable	Mean	Std. Dev.
RM	0.567	4.120
INF	0.295	0.905
DSPR	0.075	0.035
MDY	0.210	0.256
MVOL	3.546	1.803
IPG	0.303	1.180
SWG	0.226	0.565

Table 1 (Cont'd): Summary Statistics

Panel C: Industry Portfolio Returns (Rest of the World)																
Industry	United Kingdom		Australia		Canada		France		Germany		Japan		Netherlands		Switzerland	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
AERSP	0.632	8.514	-----	-----	-----	-----	0.681	10.021	-----	-----	-----	-----	-----	-----	-----	-----
AUTMB	0.267	8.812	0.326	7.998	0.399	13.516	0.470	9.035	0.249	7.444	0.506	6.572	0.805	6.235	0.000	6.127
BANKS	0.815	7.526	0.720	6.655	0.716	5.597	0.787	8.406	0.243	6.510	0.087	7.159	0.866	7.273	0.519	8.061
BEVES	0.637	6.804	0.502	7.344	1.058	6.841	0.725	7.768	-0.490	7.735	0.208	5.724	0.713	7.097	0.000	5.552
CHMCL	0.383	6.716	0.692	7.396	0.852	7.221	0.922	6.436	0.381	5.891	0.245	6.724	0.420	8.442	0.718	5.813
CNSBM	0.561	7.745	0.458	6.579	0.795	6.811	0.809	7.187	0.246	6.642	0.069	6.447	0.058	6.924	0.391	4.208
DIVIN	-----	-----	0.677	6.040	-----	-----	0.742	7.789	0.409	5.339	0.388	7.490	0.250	8.904	0.039	9.270
ELECT	0.633	5.499	0.583	8.328	0.381	10.932	-0.815	10.910	0.504	3.626	0.341	6.632	-----	-----	0.396	5.963
ELTNC	0.612	8.827	-----	-----	1.238	-----	0.838	9.017	0.434	8.100	0.402	6.568	0.367	9.380	0.249	9.226
ENGEN	0.351	7.616	-0.198	10.256	0.466	8.236	0.546	10.561	0.250	6.132	0.078	6.434	-0.082	9.890	0.193	5.275
FDRET	0.911	6.928	1.239	5.609	1.083	7.209	0.980	7.271	-0.118	7.976	1.310	8.064	0.790	7.730	0.298	9.536
FOODS	0.700	6.133	0.617	6.268	1.132	6.406	0.699	6.608	0.384	5.056	0.329	5.322	0.757	5.840	0.685	9.858
FSTPA	0.671	9.854	0.031	10.010	-----	-----	-----	-----	0.579	6.693	0.233	6.565	0.351	8.603	0.000	11.576
HLTHC	0.756	6.846	0.873	7.220	1.520	8.735	1.240	8.907	0.481	8.772	0.260	8.002	0.706	8.144	1.342	6.393
HHOLD	0.288	9.624	-6.239	25.792	1.423	8.605	0.376	8.100	0.137	6.683	0.244	6.890	0.788	7.704	0.709	6.709
INFOH	0.490	19.347	-2.007	25.732	0.166	10.276	0.723	11.341	-1.392	17.182	0.364	8.049	1.569	20.663	-0.361	8.302
INSUR	0.367	8.278	0.884	8.151	1.194	6.849	1.098	9.760	0.764	7.347	0.231	7.415	0.817	7.596	0.595	11.817
INVSC	0.621	6.833	0.879	6.642	0.094	5.556	0.886	10.254	-2.900	12.020	-12.702	22.318	0.445	3.795	0.418	19.680
LESUR	0.430	8.129	0.947	8.171	0.920	8.203	0.452	8.557	2.429	30.345	0.758	7.425	-2.096	9.289	0.305	5.731
LIFEA	0.688	8.286	0.220	7.050	0.950	7.329	0.399	7.959	0.773	8.504	-2.324	8.814	0.829	8.505	-0.375	2.500
MEDIA	0.541	8.055	1.378	12.275	0.689	6.742	0.698	9.030	-0.234	10.167	0.519	6.000	0.831	6.564	-1.057	7.920
MNING	0.701	9.292	0.629	7.767	0.413	8.882	-----	-----	-----	-----	-0.841	9.865	-----	-----	-----	-----
OILGS	0.799	7.234	0.509	10.221	0.485	7.099	0.860	7.888	-----	-----	0.178	8.609	0.697	6.066	-----	-----
PERSH	0.799	7.989	-----	-----	0.478	8.089	1.215	8.016	0.691	5.406	0.405	5.721	-----	-----	-----	-----
PHARM	1.160	7.383	1.695	11.912	2.286	13.112	0.971	7.500	0.469	6.012	0.679	6.064	-5.861	25.549	0.610	8.338
RLEST	0.456	7.798	0.664	6.511	-0.188	7.620	0.438	4.145	1.484	10.264	0.089	8.310	0.046	4.185	-0.187	8.850
RTAIL	0.592	7.434	0.687	6.989	0.234	6.416	0.858	8.655	0.114	7.121	0.282	6.155	0.195	9.037	-0.033	7.915
SFTCS	0.674	10.263	2.700	11.379	0.309	14.516	0.509	11.610	1.972	13.151	0.273	12.166	0.266	10.173	-----	-----
SPFIN	0.632	7.913	0.240	7.709	0.828	5.929	0.546	9.232	-0.694	8.873	0.208	8.473	0.438	11.250	1.591	9.357
STLOM	-0.767	12.776	0.526	10.737	0.227	7.321	0.377	10.195	-0.067	8.132	-0.001	7.880	-0.475	13.269	-----	-----
SUPSV	0.836	6.981	0.348	8.487	-0.151	6.564	0.979	7.961	-0.714	9.455	0.908	8.530	0.387	8.688	0.472	4.352
TELCM	0.989	8.009	1.331	7.665	0.642	5.234	-0.182	15.740	0.153	8.102	0.699	9.849	-0.188	15.091	-0.272	9.234
TOBAC	1.070	8.010	-----	-----	1.130	7.517	-----	-----	-----	-----	-0.342	6.977	-----	-----	-----	-----
TRNSP	0.370	6.890	0.843	7.015	1.553	9.100	-0.612	9.932	0.481	7.864	0.247	5.776	0.063	7.833	-0.665	4.345
UTILO	0.724	5.474	0.902	8.189	0.198	5.274	0.532	8.226	0.450	4.845	0.321	7.851	-----	-----	-----	-----

Table 1 (Cont'd): Summary Statistics

	United Kingdom		Australia		Panel D: Other Variables (Rest of the World)				Germany		Japan		Netherlands		Switzerland	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
RM	0.669	5.867	0.643	6.169	0.550	4.663	0.715	6.378	0.381	5.258	0.264	5.309	0.569	5.038	0.468	4.832
IPG	0.074	1.382	0.187	1.265	0.113	1.432	0.095	1.742	0.151	1.401	0.113	2.706	0.134	0.470	0.218	0.868

Table 2: Predictive Regressions Between Metal and Market Portfolios (U.S.)

This table presents the results from forecasting the market return in month t using variables at month $t-1$. METAL is the return on the primary metal industry portfolio. RM is the CRSP excess value-weighted market portfolio return. INF is the CPI inflation rate. DSPR is the default spread between BAA-rated and AAA-rated bonds. MDY is the dividend yield of the market portfolio. MVOL is the market volatility. In all columns, the least squares estimates, Newey-West t-statistics (in parentheses), adjusted R², and number of observations are displayed. Newey-West serial correlation and heteroskedasticity robust t-statistics are calculated with 3 monthly lags. The sample period is January 1946 to December 2002. *Significant at 10% level. **Significant at 5% level.

Dependent Variable ---RM (U.S.)			
	(1)	(2)	(3)
CONST	0.005 (3.074)**	-0.002 (-0.451)	-0.007 (-1.555)
METAL(-1)	-0.096 (-2.261)**	-0.077 (-2.107)**	-0.085 (-2.308)**
RM(-1)	0.053 (2.426)**	0.061 (1.752)*	0.052 (1.526)
INF(-1)		-0.578 (-3.292)**	-0.624 (-3.549)**
DSPR(-1)		0.888 (1.832)*	0.526 (1.069)
MDY(-1)		1.418 (2.342)**	1.412 (2.211)**
MVOL(-1)			0.241 (2.670)**
R ²	0.009	0.032	0.044
T	684	684	684

Table 3: Predictive Regressions Involving Various Industry and Market Portfolios

This table presents forecasts of the market return using various industry portfolio returns (separately) at various horizons: month t, month t+1 and month t+2. The other forecasting variables are INF (the CPI inflation rate), DSPR (the default spread between BAA-rated and AAA-rated bonds), MDY (the dividend yield of the market portfolio) and market volatility (MVOL). We only report the coefficients in front of the lagged industry return. The least squares estimates, Newey-West t-statistics (in parentheses), and adjusted R² are displayed. Newey-West serial correlation and heteroskedasticity robust t-statistics are calculated with 3 monthly lags. We also report Wald tests of the joint significance of the coefficients estimated in a GMM system (joint test of the null hypothesis that the lambdas are zero). The Wald tests (34 d.f.) are computed using Newey-West heteroskedasticity and serial correlation robust variance covariance matrix. The sample period is January 1946 to December 2002, except for RLEST, which is from 1972 to 2002. *Significant at 10% level. **Significant at 5% level.

Industry	Forecast of RM (U.S.) using Industry Returns at Various Horizons (H)					
	Month t		Month t+1		Month t+2	
	IND(-1)	R ²	IND(-1)	R ²	IND(-1)	R ²
RLEST	0.173 (2.643)**	0.053	-0.029 (1.034)	0.021	0.009 (0.324)	0.008
AGRIC	0.027 (0.964)	0.039	-0.010 (0.546)	0.024	-0.028 (0.998)	0.020
MINES	-0.067 (-1.999)**	0.044	-0.041 (1.708)*	0.033	-0.040 (1.659)*	0.031
OIL	-0.009 (-0.259)	0.038	-0.021 (0.932)	0.020	-0.006 (0.435)	0.002
STONE	-0.038 (-1.722)*	0.041	-0.039 (1.669)*	0.032	0.004 (0.352)	0.001
CNSTR	0.018 (0.474)	0.038	-0.020 (1.170)	0.024	-0.029 (1.133)	0.021
FOOD	0.011 (0.162)	0.038	0.039 (1.633)	0.028	-0.020 (0.594)	0.017
SMOKE	-0.025 (-0.702)	0.038	-0.001 (0.247)	0.012	-0.002 (0.102)	0.001
TXTLS	0.066 (1.574)	0.042	-0.005 (0.197)	0.009	-0.020 (0.604)	0.018
APPRL	0.093 (1.996)**	0.042	-0.008 (0.305)	0.004	-0.018 (0.564)	0.016
WOOD	-0.013 (-0.369)	0.038	-0.014 (0.982)	0.010	-0.005 (0.242)	0.002
CHAIR	0.030 (0.656)	0.038	0.001 (0.157)	0.006	-0.023 (0.734)	0.018
PAPER	-0.024 (-0.407)	0.038	-0.022 (0.687)	0.013	-0.050 (1.943)*	0.036
PRINT	0.140 (2.455)**	0.048	0.049 (1.867)*	0.035	-0.031 (1.232)	0.022
CHEMS	-0.056 (-0.704)	0.038	0.050 (1.895)*	0.034	0.035 (1.497)	0.024
PTRLM	-0.105 (-2.169)**	0.045	-0.040 (1.659)*	0.031	0.023 (0.643)	0.017
RUBBR	0.007 (0.136)	0.038	0.005 (0.495)	0.009	-0.018 (0.546)	0.014

Table 3 (Cont'd): Predictive Regressions Involving Various Industry and Market Portfolios

Forecast of RM (U.S.) using Industry Returns at Various Horizons (H)						
Industry	Month t		Month t+1		Month t+2	
	IND(-1)	R ²	IND(-1)	R ²	IND(-1)	R ²
LETHR	0.077 (2.252)**	0.042	-0.002 (0.294)	0.007	-0.015 (0.412)	0.010
GLASS	0.024 (0.492)	0.038	0.018 (0.648)	0.010	-0.026 (0.984)	0.020
METAL	-0.085 (-2.308)**	0.044	-0.045 (1.674)*	0.031	-0.042 (1.694)*	0.033
MTLPR	0.073 (1.001)	0.039	-0.023 (1.054)	0.020	-0.025 (0.878)	0.020
MACHN	0.067 (1.389)	0.041	0.021 (1.167)	0.021	0.032 (1.324)	0.024
ELCTR	0.035 (0.687)	0.038	0.003 (0.390)	0.004	-0.010 (0.435)	0.008
CARS	0.006 (0.109)	0.038	0.002 (0.278)	0.003	-0.031 (1.342)	0.023
INSTR	0.049 (0.885)	0.039	0.003 (0.305)	0.004	-0.005 (0.293)	0.002
MANUF	0.043 (1.106)	0.040	0.033 (1.280)	0.021	-0.015 (0.453)	0.009
TRANS	0.090 (2.142)**	0.040	-0.010 (0.592)	0.009	-0.033 (1.348)	0.024
PHONE	-0.018 (-0.419)	0.038	0.033 (1.378)	0.023	0.043 (1.721)*	0.035
TV	0.067 (1.769)*	0.042	0.043 (1.653)*	0.032	0.009 (0.193)	0.007
UTILS	0.088 (2.228)**	0.039	-0.035 (1.568)	0.025	0.023 (0.970)	0.018
WHLSL	0.047 (0.838)	0.039	-0.025 (1.328)	0.019	-0.033 (1.400)	0.023
RTAIL	0.091 (2.506)**	0.041	0.003 (0.246)	0.004	-0.029 (1.102)	0.018
MONEY	0.180 (2.187)**	0.046	0.058 (1.854)*	0.035	-0.019 (0.751)	0.015
SRVC	0.086 (2.223)**	0.043	0.051 (1.721)*	0.033	-0.012 (0.403)	0.011
Wald test (p-value)	75.830 <0.01		47.017 0.068		27.193 0.789	

Table 4: Industry Predictive Regressions--Economic Significance

In Panel A, the column “Economic Significance” computes the response of the market return to a two-standard-deviation shock of the corresponding industry return using the point estimates from column (1) of Table 3. Lower and upper bounds are given in parentheses below. The column “Absolute Relative Significance” computes the absolute value from dividing “Economic Significance” by the standard deviation of the market return. Same calculations are done for the other market predictors in Panel B. In Panel C, we provide performance statistics for portfolios (market and risk-free rate) formed based on the various market predictors.

Panel A: By Industries

	Economic Significance	Absolute Relative Significance		Economic Significance	Absolute Relative Significance
MONEY	1.741 (0.149, 3.332)	0.422	MANUF	0.546 (-0.441, 1.532)	0.132
RLEST	1.513 (0.368, 2.658)	0.367	INSTR	0.526 (-0.663, 1.714)	0.128
PRINT	1.496 (0.277, 2.714)	0.363	WHLSL	0.521 (-0.722, 1.764)	0.126
APPRL	1.222 (-0.002, 2.447)	0.297	CHEMS	-0.514 (-1.976, 0.947)	0.125
SRVC	1.121 (0.112, 2.130)	0.272	ELCTR	0.436 (-0.834, 1.707)	0.106
METAL	-1.039 (-1.939, -0.139)	0.252	AGRIC	0.391 (-0.420, 1.202)	0.095
PTRLM	-1.040 (-1.999, -0.081)	0.252	CHAIR	0.331 (-0.679, 1.342)	0.080
TRANS	1.031 (0.068, 1.994)	0.250	SMOKE	-0.283 (-1.090, 0.524)	0.069
LETHR	0.960 (0.107, 1.812)	0.233	GLASS	0.282 (-0.863, 1.426)	0.068
RTAIL	0.932 (0.188, 1.675)	0.226	PAPER	-0.256 (-1.516, 1.003)	0.062
TV	0.900 (0.117, 1.917)	0.218	CNSTR	0.250 (-0.804, 1.303)	0.061
MINES	-0.839 (-1.678, 0.000)	0.204	WOOD	-0.189 (-1.213, 0.835)	0.046
TXTLS	0.783 (-0.212, 1.779)	0.190	PHONE	-0.169 (-0.976, 0.638)	0.041
MACHN	0.780 (-0.343, 1.903)	0.189	OIL	-0.120 (-1.047, 0.807)	0.029
MTLPR	0.712 (-0.711, 2.134)	0.173	FOOD	0.093 (-1.060, 1.247)	0.023
UTILS	0.681 (0.070, 1.292)	0.165	RUBBR	0.084 (-1.155, 1.323)	0.020
STONE	-0.582 (-1.257, 0.094)	0.141	CARS	0.065 (-1.125, 1.255)	0.016

Table 4 (Cont'd): Industry Predictive Regressions--Economic Significance

Panel B: By Other Market Predictors

	Economic Significance	Absolute Relative Significance
RM(-1)	0.428 (0.990, -0.133)	0.104
INF(-1)	-1.129 (-1.765, -0.493)	0.274
DSPR(-1)	0.036 (0.104, -0.032)	0.009
MDY(-1)	0.723 (0.069, 1.377)	0.175
MVOL(-1)	0.869 (0.218, 1.520)	0.211

Panel C: Performance Statistics of Various Portfolios (figures in annualized %, with the exception of the Relative MSE)

	Market	Timing Market and Risk- Free Using only Z_{t-1}	Timing Market and Risk- Free Using Z_{t-1} and R_{t-1}
Relative MSE	-	1.000	0.985
Mean	13.61	11.62	9.55
Standard Deviation	15.62	10.65	6.10
Sharpe Ratio	0.46	0.49	0.52
Certainty Equivalent (Relative Risk Aversion=1)	8.24	9.40	10.99
Certainty Equivalent (Relative Risk Aversion=5)	7.93	8.32	8.64

Table 5: Predictive Regressions of Measures of Economic Activity using Industry Portfolios

Panel A presents results of forecasting IPG, industrial production growth in month t, using various industry portfolio returns at month t-1 separately and other information available at month t-1. Panel B presents analogous results from forecasting SWG, the growth rate of the Stock and Watson (1989) index of economic activity in month t. The other forecasting variables are lagged RM, INF, DSPR, MDY, MVOL, and three monthly lags of IPG (and correspondingly SWG). We only report the coefficient in front of the lagged industry return. The least squares estimates, Newey-West t-statistics (in parentheses), and adjusted R2 are displayed for each industry. Newey-West serial correlation and heteroskedasticity robust t-statistics are calculated with 3 monthly lags. We also report Wald tests of the joint significance of the coefficients estimated in a GMM system (joint test of the null hypothesis that the lambdas are zero). The Wald tests (34 d.f.) are computed using Newey-West heteroskedasticity and serial correlation robust variance covariance matrix. The sample period for the IPG regressions is January 1946 to December 2002 and for SWG is March 1959 to December 2002, with the exception of RLEST, which is from January 1972 to December 2002. *Significant at 10% level. **Significant at 5% level.

Panel A: Forecast of IPG (U.S.) using Industry Returns					
Industry	IND(-1)	R2	Industry	IND(-1)	R2
RLEST	0.020 (1.889)*	0.209	LETHR	0.016 (1.675)*	0.109
AGRIC	0.002 (0.331)	0.106	GLASS	0.012 (0.934)	0.107
MINES	-0.001 (-0.137)	0.106	METAL	-0.023 (-1.999)**	0.111
OIL	0.002 (0.171)	0.106	MTLPR	0.006 (0.346)	0.106
STONE	-0.009 (-1.330)	0.108	MACHN	0.017 (1.363)	0.108
CNSTR	-0.021 (-2.227)**	0.112	ELCTR	0.031 (2.438)**	0.113
FOOD	-0.037 (-2.140)**	0.112	CARS	0.038 (2.868)**	0.116
SMOKE	-0.012 (-1.309)	0.108	INSTR	-0.026 (-1.934)*	0.110
TXTLS	0.027 (2.649)**	0.115	MANUF	0.006 (0.563)	0.106
APPRL	0.012 (1.237)	0.108	TRANS	0.014 (1.062)	0.107
WOOD	0.010 (1.202)	0.107	PHONE	-0.004 (-0.382)	0.106
CHAIR	0.005 (0.475)	0.106	TV	0.006 (0.623)	0.106
PAPER	0.009 (0.604)	0.106	UTILS	0.042 (2.827)**	0.116
PRINT	-0.021 (-1.489)	0.108	WHLSL	0.007 (0.499)	0.106
CHEMS	-0.041 (-2.087)**	0.111	RTAIL	0.009 (0.606)	0.106
PTRLM	-0.002 (-0.175)	0.106	MONEY	0.052 (2.532)**	0.114
RUBBR	0.014 (1.150)	0.107	SRVC	0.011 (1.029)	0.107
Wald Test	61.292				
p-value	<0.01				

Table 5 (Cont'd): Predictive Regressions of Measures of Economic Activity using Industry Portfolios

Panel B: Forecast of SWG (U.S.) using Industry Returns					
Industry	IND(-1)	R2	Industry	IND(-1)	R2
RLEST	0.011 (0.100)	0.215	LETHR	0.004 (0.896)	0.211
AGRIC	-0.005 (-1.350)	0.213	GLASS	0.006 (1.020)	0.212
MINES	-0.004 (-0.863)	0.211	METAL	-0.021 (-0.054)	0.210
OIL	0.021 (0.041)	0.210	MTLPR	-0.001 (-0.156)	0.210
STONE	-0.003 (-0.816)	0.211	MACHN	0.013 (2.123)**	0.217
CNSTR	0.009 (0.080)	0.210	ELCTR	0.017 (2.718)**	0.221
FOOD	-0.020 (-2.370)**	0.218	CARS	0.013 (1.933)*	0.216
SMOKE	-0.004 (-1.002)	0.211	INSTR	0.003 (0.498)	0.210
TXTLS	0.010 (1.923)*	0.216	MANUF	0.005 (1.016)	0.211
APPRL	0.006 (1.259)	0.212	TRANS	-0.004 (-0.652)	0.211
WOOD	0.001 (0.341)	0.210	PHONE	-0.004 (-0.716)	0.211
CHAIR	0.002 (0.266)	0.210	TV	0.006 (1.109)	0.212
PAPER	0.004 (0.538)	0.210	UTILS	0.017 (2.368)**	0.218
PRINT	-0.001 (-0.109)	0.210	WHLSL	0.002 (0.264)	0.210
CHEMS	-0.021 (-2.111)**	0.217	RTAIL	0.004 (0.527)	0.210
PTRLM	-0.043 (-0.026)	0.210	MONEY	0.017 (1.598)	0.214
RUBBR	0.010 (1.679)*	0.214	SRVC	0.001 (0.136)	0.210
Wald test (p-value)	63.396 <0.01				

Table 6: Predictive Regressions of Market Returns and Economic Activity using All Industry Returns Simultaneously (U.S.)

This table presents results of forecasting the market and various indicators of economic activity in month t using all industry portfolio returns at month $t-1$ (excluding RLEST) and other information available at month $t-1$. RM is the market return, IPG is industrial production growth and SWG is the growth rate of the Stock and Watson (1989) index of economic activity. The other control variables are lagged RM, INF (the CPI inflation rate), DSPR (the default spread between BAA-rated and AAA-rated bonds), MDY (the dividend yield of the market portfolio), MVOL (the market volatility), and in the cases of IPG and SWG, three lags of the dependent variable. In Panel A, for each of the three regressions, we test three hypotheses: (1) whether the coefficients on the lagged industry returns are jointly zero; (2) whether the coefficients on the other lagged control variables are jointly zero; and (3) whether the coefficients on the lagged industry returns and the lagged controls are jointly zero. The F-tests and p-values are reported below. The degrees of freedom of the F-tests are determined as in Hayashi (2000, p. 43) In all tests, we use Newey-West serial correlation and heteroskedasticity robust standard errors calculated with 3 monthly lags. In Panel B, we report the parameter of proportionality obtained from estimating the predictive regressions of RM and IPG using all industry returns jointly and testing whether the predictive coefficients from these two regressions are proportional. *Significant at 10% level. **Significant at 5% level.

Panel A: Forecasting using All Industry Returns Simultaneously						
Null Hypothesis	RM		IPG		SWG	
	F-test	P-value	F-test	P-value	F-test	P-value
Coefficients of all industries equal zero	1.947	<0.01	2.085	<0.01	1.652	0.015
Coefficients of all controls equal zero	2.236	<0.01	8.482	<0.01	8.446	<0.01
Coefficients of all industries and controls equal zero	2.321	<0.01	2.828	<0.01	2.450	<0.01

Panel B: Alternative Method of Estimating the Relation between an Industry's Ability to Lead the Market and Its Ability to Predict Fundamentals		
	RM and IPG	RM and SWG
Parameter of Proportionality	2.142 (2.644)**	3.620 (2.488)**

Table 7: Predictive Regressions of Market Returns and Economic Activity by Industry Returns (International Evidence)

This table presents results of forecasting the market (RM) and industrial production growth (IPG) in month t using industry portfolio returns at month $t-1$ and lagged RM as a control for each of the eight largest markets outside of the U.S. In Panel A, we report the results of forecasting the market and IPG using individual industry portfolios separately with the lagged market as a control: the number of industry portfolios that can significantly predict the market at the 10% level of significance. In Panel B, we report the F-test of forecasting the market and IPG using all industry portfolios simultaneously with the lagged market as a control. The F-tests and p-values are reported below. The degrees of freedom of the F-tests are determined as in Hayashi (2000, p. 43) In all tests, we use Newey-West serial correlation and heteroskedasticity robust standard errors calculated with 3 monthly lags. In Panel C, we report the parameter of proportionality obtained from estimating the predictive regressions of RM and IPG using all industry returns jointly and testing whether the predictive coefficients from these two regressions are proportional. The sample is from 1973 to 2002. *Significant at 10% level. **Significant at 5% level.

Country	Panel A: Individual Regressions		Panel B: Joint Regressions	
	RM	IPG	RM	IPG
	At 10 percent	At 10 percent	F-test	F-test
United Kingdom	8 of 34	6 of 34	4.584**	5.753**
Australia	18 of 31	11 of 31	5.458**	6.764**
Canada	11 of 31	12 of 31	5.049**	6.077**
France	14 of 32	12 of 32	4.923**	6.983**
Germany	12 of 31	10 of 31	5.167**	5.403**
Japan	8 of 34	10 of 34	4.735**	5.790**
Netherlands	16 of 29	11 of 29	5.274**	6.248**
Switzerland	14 of 27	10 of 27	4.753**	6.452**

Country	Panel C: Parameter of Proportionality (RM and IPG)	
	Parameters	T-statistic
United Kingdom	2.491	(3.795)**
Australia	1.036	(2.159)**
Canada	1.343	(3.353)**
France	1.634	(3.140)**
Germany	1.517	(3.037)**
Japan	-0.780	(-1.190)
Netherlands	1.004	(2.450)**
Switzerland	3.248	(5.453)**

Figure 1: The Relationship between an Industry's Ability to Lead the Market and Its Ability to Predict Economic Activity (U.S.)

Figure 1(a) presents a scatter-plot of the coefficients λ_i obtained by forecasting RM using 34 industry returns and other conditioning information on the coefficients γ_i obtained by forecasting industrial production growth (IPG) using the same 34 industry returns. The linear relationship between the two sets of coefficients is plotted with a solid line. The slope of the line, Newey-West t-statistic, and R2 are also presented. Figures 1(b) display the same relationship between the coefficients λ_i and γ_i , where γ_i are obtained by forecasting the growth rate of the Stock and Watson (1989) index of economic activity (SWG).

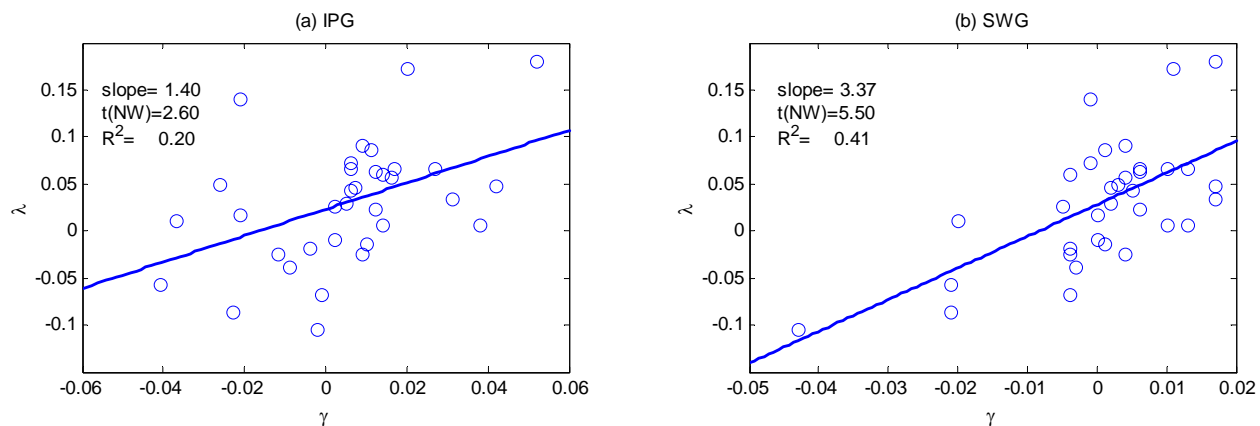
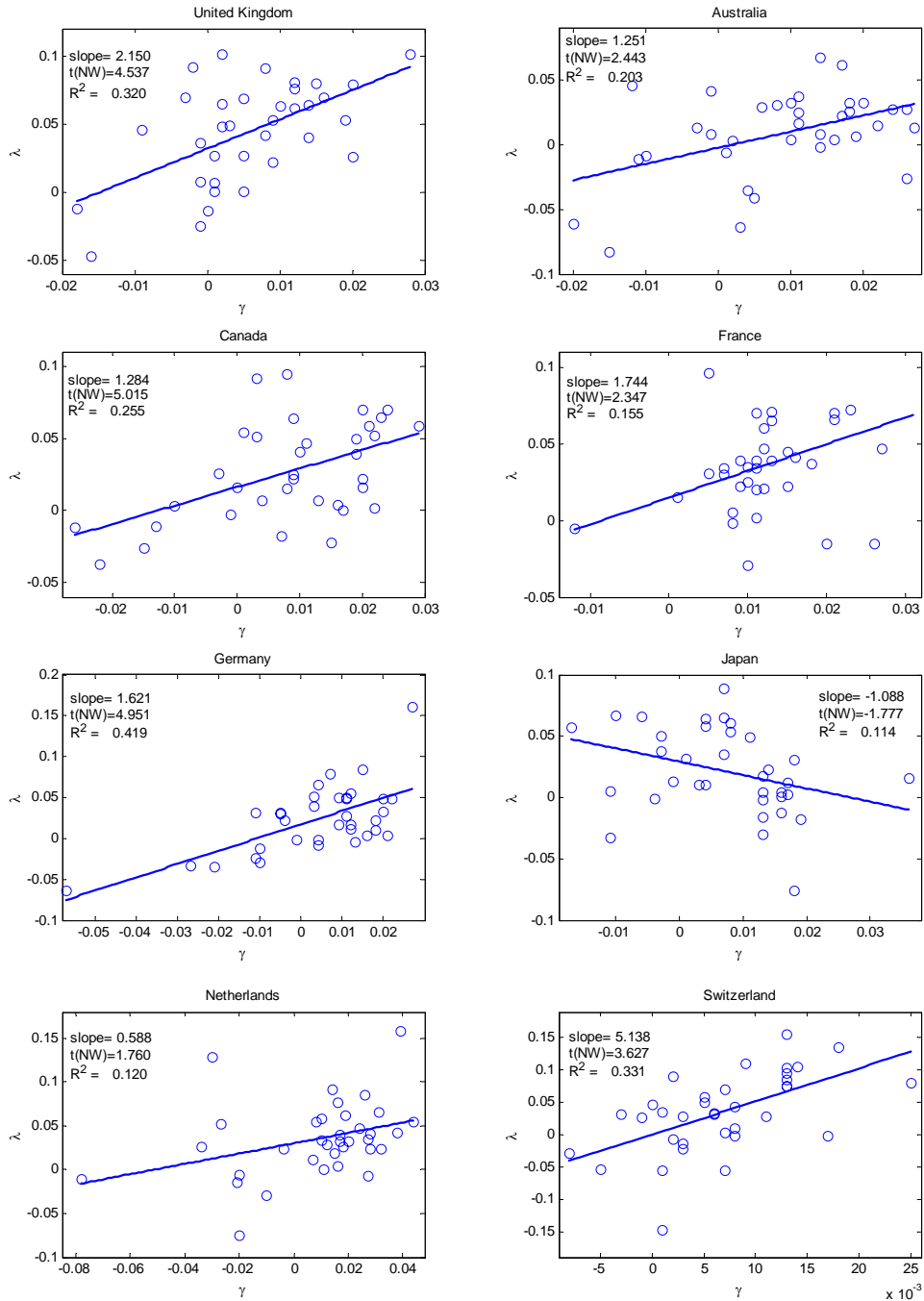


Figure 2: The Relationship between an Industry's Ability to Lead the Market and Its Ability to Predict Economic Activity (International Evidence)

This figure presents a scatter-plot, for each of the eight largest stock markets outside of the U.S., of the coefficients λ_i obtained by forecasting RM using industry returns and other conditioning information on the coefficients γ_i obtained by forecasting industrial production growth (IPG) using the same industry returns. The linear relationship between the two sets of coefficients is plotted with a solid line. The slope of the line, Newey-West t-statistic, and R2 are also presented.



Appendix: US and International Industry Portfolio Dataset

Detailed definitions of the US industry portfolios are available at Kenneth French's website http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

The international industry portfolio returns, market returns, and industrial production growth data is from the world, region, and country code indices of Datastream International. The international industry acronyms are:

ACRONYM	Description
1 AERSP	Aerospace and Defense.
2 AUTMB	Automobiles, Auto Parts, Tires & Rubber, Vehicle Distribution.
3 BANKS	Banks.
4 BEVES	Beverages & Brewers & Distillers & Vintners Distillers & Vintners& Soft Drinks.
5 CHMCL	Chemicals, Commodity Chemicals, Speciality Chemicals, Advanced Chemicals.
6 CNSBM	Builders Merchants, Building & Building Materials, Construction Materials, Building, Other Construction.
7 DIVIN	Diversified, Industrials.
8 ELECT	Electricity.
9 ELTNC	Electronic & Electrical Equipment.
10 ENGEN	Engineering, Commercial Vehicles, Machinery Engineering, Contractors, Fabricators.
11 FDRET	Food & Drug, Food & Drug Retailers.
12 FOODS	Food, Farming & Fishing Producers, Food Processors.
13 FSTPA	Forestry, Paper and Paper Processing.
14 HLTHC	Health, Health Maintenance, Organizations, Hospital Management, Medical Equipment, Medical Equipment & Supplies, Other Health Care.
15 HHOLD	Clothing & Footwear, Clothing & Footwear Goods, Furnishings & Floor, Textiles coverings , Consumer Electronics, Household Appliances & Houseware, Leisure Equipment, Textiles & Leather.
16 INFOH	Information, Computer Hardware, Semiconductors, Hardware Telecom Equipment.
17 INSUR	Insurance, Brokers, Insurance Brokers, Insurance Non-Life, Re-insurance, Other Insurance.
18 INVSC	Investment, Investment Companies, Investment Trust, International Investment. Investment Trust, Emerging Markets, Venture Investment Companies, Exchange Traded Funds, Open Ended investment.
19 LESUR	Leisure, Gambling & Hotels, Leisure Facilities, Restaurants & Pubs.
20 LIFEA	Life Assurance Companies.
21 MEDIA	Media, Television, Radio, Entertainment, Filmed Entertainment, Entertainment Networks, Media Agencies, Photography, Publishing & Printing.
22 MNING	Mining, Gold Mining, Mining Finance, Other Mineral, Extractors.
23 OILGS	Oil & Gas, Exploration, Production, Oil Services, Oil Integrated.
24 PERSH	Personal Care Household Products, Personal Products.
25 PHRMC	Pharmaceuticals, Biotechnology.
26 RLEST	Real Estate Development, Development Property Agencies, Real Estate Investment Trusts.
27 RTAIL	Retailers, Discount & Super, Stores & Warehouses, Retailers e-commerce, Retailers, Hardline Retailers, Multi Department, Retailers-Soft Goods.
28 SFTCS	Software & Computer Services, Internet Services, Software.
29 SPFIN	Speciality & Asset Managers, Asset Managers, Other Consumer Finance, Consumer Finance, Finance, Investment Banks, Mortgage Finance, Other Financial.
30 STLOM	Steel & Non-Ferrous Metals, Other Metals.
31 SUPSV	Business Support, Business Support Services, Delivery Services, Education, Training, Environmental Control, Environmental Control, Transaction & Payroll Services, Security & Alarm Services.
32 TELCM	Telecom, Telecom Fixed Line Services, Telecom Wireless.
33 TOBAC	Tobacco.
34 TRNSP	Transport, Airlines & Airports, Rail, Road & Freight, Shipping & Ports.
35 UTILO	Gas Distribution, Multi-Utilities, Water.