Three Essays on Financial Markets

by

Lu Zhang

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Finance

Faculty of Business
University of Alberta

© Lu Zhang, 2015

Abstract

This thesis consists of three essays. The first essay studies the ability of stock return idiosyncrasy to predict future economic conditions over time. The second essay investigates the technological innovation and creative destruction during the 1920s and the 1930s, one of the most innovative periods in the 20th century. The third essay tests the performance of an investment strategy using information about past market-wide comovement.

Stock return idiosyncrasy, defined as the ratio of firm-specific to systematic risk in individual stock returns, contains information about future growth rate in real GDP, industrial production, real fixed assets investment, and unemployment. Forecasts are generally significant one-quarter-ahead, particularly after World War II. These effects persist after controlling for other potential leading economic indicators, both in-sample and out-of-sample. These findings are consistent with information generating firms, presumably uniquely well-informed about economic conditions because their core business is information, adjusting their information production before downturns.

The second essay studies the process of creative destruction during the technological revolution in the 1920s and 1930s. Intensified creative destruction magnifies the performance gap between winner and loser firms, and thus elevates firm-specific stock return variation. We find high firm-specific return variation in innovative industries and firms during the 1920s boom and the subsequent depression. We also find some evidence of elevated firm-specific return variation in manufacturing sectors with higher labor productivity, more research staff and more extensive electrification.

In the third essay, we define the directional market-wide comovement measure as the proportion of stocks moving up together. Positing that high comovement reflects large fund inflows, we devise an investment strategy of entering the market whenever positive directional market-wide comovement passes a certain threshold. Specifically, this comovement-based investment strategy holds the market index when the market-wide upward comovement in the prior one to four weeks is above the fourth decile of the historical comovement distribution, and invests in the risk-free asset otherwise. During the sample period of 1954 to 2014, this strategy outperforms the NYSE value-weighted market index by 6.42% per year. Out of sample tests using NASDAQ stocks and TSE stocks validate the strategy. Our findings suggest that market-wide upward comovement identifies periods of market run-ups, when unsophisticated investor buying is apt to be driven by herding or information cascades.

Preface

Part of the research conducted for this thesis is joint work. Chapter 2 is joint research with Dr. Randall Morck and Chapter 3 is with Dr. Mark Huson and Dr. Randall Morck, both at University of Alberta. In both papers, we co-developed the research ideas and methodologies. I was responsible for the data collection and analysis, and contributed the majority of the manuscript composition.

Acknowledgments

I would like to thank my advisor Dr. Randall Morck, for his invaluable mentorship and support in developing this thesis. I have benefited greatly from his knowledge, experience, and enthusiasm. I am grateful to the guidance of Dr. Mark Huson in developing the third chapter of this thesis. I thank Dr. David McLean and Dr. Rick Szostak for serving on my dissertation committee and Dr. William Goetzmann as the external examiner. I would like to give special thanks to Dr. Alice Nakamura for her support and help in my scholarly development and job search.

I have also benefited from the faculty and my fellow PhD students of the Finance group at the University of Alberta. I thank the support staff at the Finance Department and the Ph.D. Office for their assistance.

Finally, I am grateful to my family and friends, for their encouragement and support.

Table of Contents

List of Tables

List of Figures

Introduction

Stock return variation can be decomposed into firm-specific and systematic components. Roll (1988) observes that a significant proportion of the return variation in U.S. stocks is firm-specific. Morck, Yeung, and Yu (2000) and Campbell et al. (2001) document rising firm-specific return variation in the second half of the twentieth century in the U.S. The first two chapters of this thesis explore the relationship between firm-specific return variation and economic dynamism over time and across firms, respectively. The third chapter designs an investment strategy that profits from time varying systematic return variation.

Chapter 1 studies stock return idiosyncrasy over the business cycle. A variance decomposition of U.S. firm-level stock returns from 1921 to 2011 reveals that return idiosyncrasy, defined as the ratio of firm-specific to market-related return volatility, statistically significantly predicts the growth rates of real GDP, industrial production, real fixed assets investment, and the unemployment rate, at least one-quarter-ahead. These findings give new empirical support for Veldkamp's (2005, 2006) papers that information production specialists provide less firm-specific information when the economy is weaker. These findings further suggest that this shift towards producing less firm-specific information precedes actual economic downturns.

Chapter 2 focuses on one of the major business cycles of the past century: the 1920s boom and the subsequent depression in the 1930s. We investigate the cross-sectional difference in firm-specific return variation during this era of intensive innovative era. We argue that elevated firm-specific return variation reflects a widened performance gap between successful and failing firms, and thus can measure the intensity of creative destruction. We find high firm-specific return variation in innovative industries and firms during this time period, consistent with Schumpeter's (1912) creative destruction occurring during this time. New and creative firms arise to successfully apply new technologies, thereby destroy non-innovative firm and unsuccessful innovators.

Previous studies offer explanations for rising firm-specific volatility in the second half of the 20th century. Candidates include intensified innovation (Chun et al., 2008, 2011), more active arbitrage trading (Bris et al., 2007), increased number of young listed firms (Pastor and Veronesi, 2003; Fama and French, 2004; Brown and Kapadia, 2007), more volatile cash flows (Pastor and Veronesi, 2003; Wei and Zhang, 2006), and stronger competition (Bennett and Sias 2006; Irvine and Pontiff, 2009). The time series and cross-sectional evidence of my papers helps to reconcile these seemingly inconsistent findings: all of the above are characteristics of intensified creative destruction (Morck et al., 2013).

Chapter 3 investigates the ability of systematic return variation to predict market returns over short horizons. We define a directional market-wide upward comovement measure as the proportion of stocks moving up together. We find that the strategy of using upward comovement in the prior week as the signal to be in the NYSE value-weighted index, as opposed to being in three-month T-bills, generates an annualized alpha of 6.42% from 1954 to 2014. Repeating the exercise using stocks listed on NASDAQ or on the Tokyo Stock Exchange also yields positive and significant alphas.

Our findings that high upward comovement predicts high future returns in short windows are consistent with both the limited arbitrage theory and the information cascades model. These theories suggest that the coordinated behavior among investors can affect asset prices. The upward comovement measure detects periods of investors following each other into stock markets, due to rational or behavioral reasons, and profits from riding the waves of market run-ups.

Chapter 1 Stock Return Idiosyncrasy and the Business Cycle

1.1 Introduction

The firm-specific (idiosyncratic) component of stock return volatility is known to vary with time. Schwert (1989) finds elevated overall volatility during economic downturns, while Morck, Yeung, and Yu (2000) and Campbell, Lettau, Malkiel and Xu (2001) both document rising firm-specific volatility during the so-called Great Moderation, the mid to late 20th century, though the former also document high firm-specific volatility in the 1920s boom. Motivated by the financial historian Charles Kindleberger's (1976) analysis of centuries of financial boom and bust cycles in many different countries, this paper explores how the ratio of firm-specific volatility to systematic volatility, i.e. stock return idiosyncrasy, changes with the business cycle and whether this ratio contains useful information about future economic conditions.

Kindleberger posits that a boom and bust cycle begins with a *displacement*, an exogenous shock to the economy, usually a new technology, but sometimes a political shock or other abrupt change in the business environment. The displacement puts the economy in disequilibrium, letting entrepreneurs (and/or early investors who finance them) earn economic profits (abnormal returns) until equilibrium is regained. Chun et al. (2008) interpret high firm-specific return variation (returns heterogeneity) as reflecting a broadened performance gap between firms successfully taking advantage of these profit opportunities and firms left behind.

The next phase of the Kindleberger cycle is a positive feedback loop: uninformed capital, seeking high returns, flows into the affected sectors (or the whole economy if the displacement is at that level), lifting asset prices above fundamentals. While Kindleberger explicitly mentions behavioral biases (mania) in this context, information cascades or other forms of rational herding (Bikhchandani, Hirshleifer and Welch, 1992, 1998) might also arise. I hypothesize that the inflows of uninformed capital elevate stock return co-movement or reduce idiosyncrasy in this stage of the cycle.

Ultimately, the mania ends and a panic phase begins: uninformed investors, often quite abruptly, recognize asset prices as unrealistically elevated and rush to sell. Again, a feedback loop can arise as dropping asset prices panic other investors, whose selling further depresses asset prices. This, Kindleberger notes, can also paralyzes financial institutions, constricts normal financing channels to fundamentally sound firms, and further depresses valuations. Because these patterns of events all involve large numbers of investors trading in the same ways at the same times, I hypothesize that these phases might also be characterized by elevated returns co-movement.

In an ensuing recession, with uninformed traders' presence in the market diminished, prices might readjust idiosyncratically as arbitrageurs regain dominance over the market (Bris et al., 2007) and share prices become more information-laden. Firms that successfully adapted to the new technology or other shift in fundamentals that set the chain of events in motion emerge reinvigorated, though financial system constrictions can harm even fundamentally successful adaptors. Firms that failed to adapt successfully (or even fraudulently sought to disguise their failures) emerge more damaged. I hypothesize that firm-specific variation in stock returns should reemerge as this happens.

Kindleberger documents this cycle repeating at irregular intervals over centuries and across all free-market economies, and suggests that it is an unavoidable feature of a capitalist economy.

Figure 1.1 shows the time series pattern of stock return idiosyncrasy over U.S. boom and bust cycles from 1921 to 2011. Notice that stock return idiosyncrasy falls before each recession, suggesting that falling idiosyncrasy mighty predict downturns. In other contexts, Scheffer et al. (2009) model abruptly elevated comovement within a complex adaptive system as presaging an approaching "tipping point" and posit ecosystems, neural networks, and economic systems as potential applications of their model.

A variance decomposition of U.S. firm-level stock returns from 1921 to 2011, following the approaches of either Morck et al. (2000) or Campbell et al. (2001), reveals that return idiosyncrasy, defined as the ratio of firm-specific to market-related return volatility, statistically significantly predicts the growth rates of real GDP, industrial production, and real fixed assets investment, as well as the unemployment rate, at least one-quarter-ahead. This predictability is especially strong in the postwar period. Granger causality tests show that past stock return idiosyncrasy predicts future economic activity, but that the converse is not true.

These results are robust to controlling for other documented leading indicators of economic conditions: term spread (Harvey, 1988; Stock and Watson, 1989), default spread (Fama, 1986), stock market index returns (Fisher and Merton, 1984; Barro, 1990), stock market volatility (Campbell et al., 2001), dividend yield (Chen, Roll, and Ross, 1986), and stock market liquidity (Næs, Skjeltorp, and Ødegaar, 2011). This literature is largely consistent with the stock market serving as a leading indicator of economic activity, however Samuelson's (1966) quip that "Wall Street indices predicted nine out of the last five recessions!" could still be made today. The stock market crashes of 1929 and 2008 foreshadowed a deep recession, but those of 1907, 1987, and 2000 did not. This paper shows that stock return idiosyncrasy helps predict the business cycle more reliably. In fact, the idiosyncrasy measures significantly improve the performance of forecasting models of economic conditions out-of-sample and up to four quarters ahead.

This paper also contributes to the literature on firm-specific volatility by linking changes in firm-specific volatility to phases of the business cycle. Various studies explain rising firm-specific volatility in the second half of the 20th century (Morck et al., 2000; Campbell et al., 2001). Candidates are intensified innovation (Chun et al., 2008, 2011), increased number of young listed firms (Pastor and Veronesi, 2003; Fama and French, 2004; Brown and Kapadia, 2007), more volatile cash flows (Pastor and Veronesi, 2003; Wei and Zhang, 2006), stronger competition (Bennett and Sias 2006; Irvine and Pontiff, 2009); and all of the above as characteristics of intensified creative destruction (Morck et al., 2013).[1]

The findings in this paper are also consistent with the models of Veldkamp (2005) and Van Nieuwerburgh and Veldkamp (2006). They argue that information intermediaries, profit maximizing firms that produce information and sell it to investors, generate more and higher quality firm-specific information during booms than during recessions. Consistent with investors obtaining less or less reliable firm-specific information in downturns, I replicate the Brockman, Liebenburg, and Schutte (2011) finding of higher returns idiosyncrasy during downturns. My finding that stock return idiosyncrasy also serves as a leading indicator of future economic conditions is consistent with information generating firms, presumably themselves uniquely well-informed about economic conditions because their core business is information, adjusting their information production before downturns.

---

[1] Davis et al. (2006) find that these findings do not carry across to unlisted firms, reaffirming Schumpeter's (1911) view of the importance of risk-tolerant equity finance for new firms in the process of creative destruction and subsequent work along similar lines (e.g. Levine and King 1993).

The rest of this paper is organized as follows. Section 1.2 defines the stock return idiosyncrasy measures and business cycle indicators, and summarizes the data. Section 1.3 presents in-sample regressions and Granger causality test results. Section 1.4 conducts out-of-sample analysis and Section 1.5 concludes.

1.2 Variables Definition and Data Summary

1.2.1 **Measures of Stock Return Idiosyncrasy**

I use two methods to decompose stock return volatility into a systematic and a firm-specific component. The first, following Morck et al. (2000), uses the regression

[1] $$R_{js} = \alpha_{jt} + \beta_{m,jt} R_{ms} + \beta_{i,jt} R_{is} + \epsilon_s \ ,$$

where R_{js} is the return of stock *j* on day *s*, and R_{ms} and R_{is} are value-weighted market and industry returns on day *s*. Stock *j*'s return is excluded in the calculation of both the market and industry returns. The stock's systematic return volatility is defined as the sum of squared variation explained by the regression, *SSM*; and the stock's firm-specific volatility is the sum of squared variation attributed to the regression errors, *SSE*. That is,

[2] $$SSM_{jt} = \sum_{s \in t} \left(\hat{R}_{js} - \bar{R}_{js} \right)$$

[3] $$SSE_{jt} = \sum_{s \in t} \left(\hat{R}_{js} - R_{js} \right)$$

Stock *j*'s *idiosyncrasy ratio*, denoted $\phi_{j,t}$, is defined as the log of the ratio of its firm-specific volatility over its systematic volatility – that is, $\phi_{j,t}$= *log(SSE/SSM)*. To obtain aggregate measures of *SSE*, *SSM*, and idiosyncrasy, I take value-weighted averages of firm-level measures.

Campbell et al. (2001) develop an alternative decomposition methodology that decomposes firm-level stock return volatility into three components: market volatility, industry volatility, and firm-specific volatility. On day s of quarter t, the market return R_{ms} is the value-weighted average return of all common stocks, i.e. $R_{ms} = \sum_j w_j R_{js}$ where the weight, w_j, is the market capitalization of firm j as a fraction of the value of the market portfolio. Likewise, each industry return R_{is} is the value-weighted average return of industry i, i.e. $R_{is} = \sum_{j \in i} w_j R_{js}$. The market volatility of quarter *t* is then defined as:

[4] $$MKT_t = \sum_{s \in t} (R_{ms} - \mu_{mt})^2,$$

where μ_{mt} is the average daily market index return in quarter t. Industry-specific return volatility for industry i is defined as the sum of the squares of the difference between industry and market returns, and industry-specific volatility for the whole market IND_t is the weighted average of this across all the industries:

[5] $$IND_{it} = \sum_{s \in t} (R_{is} - R_{ms})^2$$

[6] $$IND_t = \sum_i w_{it} IND_{it}$$

Firm-specific volatility is first calculated at the firm-level as the sum of the squares of the firm's returns residuals after subtracting the industry returns. A value-weighted average is taken at the industry and then market level.

[7] $$FIRM_{jit} = \sum_{s \in t} \left(R_{jis} - R_{is}\right)^2,$$

[8] $$FIRM_t = \sum_i w_{it} \sum_{j \in i} w_{jit} FIRM_{jit}$$

Using this decomposition methodology, idiosyncrasy is again the log ratio of firm-specific volatility to market volatility, *log(FIRM/MKT)*. This approach has the advantage that it does not depend on any specific regression model, but the drawback that differences in individual stocks' sensitivities to systematic risk are suppressed. Campbell et al. (2001) show that this suppression is unlikely to induce serious bias in time-series analysis of market-wide idiosyncrasy measures.

The major functional difference between these two idiosyncrasy measures is the treatment of the industry component. In the regression based method, industry volatility is taken as part of systematic risk. While in the Campbell et al. (2001) methodology, industry volatility is a separate component and not included in the idiosyncrasy ratio. Tests using the two alternative measures generate very similar results. This is perhaps because the industry-level volatility from the Campbell et al. methodology is far smaller than the other two components. The correlation between the two idiosyncrasy measures is as high as 84.5% and 90.3%, depending on the choice of sample period (See Table 1.2).

1.2.2 **Stock Return Data**

Daily stock returns, trading volumes, prices, and shares outstanding of all NYSE listed common stocks from Dec 31st, 1925 to Dec 31st, 2011 are from CRSP. Using data from Zhang (2014), I extend daily stock returns back to Jan 1st, 1921 to capture the complete episodes of the 1920s economic boom.[2] My results do not depend on the inclusion of these earlier years' data; so, I use the full range of years in the tables and rerun my tests on CRSP data alone to check robustness.

---

[2] I am grateful to Blake Phillips for providing these data, which consist of daily stock prices, dividends, and ex-dividend dates as well as end-of-week shares outstanding, all hand-collected from the stock

1.2.3 **Macroeconomic Series**

I use growth rates (log differences) in real GDP (*GDP*), industrial production (*IP*), real fixed assets investment (*INVEST*), and unemployment (*UE*) to measure economic conditions. The *GDP* and *IP* series go back to 1921. The unemployment series starts in 1929, and the investment series starts in 1947. I then show results with the growth rate of GDP *(dlogGDP)* and industrial production (*dlogIP*) from 1921 to 2011 and results of all four variables from 1947 to 2011. Financial series[3] that have been shown to help predict future economic conditions by the literature are added in the regressions as control variables. Term spread (*TERM*) is the difference between the ten-year government bond yield and the three-month Treasury bill rate. Default spread (*DEF*) is calculated as the yield spread between the Moody's Baa corporate bonds and Aaa corporate bonds. Excess market index return (*MRET*) is the value-weighted NYSE market index return in excess of the three-month T-bill rate. Market volatility (*VOL*) is the average quarterly return variance of individual stocks calculated with daily returns. Dividend yield (*DIV*) is defined as the sum of the dividend payments of all NYSE firms in the prior four quarters, divided by the quarter-end NYSE index level (Campbell and Shiller, 1988). Following Amihud (2002), I define illiquidity (*ILLIQ*) for stock *j* in quarter *t* as

$$ILLIQ_{j,t} = \frac{1}{D_t} \sum_{s \in t} \frac{|R_{j,s}|}{VOL_{j,s}}, \quad [9]$$

where D_t is the number of trading days during quarter *t*, $R_{j,s}$ is the return of stock *j* on day *s*, and $VOL_{j,s}$ is the dollar trading volume of stock *j* on day *s*. The market-wide illiquidity is the

---

listings of *The New York Times* (1921-1925) and *Commercial and Financial Chronicle* (1921-1925). Following previous studies, I classify all stocks into 48 industries according to Fama and French (1997). Firms that do not fall into any category are classified as industry 49.

[3] See the appendix for the data source of the macroeconomic series.

average of the firm-level measure. Næs, Skjeltorp, and Ødegaar (2011) present evidence that this measure can help predict future economic growth.

1.2.4 **Data Summary**

Figure 1.1 graphs stock return idiosyncrasy measures and U.S. business cycles from 1921 to 2011. Both return idiosyncrasy measures tend to be high in expansions, start to drop as the economy peaks, and remain low during recessions. The 1960s and the 1990s are high idiosyncrasy decades, in which the U.S. economy was relatively stable and growing rapidly. The low point for idiosyncrasy is in the recession of 2008 to 2009.

Figure 1.2 shows that both the systematic and firm-specific components of stock return volatility move counter-cyclically. The most prominent volatility spikes appear in the two highest amplitude business cycles of the past century: the 1920s-1930s, and 1990s-2000s. The first decade of each episode contains an unusually large and sustained economic boom, driven by new technology; the second decade of each contains an unusually deep and prolonged downturn subsequent to a major financial crisis.

The high systematic volatility in these crisis and subsequent recessions is suggestive, though by no means conclusive evidence, of herding. In rational herding models, investors imitate other investors' trades to save on information costs (e.g. Bikhchandani¸ Hirshleifer, and Welch, 1998), resulting in asset price comovement. In noise trader models (e.g. De Long et al., 1990; Shleifer and Summers, 1990; Bennett, Sias, and Starks, 2003; Brand, Brav and Graham, 2010), large numbers of homogenously misinformed investors move the entire market, again elevating systematic risk. As noted above, high firm-specific return volatility can reflect investors distinguishing winning from losing firms (Caballero and Hammour, 1994), which increases returns heterogeneity as they do so (Chun et al., 2008).

Table 1.1 reports summary statistics of the idiosyncrasy measures, growth rates of the macroeconomic time series, and other financial variables over the full sample period of 1921-2011 and the post-war period, 1947-2011. The aggregate fixed assets investment, *INVEST*, and the unemployment rate, *UE*, become available in the latter. *DlogGDP, dlogIP, dlogINVEST, and dlogUE* denote the growth rates of the original series.

Augmented Dicky-Fuller (*ADF*) statistics provide unit-root test results. The numbers of lags in the *ADF* tests are determined by the Schwartz information criteria. *ADF* tests cannot reject the existence of a unit-root in the dividend yield and illiquidity series. I thus use log-differenced dividend yield (*DIV*) and illiquidity (*ILLIQ*) in all analyses. The idiosyncrasy measures pass the *ADF* tests in both sample periods. The macroeconomic series *GDP, IP, INVEST,* and *UE* are stationary after log difference transformation. The control variables *TERM, DEF,* and *MRET* are stationary in both sample periods as well. Calculating *DIV* and *ILLIQ* requires CRSP data, which are not available before 1926. These two series are included as controls only for the postwar (1947-2011) sample period, when the additional macroeconomic series *INVEST* and *UE* are also available.

Table 1.2 shows that the two idiosyncrasy measures are highly contemporaneously correlated with a correlation coefficient of 84.5% over 1921-2011 and 90.3% over 1947-2011. The idiosyncrasy measures are contemporaneously and significantly correlated with the four economic indicators in both sample periods. High stock return idiosyncrasy correlates with prosperous economic conditions, as measured by high growth rates in real GDP, industrial production, and investment, and a low growth rate in unemployment. Among the control variables, the contemporaneous correlation between default spread (*DEF*), market volatility (*VOL*), illiquidity (*ILLIQ*) and the macro-economic prosperity is negative. Higher market index returns (*MRET*) are contemporaneously positively correlated with favorable economic conditions.

1.3 Idiosyncrasy and Economic Conditions, In-Sample Analysis

This section examines whether or not stock return idiosyncrasy helps predict future economic prospects in in-sample regressions. I start with OLS regressions with one-quarter lag of idiosyncrasy; and then conduct vector autoregressions (*VAR*) where multiple lags are added to the regressions. Granger causality is used to test whether or not the lags of the idiosyncrasy measures are jointly significant in predicting future economic conditions.

1.3.1 **Predicting Economic Conditions with Stock Return Idiosyncrasy**

The left-hand side variable in each OLS regression is one of the economic conditions indicators in quarter *t*; and the right-hand side variable I am interested in is stock return idiosyncrasy in quarter *t-1*. I use growth rates of real GDP (*dlogGDP*) and industrial production (*dlogIP*) as indicators of economic conditions for the sample period 1921-2011, and supplement these with growth rates in real fixed investment (*dlogINVEST*) and the unemployment rate (*dlogUE*) for the post-war period 1947-2011. I control for financial variables shown elsewhere to contain leading information about economic growth: term spread (*TERM*), default spread (*DEF*), market index return (*MRET*), market volatility (*VOL*), dividend yield (*DIV*), and illiquidity (*ILLIQ*). The last two variables are only included for the sub-period of 1947-2011. Standard errors are adjusted by the Newey-West (1987) procedure.

Over the 1921-2011 period, high return idiosyncrasy predicts high growth rates in real GDP and industrial production in the subsequent quarter. After controlling for *TERM, DEF, MRET,* and *VOL*, the coefficient of *log(FIRM/MKT)* remains positive and significant. To gauge economic significance, I consider a one standard deviation change in the idiosyncrasy measures. After controlling for the financial variables, a one standard deviation increase in *log(FIRM/MKT)* predicts a 0.21 percentage points higher growth rate of real GDP in the subsequent quarter (roughly one fourth of the long-run mean growth rate of 0.8% per quarter) and 0.41 percentage points higher growth in industrial production (roughly half of the long-run mean of 0.9% per quarter). The corresponding increases are 0.12 and 0.24 percentage points following a one standard deviation increase in the other idiosyncrasy measure, *log(SSE/SSM)*. The coefficients of both idiosyncrasy measures remain positive and are highly significant in the regressions over the post-war period 1947-2011, and their economic significance is comparable to that in the full sample period. Both idiosyncrasy measures also predict a higher growth rate in fixed-assets investment and a lower growth rate in unemployment the next quarter in the post war period, although the coefficients are statistically insignificant in the regressions predicting *dlogUE* after controlling for the other financial variables. To estimate economic significance, I predict a 0.44 percentage points increase and a 0.45 percentage points drop in the growth rates of investment and unemployment, respectively, given a one standard deviation increase in *log(FIRM/MKT).* A similar exercise for *log(SSE/SSM)* predicts a 0.24 percentage points increase and 0.35 percentage points drop in the growth rates of investment and unemployment, respectively. These changes are economical significant: the mean growth rates in real investment and unemployment are 0.9% and 0.3%, respectively, during the sample period.

1.3.2 **Vector Autoregression (VAR) and Granger Causality Test**

The Newey-West regression results show that stock return idiosyncrasy in quarter *t-1* correlates with economic activity in quarter *t*. Granger causality using vector auto regressions (*VAR*) entails tests of the null hypothesis that, controlling for lags of the all other variables in the *VAR* system, lags of the variable of interest are jointly zero. In this case, the question of interest is whether or not lags of the idiosyncrasy measures are jointly significant in regressions predicting real economic series, as well as whether or not the lags of the economic series are jointly insignificant in predicting the idiosyncrasy measures. Granger causality test results from both bivariate *VAR* and multivariate *VAR* are explored.

Table 1.4 reports the coefficients of the idiosyncrasy measures from the vector autoregressions of the economic indicators. The bivariate VARs include one of the economic series and an idiosyncrasy measure only, while the multivariate VARs also include financial variables used in the Newey-West regressions[4]. I use the Akaike Information Criteria (AIC) to determine the number of lags of each variable in the VAR specifications. The selected number of lags varies from 3 to 6. VARs are standard OLS regressions without adjusting for auto-correlation in the error terms, and the significance levels of individual coefficients thus convey little information. However, the signs of the one quarter lags of the idiosyncrasy measures are

[4] The complete set of VAR regression coefficients and impulse response functions are available upon request.

always consistent with those in the Newey-West regressions. That is, higher stock return idiosyncrasy predicts favorable future economic conditions.

Table 1.5 reports the Granger causality test results from VARs for the economic series and return idiosyncrasy. Panel A shows Granger causality p-values from the bivariate VARs. The variable in each column denotes a right-hand side lagged variable and the variable in each row denotes the left-hand side variable being predicted. Significant p-values in the regressions of $dlogY_t$ on $Idio_{t-k,\ t-1}$ reject the null hypothesis that the lags of the idiosyncrasy measure from t-k to t-1 do not Granger cause *dlogY* at time *t*, where *Y=GDP* or *IP* over 1921-2011; and *Y=GDP*, *IP, INVEST*, or *UE* over 1947-2011. Significant p-values in the regressions of $Idio_t$ on $dlogY_{t-k,\ t-1}$ reject the null hypothesis that the lags of *dlogY* do not Granger cause the idiosyncrasy measure in quarter *t*. Smaller *p*-values indicate that the null hypothesis can be rejected with greater confidence.

The bivariate Granger causality test shows that lagged *log*(*FIRM*/*MKT)* Granger causes all four of the economic series in both sample period. In contrast, none of the economic series Granger cause *log(FIRM/MKT)* in either sample period. The bivariate Granger causality test results using *log*(*SSE*/*SSM*) show that *log*(*SSE*/*SSM*) does not Granger cause *dlogGDP* or *dlogIP* in the 1921-2011 period, but significantly Granger causes both in the post-war period. Two-way Granger causality is detected in the bivariate VAR between *log(SSE/SSM)* and d*logUE.* However, this two-way causality disappears after controlling for other financial variables in the multivariate VAR.

Panel B reports the Granger causality test results from multivariate VARs that also include term spread, default spread, excess market return, market volatility, dividend yield, and market illiquidity. In the full sample period, *Log(FIRM/MKT)* still marginally ($p = 0.10$) Granger causes the real GDP growth rate, while the other Granger causality test statistics become insignificant. In the post-war period (1947-2011), both idiosyncrasy measures strongly Granger cause *dlogGDP, dlogIP, and dlogINVEST*; while none of these economic series Granger cause return idiosyncrasy. Note that the ability of the idiosyncrasy measures to predict these economic variables is above and beyond Granger causation from the other potential leading indicators: the term structure, *TERM*; the default spread, *DEF*; the market risk premium, *MRET;* stock market volatility, *VOL;* dividend yield, *DIV;* and market illiquidity, *ILLIQ*. Only one idiosyncrasy measure, *dlog(SSE/SSM)*, marginally ($p = 0.06$) Granger causes *dlogUE,* while *dlogUE* does not Granger cause either idiosyncrasy measure.

Figure 1.3 plots the impulse response functions between the economic series and the idiosyncrasy measure from the multivariate VARs. The impulse response functions are calculated based on the general approach of Pesaran and Shin (1998), which does not require orthogonal shocks or the ordering of the variables in the VAR. The graphs show how the shocks to stock return idiosyncrasy affect economic conditions in the subsequent three years. The ability of stock return idiosyncrasy to predict future economic conditions is evident in the impulse response functions. In the post-war period, a shock to stock return idiosyncrasy strongly flows through to the economic series in the next four to six quarters; while a shock to the economic series has only modest impact on idiosyncrasy.

Overall, both the regression and Granger causality test results are consistent with stock return idiosyncrasy containing useful information about future economic prospects one-quarter-ahead. This is especially evident in the post-war period. More idiosyncratic stock returns predict greater economic prosperity; less idiosyncrasy (more comovement) in stock returns predicts weaker economic conditions.

1.3.3 **Robustness Checks**

This section tests the robustness the results presented above by rerunning the tests in a variety of different ways. Where identical patterns of signs and significance are found, I say that the robustness check generates results "consistent with" those in the corresponding table. Where the robustness check produces results not consistent with the tables, I explain what is different.

The results in the tables might be evident only in NYSE firms. However, consistent results arise from rerunning the Newey-West regressions and Granger causality tests using the combined NYSE/NASDAQ/AMEX sample.

To test whether the results are driven by the coinciding cyclical movement of the idiosyncrasy measures and the macroeconomic series, I use a range of methods to remove the cyclical component from the idiosyncrasy measures: subtracting the moving average of the past two years, applying a Hodrick-Prescott filter, and taking residuals from ARMA (1, 1) models[5]. These exercises generate quantitatively similar results with those in the Newey-West regressions, though some are less statistically significant.

I use hand-collected data for the early 1920s and CRSP data for 1926 to 2011. To ensure that the results do not depend on the hand collected data, I repeat all my tests using CRSP data alone. This produces results consistent with the tables.

[5] ARMA (1, 1) model produces the lowest AIC for both idiosyncrasy measures among all AR, MA, and ARMA models.

1.4 Predicting Future Economic Conditions Out-of-Sample

The in-sample analysis above shows that the degree of idiosyncrasy in stock returns can help predict the movement of four important macroeconomic indicators the subsequent quarter, particularly in the post-war period. All the tests above are "in-sample", in the sense that each uses all available data. "Out-of-sample" analysis is considered more robust to selection bias and the over-fitting problem (Fair and Shiller, 1990; Estrella and Mishkin, 1998). This section tests the predictive power of stock return idiosyncrasy out-of-sample. The out-of-sample method means that the prediction is made using data other than those over which the model is estimated.

I apply a recursive estimation scheme to the in-sample regression and then predict the economic series at quarterly frequency for up to a year following the estimation window. The recursive approach lets the estimation window expands one period at a time as the predicted point moves one period forward. I then investigate whether including stock return idiosyncrasy can again improve the accuracy with which the model forecasts future macroeconomic conditions.

1.4.1 **Out-of-Sample Test Methodology**

I start by estimating two sets of nested regressions where the restricted regression sets the coefficient of the idiosyncrasy series (γ_t) equal to zero.

[10] $$dlogY_{t+h} = \alpha_t + \beta_t FIN\_VAR_t + \gamma_t Idio_t + dlogY_t + \epsilon_{t+h},$$

where *DlogY* denotes the predicted economic series, *h* denotes the forecast horizon, *FIN_VAR* represents one or all of the financial variables, and *Idio* is the idiosyncrasy measure. Eq. [10] is estimated recursively starting with the first half of the sample period as the estimation window − that is, 1921Q1-1966Q4 for the full sample period and 1947Q1-1979Q4 for the post-war period.

I use the *MSE-F* and *ENC-NEW* tests to assess the out-of-sample forecasting performance of the idiosyncrasy measures. The *MSE-F* test is based on the equal mean standard error (*MSE*) as defined by McCracken (2007):

$$[11] \qquad MSE - F = (p - h + 1)\frac{MSE_r - MSE_u}{MSE_u}$$

where p is the number of out-of-sample forecasts and h is the forecasting horizon. The subscripts r and u denote the restricted and the unrestricted predictive models, respectively. The null hypothesis is that the *MSE* of the restricted model is no greater than that of the unrestricted model. Significant one-tail *MSE-F* statistics suggest that the unrestricted model has better forecasting accuracy than the restricted model.

The *ENC-NEW* statistic, in contrast, asks if forecasts of the restricted model (economic and financial variables only) "encompass" the forecasts made by the unrestricted model (idiosyncrasy measure also included). If the unrestricted model is not encompassed by the restricted model, the idiosyncrasy measure contains useful information in predicting future economic conditions out-of-sample. Clark and McCracken (2001) show that the *ENC-NEW* test has higher power than the *MSE-F* test. The *ENC-NEW* statistic by Clark and McCracken (2001) is defined as

$$[12] \qquad ENC - NEW = (p - h + 1)\frac{\sum_{t+h}(\epsilon_{r,t+h}^2 - \epsilon_{r,t+h}\epsilon_{u,t+h})}{\sum_{t+h}\epsilon_{u,t+h}^2}$$

The *MSE-F* and *ENC-NEW* test results, with forecast horizons of one, two, and four quarters, are shown in Table 1.6. The significant test statistics suggest that including the idiosyncrasy measures in the forecasting regressions significantly improves forecasting accuracy out-of-sample.

1.4.2 **Forecasting Performance of Stock Return Idiosyncrasy**

In the one quarter ahead forecasts, the *ENC-NEW* test shows that both idiosyncrasy measures, *log*(*FIRM/MKT)* and *log(SSE/SSM),* significantly improves the ability of the regressions to forecast future conditions in both the 1921-2011 and 1947-2011 periods. The *MSE-F* test statistics are less significant. However, in the regressions controlling for all of the financial variables, the *MSE-F* statistics show the unrestricted model with *log(SSE/SSM)* having a significantly lower mean standard error.

If the forecast horizon increases from one to two quarters, the forecasting performance of the idiosyncrasy measures becomes weaker. However, the *MSE-F* and *ENC-NEW* tests show that both idiosyncrasy measure significantly improve the forecasting performance of the predictive models of *dlogGDP, dlogIP*, and *dlogINVEST* over 1947-2011. The *ENC-NEW* statistics are also significant across most of the predictive models of *dlogUE*, particularly in the regressions with *log(SSE/SSM)* included on the right-hand side. If the forecasting horizon is lengthened to four quarters, *log(FIRM/MKT)* retains power to predict *dlogGDP* and *dlogIP* in both the full and postwar time windows, while *log(SSE/SSM)* remains strongly predictive of *dlogGDP, dlogIP, and dlogINVEST*. Neither idiosyncrasy measure improves the forecasting accuracy of *dlogUE* out-of-sample.

In some cases, the *MSE-F* test and the *ENC-NEW* test suggest inconsistent results, meaning that only one of the two test statistics is significant. Clark and McCracken (2001) prove that in such situations, the *ENC-NEW* test has higher power and better size properties in other out-of-sample tests commonly used in the literature. Rogoff and Stavrakeva (2008) further show that the power of the *ENC-NEW* test dominates that of the *MSE-F* test only if there is no severe forecasting bias. They suggest that a forecast has severe scale bias if the regression coefficient of $Y_{t+h} = \mu \hat{Y}_{u,t+h}$ lies between 0 and 0.5[6*]. In unreported results, I calculate μ for the forecasts where *MSE-F* and *ENC-NEW* tests show inconsistent results. Only the four-quarter ahead predictive regression of *dlogGDP* on *log(FIRM/MKT)* and the financial variables gives $\mu < 0.5$. No severe bias is detected in the result of the other forecasting regressions.

1.5 Conclusions

Stock return idiosyncrasy, as measured by the ratio of the firm-specific to systematic components of stock return variation, contains useful information for predicting future economic conditions, particularly in the post-war era. Greater idiosyncrasy in stock returns predicts better economic prospects. In-sample Newey-West regressions and Granger causality tests show that lags of stock return idiosyncrasy help predict changes in real GDP, industrial production, fixed assets investment, and unemployment. Out-of-sample tests show idiosyncrasy also significantly improving the forecasting accuracy of simple models predicting these four macroeconomic variables as far as four quarters ahead.

---

[*] There is no constant in the regression. H denotes the forecasting horizon.

These findings provide new empirical support for Veldkamp's (2005, 2006) thesis that information production specialists provide less firm-specific information when the economy is weaker. The findings presented are new, in that this shift towards producing less firm-specific information leads actual economic downturns. This is reasonable, in that Veldkamp's information production specialists, by virtue of that specialization, might be uniquely informed about impending economic conditions and adjust their information offerings in advance.

The findings presented above may also be of interest to financial historians, in that my findings suggest possible, if still very speculative, underpinnings in the microeconomics of information production for Kindleberger's financial mania, panic and crash cycles. Perhaps the early phases of Kindleberger's cycle, in which some firms profit from new technologies or other new developments, correspond to a period of prosperity, in which information production specialists opt to generate firm-specific information unusually energetically. Demand for firm-specific information might reflect investors' desire to distinguish firms that are exploiting the associated profit opportunities successfully from those that are not. When these profit opportunities are nearly all exploited, information specialists realize this before others, and refocus their efforts on generating more economy-level information. The mania and panic phases of Kindleberger's cycle might then ensue as investors enthusiastic about the new profit opportunities increasingly trade on the same market-wide information and, only somewhat after the information specialists, comprehend that that, but those opportunities have been exploited and the economy is slowing. This set of connections, though consistent with the above findings and with Veldkamp's models of the microeconomics of information production, is obviously highly speculative. I hope to pursue this line of inquiry in future work.

More concrete policy implications of my findings follow from the time-series of stock return idiosyncrasy helping to predict subsequent macroeconomic conditions. First, stock market idiosyncrasy should not be overlooked in the economic forecasting models of governments, central banks, financial institutions, and nonfinancial corporations. A death of idiosyncrasy cautions that the economy could be overheated and that information production specialists perceive diminished opportunities for individual firms to capture abnormal profits. Portfolio managers may find these results useful as well. The unsystematic volatility component should not be omitted in the models that predict future macroeconomic risk.

1.6 Appendix: Data Sources of the Macroeconomic Series

Nominal GDP

- 1921-1946: quarterly Gross National Product (GNP); the Appendix B of "The American Business Cycle: Continuity and Change" Edited by Robert J. Gordon. National Bureau of Economic Research Studies in Business Cycles Volume 25, University of Chicago Press 1986[7].
- 1947-2011: quarterly Gross Domestic Product (GDP), quantity index; The Bureau of Economic Analysis

Industrial Production (IP)

- 1921-2010: monthly industrial production index, FRED economic data, Federal Reserve Bank of St. Louis. The log difference of IP index $dlog(IP)_t$ is calculated with the index levels at the end of quarter t and quarter $t-1$.

Real Fixed Investment

- 1947-2011: quarterly real private fixed investment quantity index, quantity index; Bureau of Economic Analysis

Unemployment rate

- 1947: U.S. monthly unemployment Rate 01/1947-12/1966, NBER macro-history database
- 1948-2011: monthly civilian unemployment rate; FRED economic data, Federal Reserve at St. Louis. The log difference of unemployment rate $dlog(UE)_t$ is calculated with the monthly unemployment rate at the end of quarter t and quarter $t-1$.

Three-Month T-bill Rate

- 1921-1933: U.S. yields on short-term United States securities, three month treasury, NBER macro-history database

[7] Can be downloaded at http://www.nber.org/data/abc/

- 1934-2011: three-month Treasury bill: secondary market rate; FRED economic data, Federal Reserve at St. Louis.
- The T-bill rate in each quarter is defined as the average of the monthly T-bill rate within a quarter.

Ten-year Government Bond Yield

- 1921-1924: U.S. yields on short-term United States securities, NBER macro-history database
- 1925-Mar 1953: Interest rates: Treasury Constant Maturities; 10-year; Federal Reserve Data, Wharton Research Data Services
- Apr 1953- 2011: Interest rates: Treasury composite; over 10 years; Federal Reserve Data, Wharton Research Data Services
- The quarterly rate is defined as the average of the monthly rate within each quarter.

Baa and Aaa Corporate Bond Yield

- 1921-2011: monthly interest rate data; CRSP Federal Reserve Bank Database
- The quarterly rate is defined as the average of the monthly rate within each quarter.

Table 1.1 Summary Statistics

Table 1.1 lists the summary statistics of the stock return idiosyncrasy, macroeconomic series, and other financial variables used in the analysis over the 1921-2011 and 1947-2011 period. The idiosyncrasy measures *log(SSE/SSM)* and *log(FIRM/MKT)* are the ratio of the firm-specific to systematic return variation following the methods of Morck et al. (2000) and Campbell et al. (2001), respectively. The quarterly real economic series *dlogGDP* and *dlogIP* are available during 1921-2011; and *dlogINVEST* and *dlogUE* are added in for the 1947-2011 period. *dlogGDP, dlogIP, dlogINVEST,* and *dlogUE* are the log difference of real GDP, industrial production, fixed assets investment, and unemployment rate from the previous quarter to the current one. Term spread (*TERM)* is defined as the difference between the 10-year government bond yield and the three-month T-bill rate. Default spread (*DEF)* is the difference between the Moody's Baa and Aaa corporate bond yields. *MRET* is the value-weighted NYSE index return in excess of the three-month T-bill rate. Market volatility (*VOL*) is the average quarterly return variance of all the NYSE stocks calculated with daily return data. Dividend yield (*DIV*) is the total dividend of all NYSE stocks over the past four quarters divided by the quarter-end index level. Illiquidity measure (*ILLIQ)* is the Amihud (2002) illiquidity measure averaged across all NYSE stocks. The log difference of *DIV* and *ILLIQ* are used in the regressions because the original series do not pass the Augmented Dicky-Fuller (ADF) test. The ADF test allows a drift and a time trend in the regression equations, and the number of lags used in the test is chosen by the Schwarz information criteria. "***", "**", and "*" denote the significance at 1%, 5%, and 10% level, respectively. All variables are winsorized at 1% level.

	Obs.	*Mean*	*Std. Dev.*	*Median*	*Min*	*Max*	*ADF Stats.*
1921-2011							
log(SSE/SSM)	364	1.953	0.686	2.028	-0.131	3.403	-5.336***
log(FIRM/MKT)	364	1.044	0.609	1.081	-0.670	2.499	-5.042***
dlog(GDP)	363	0.008	0.020	0.008	-0.058	0.0748	-7.102***
dlog(IP)	363	0.009	0.036	0.009	-0.128	0.121	-8.186***
TERM	364	0.015	0.011	0.015	-0.012	0.036	-5.136***
DEF	364	0.012	0.007	0.009	0.004	0.039	-3.750**
MRET	364	0.010	0.095	0.016	-0.278	0.292	-9.527***
VOL	364	0.001	0.001	0.000	0.000	0.006	-3.544**

	Obs.	*Mean*	*Std. Dev.*	*Median*	*Min*	*Max*	*ADF Stats.*
1947-2011							
log(SSE/SSM)	260	1.928	0.725	1.990	-0.131	3.403	-4.016***
log(FIRM/MKT)	260	1.119	0.591	1.128	-0.670	2.499	-3.541**
dlog(GDP)	260	0.008	0.010	0.008	-0.027	0.040	-8.744***
dlog(IP)	260	0.008	0.020	0.009	-0.068	0.081	-10.073***
dlog(INVEST)	259	0.009	0.026	0.010	-0.073	0.073	-7.787***
dlog(UE)	260	0.003	0.075	-0.007	-0.247	0.353	-7.825***
TERM	260	0.014	0.011	0.013	-0.012	0.036	-5.094***
DEF	260	0.009	0.004	0.008	0.004	0.030	-4.639***
MRET	260	0.001	0.085	0.008	-0.278	0.292	-10.285***
VOL	260	0.000	0.000	0.000	0.000	0.004	-7.076***
DIV	260	-0.015	0.083	-0.020	-0.223	0.309	-7.875***
ILLIQ	260	-0.019	0.245	-0.023	-0.827	0.649	-11.552***

Table 1.2 Correlation Matrix

Table 1.2 shows the pair-wise correlation coefficients between all the variables used in the regression analysis over the 1921-2011 and 1947-2011 period. The idiosyncrasy measures *log(SSE/SSM)* and *log(FIRM/MKT)* are the ratio of the firm-specific to systematic return variation following the methods of Morck et al. (2000) and Campbell et al. (2001), respectively. The economic indicators are the growth rate of real GDP, *dlog(GDP)*, and industrial production, *dlog(IP),* during 1921-2011. The growth rate of real fixed-investment, *dlogINVEST,* and unemployment rate, *dlogUE*, are added in for the 1947-2011 period. Term spread (*TERM)* is defined as the difference between the 10-year government bond yield and the three-month T-bill rate. Default spread (*DEF)* is the difference between the Moody's Baa and Aaa corporate bond yields. *MRET* is the value-weighted NYSE index return in excess of the three-month T-bill rate. Market volatility (*VOL*) is the average quarterly return variance of all the NYSE stocks calculated with daily returns. Dividend yield (*DIV*) is the total cash dividends of all NYSE stocks over the past four quarters divided by the quarter-end index level. Illiquidity measure (*ILLIQ)* is the Amihud (2002) illiquidity measure averaged across all NYSE stocks. The log difference of *DIV* and *ILLIQ* are used in the regressions because the original series do not pass the Augmented Dicky-Fuller (ADF) test. "***", "**", and "*" denote the significance at 1%, 5%, and 10% level, respectively. All variables are winsorized at 1% level.

1921-2011

	log(SSE/SSM)	*log(FIRM/MKT)*	*dlog(GDP)*	*dlog(IP)*	*TERM*	*DEF*	*MRET*	*VOL*
log(SSE/SSM)	1.000							
log(FIRM/MKT)	0.845***	1.000						
	(0.00)							
dlog(GDP)	0.160***	0.206***	1.000					
	(0.00)	(0.00)						
dlog(IP)	0.097*	0.153***	0.844***	1.000				
	(0.07)	(0.00)	(0.00)					
TERM	-0.223***	-0.199***	0.030	0.004	1.000			
	(0.00)	(0.00)	(0.56)	(0.93)				
DEF	-0.202***	-0.387***	-0.132**	-0.145**	0.311***	1.000		
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)			
MRET	0.194***	0.238***	0.236***	0.180***	0.154***	0.002	1.000	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.96)		
VOL	-0.227***	-0.431***	-0.223***	-0.166***	0.210***	0.820**	-0.085	1.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)	

1947-2011	*log(SSE/SSM)*	*log(FIRM/MKT)*	*dlog(GDP)*	*dlog(IP)*	*dlog(INVEST)*	*dlog(UE)*	*TERM*	*DEF*	*MRET*	*VOL*	*DIV*	*ILLIQ*
log(SSE/SSM)	1.000											
log(FIRM/MKT)	0.903***	1.000										
	(0.00)											
dlog(GDP)	0.225***	0.220***	1.000									
	(0.00)	(0.00)										
dlog(IP)	0.145**	0.155**	0.802***	1.000								
	(0.02)	(0.01)	(0.00)									
dlog(INVEST)	0.233***	0.269***	0.695***	0.723***	1.000							
	(0.00)	(0.00)	(0.00)	(0.00)								
dlog(UE)	-0.099	-0.137**	-0.631***	-0.740***	-0.588***	1.000						
	(0.11)	(0.03)	(0.00)	(0.00)	(0.00)							
TERM	-0.195***	-0.125**	0.035	-0.005	0.099	-0.006	1.000					
	(0.00)	(0.04)	(0.57)	(0.94)	(0.11)	(0.93)						
DEF	-0.374***	-0.312***	-0.313***	-0.383***	-0.335***	0.297***	0.318***	1.000				
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					
MRET	0.167**	0.193**	0.103*	0.018	0.055	-0.016	0.214***	-0.046	1.000			
	(0.01)	(0.00)	(0.10)	(0.77)	(0.38)	(0.79)	(0.00)	(0.46)				
VOL	-0.501***	-0.432***	-0.327***	-0.343***	-0.378***	0.278***	0.160**	0.563***	-0.190***	1.000		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)			
DIV	-0.318***	-0.344	-0.085	-0.012	-0.078	-0.002	-0.131**	-0.087	-0.884***	0.248***	1.000	
	(0.00)	(0.00) ***	(0.17)	(0.85)	(0.21)	(0.97)	(0.04)	(0.16)	(0.00)	(0.00)		
ILLIQ	-0.260***	-0.332***	-0.185***	-0.114*	-0.139**	0.087	-0.215***	-0.023	-0.539***	0.241***	0.535***	1.000
	(0.00)	(0.00)	(0.00)	(0.07)	(0.03)	(0.16)	(0.00)	(0.71)	(0.00)	(0.00)	(0.00)	

Table 1.3 Stock Return Idiosyncrasy and the Business Cycle – Newey-West Regressions

This table reports the Newey-West regression results of economic condition variable *dlogGDP, dlogIP, dlogINVEST*, and *dlogUE* in quarter *t+1* on the stock return idiosyncrasy in quarter *t*. The idiosyncrasy measures *log(SSE/SSM)* and *log(FIRM/MKT)* are the log ratio of the firm-specific to systematic return variation following the methods of Morck et al. (2000) and Campbell et al. (2001), respectively. *DlogGDP, dlogIP, dlogINVEST,* and *dlogUE* are the log difference of real GDP, industrial production, real fixed-assets investment, and unemployment rate from the previous quarter to the current one. Term spread (*TERM)* is defined as the difference between the 10-year government bond yield and the three-month T-bill rate. Default spread (*DEF)* is the difference between the Moody's Baa and Aaa corporate bond yields. *MRET* is the value-weighted NYSE index return in excess of the three-month T-bill rate. Market volatility (*VOL*) is the average quarterly return variance of all the NYSE stocks calculated with daily return data. Dividend yield (*DIV*) is the total dividend of all NYSE stocks over the past four quarters divided by the quarter-end index level. Illiquidity measure (*ILLIQ)* is the Amihud (2002) illiquidity measure averaged across all NYSE stocks. The log difference of *DIV* and *ILLIQ* are used in the regressions because the original series do not pass the Augmented Dicky-Fuller (ADF) test. Panels A and B shows the regression results during 1921Q1-2011Q4 and 1947Q1-2011Q4, respectively. The Newey-West adjusted t-statistics are reported in the parentheses. The maximum order of the lagged autocorrelation is chosen by Newey and West's (1994) automatic bandwidth-selection procedure. "***", "**", and "*" denote the significance at 1%, 5%, and 10% level, respectively.

Panel A 1921-2011

1921Q1-2011Q4	*Yt=dlogGDP*				*Yt=dlogIP*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(FIRM/MKT)	0.005**		0.003**		0.008**		0.006**	
	(2.32)		(2.40)		(2.44)		(2.10)	
log(SSE/SSM)		0.003***		0.002		0.006***		0.004*
		(2.79)		(1.47)		(3.37)		(1.85)
TERM			0.073	0.075			0.114	0.123
			(1.19)	(1.13)			(0.98)	(0.90)
DEF			0.511*	0.482*			0.582	0.534
			(1.71)	(1.81)			(1.37)	(1.15)
MRET			0.053**	0.055***			0.094***	0.097***
			(2.53)	(2.95)			(2.81)	(3.46)
VOL			-3.358*	-3.708**			-1.870	-2.482
			(-1.70)	(-2.05)			(-0.82)	(-0.93)
dlogY$_{t-1}$	0.294***	0.305***	0.231**	0.232**	0.405***	0.413***	0.373***	0.375***
	(3.58)	(2.81)	(2.16)	(2.18)	(3.53)	(3.96)	(2.80)	(2.97)
Adj_R^2	0.116	0.110	0.184	0.182	0.191	0.188	0.252	0.251

Panel B 1947-2011

1947Q1-2011Q4	*Yt=dlogGDP*				*Yt=dlogIP*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(FIRM/MKT)	0.004***		0.003**		0.008***		0.004**	
	(3.42)		(1.98)		(3.81)		(2.02)	
log(SSE/SSM)		0.003***		0.003***		0.006***		0.004***
		(4.66)		(3.26)		(4.13)		(2.69)
TERM			0.129**	0.137**			0.214**	0.225***
			(2.02)	(2.20)			(2.47)	(2.75)
DEF			-0.208	-0.170			-0.201	-0.125
			(-1.14)	(-0.91)			(-0.72)	(-0.43)
MRET			0.014	0.017			0.015	0.021
			(0.98)	(1.27)			(0.69)	(0.86)
VOL			-1.914	-1.283			-4.827	-3.784
			(-0.96)	(-0.65)			(-1.46)	(-1.25)
DIV			0.004	0.008			-0.009	-0.002
			(0.28)	(0.65)			(-0.42)	(-0.08)
ILLIQ			-0.004*	-0.005*			-0.011	-0.012**
			(-1.90)	(-1.90)			(-1.64)	(-2.34)
$dlogY_{t-1}$	0.270***	0.266***	0.205***	0.204**	0.489***	0.490***	0.444***	0.452***
	(4.20)	(4.07)	(2.61)	(2.52)	(7.76)	(7.30)	(6.63)	(5.86)
Adj_R^2	0.150	0.157	0.199	0.209	0.318	0.321	0.375	0.381

1947Q1-2011Q4	*Yt=dlogINVEST*				*Yt=dlogUE*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(FIRM/MKT)	0.010***		0.006**		-0.016**		-0.006	
	(3.67)		(2.31)		(-2.48)		(-0.64)	
log(SSE/SSM)		0.007***		0.004**		-0.013***		-0.006
		(4.05)		(2.33)		(-2.70)		(-0.76)
TERM			0.423***	0.430***			-0.904***	-0.919***
			(3.26)	(3.37)			(-3.75)	(-3.86)
DEF			-0.187	-0.210			0.796	0.679
			(-0.55)	(-0.55)			(0.57)	(0.45)
MRET			0.018	0.014			-0.051	-0.061
			(0.43)	(0.27)			(-0.44)	(-0.48)
VOL			-8.552*	-7.696*			19.349	17.771
			(-1.84)	(-1.66)			(1.49)	(1.35)
DIV			-0.005	-0.010			0.022	0.011
			(-0.14)	(-0.21)			(0.23)	(0.10)
ILLIQ			-0.016**	-0.017**			0.024	0.024
			(-2.45)	(-2.48)			(0.96)	(0.97)
$dlogY_{t-1}$	0.465***	0.480***	0.392***	0.401***	0.490***	0.495***	0.454***	0.458***
	(8.74)	(9.86)	(6.87)	(7.12)	(5.64)	(5.70)	(4.23)	(4.10)
Adj_R^2	0.319	0.308	0.398	0.395	0.269	0.269	0.305	0.306

Table 1.4 Predicting Business Cycles with Stock Return Idiosyncrasy, Vector Autoregression

This table reports the selected coefficients of the lagged stock return idiosyncrasy from the vector autoregressions. Only the results using one of the four economic condition indicators as left-hand side variable are reported here. These economic indicators are the growth rate of real GDP (*dlogGDP*), industrial production (*dlogIP*), real fixed investment (*dlogINVEST*), and unemployment rate (*dlogUE*). *Log(FIRM/MKT)* and *log(SSE/SSM)* denote the measures of stock return idiosyncrasy following the method by Campbell et al. (2001) and Morck et al. (200), respectively. The Bivariate VARs in Panel A only include one economic indicator and one idiosyncrasy measure. The multivariate VARs in Panel B also include the following financial variables (unreported): term spread (*TERM*), default spread (*DEF*), excess market return (*MRET*), market volatility (*VOL*), dividend yield (*DIV*), and illiquidity (*ILLIQ*). The last two variables are only included in the regressions for the period of 1947-2011 when *dlogINVEST* and *dlogUE* become available. The number of lags in each VAR is selected by the AIC. "***", "**", and "*" denote the significance at 1%, 5%, and 10% level, respectively.

Panel A Bivariate VAR

1921-2011	*Dlog(GDP)*	*Dlog(IP)*		*Dlog(GDP)*	*Dlog(IP)*
Lag of log(FIRM/MKT)			*Lag of log(SSE/SSM)*		
1	0.007***	0.011***	1	0.005***	0.009***
	(3.44)	(3.50)		(2.76)	(2.86)
2	-0.004*	-0.004	2	-0.003	-0.002
	(-1.94)	(-1.26)		(-1.37)	(-0.58)
3	-0.000	-0.000	3	0.000	-0.001
	(-0.18)	(-0.13)		(0.11)	(-0.39)
4	0.001	-0.000	4	0.002	-0.002
	(0.57)	(-0.03)		(0.79)	(-0.51)
5	-0.001	-0.001	5	-0.002	-0.001
	(-0.69)	(-0.29)		(-0.76)	(-0.29)
6	-0.002	-0.004	6	-0.002	
	(-1.02)	(-1.15)		(-1.27)	
Adj. R^2	0.21	0.20		0.33	0.33

1947-2011	Dlog(GDP)	Dlog(IP)	Dlog(INVEST)	Dlog(UE)		Dlog(GDP)	Dlog(IP)	Dlog(INVEST)	Dlog(UE)
lag of log(FIRM/MKT)					lag of log(SSE/SSM)				
1	0.005***	0.009***	0.013***	-0.019**	1	0.004***	0.007***	0.008***	-0.014*
	(4.14)	(4.33)	(4.53)	(-2.23)		(4.32)	(4.07)	(3.38)	(-1.97)
2	-0.000	-0.000	0.004	-0.005	2	0.000	0.001	0.006**	-0.011
	(-0.02)	(-0.19)	(1.38)	(-0.56)		(0.08)	(0.32)	(2.13)	(-1.34)
3	-0.003*	0.000	-0.005	-0.001	3	-0.002	-0.001	-0.004*	0.002
	(-1.94)	(0.01)	(-1.52)	(-0.11)		(-1.38)	(-0.53)	(-1.65)	(0.30)
4	0.001	-0.002	-0.004	0.011	4	0.001	-0.001	-0.003	0.010
	(0.83)	(-1.13)	(-1.50)	(1.30)		(0.97)	(-0.40)	(-1.29)	(1.43)
5					5	-0.000			
						(-0.18)			
Adj. R^2	0.21	0.4	0.35	0.29		0.21	0.41	0.34	0.29

Panel B Multivariate VAR

1921-2011	Dlog(GDP)	Dlog(IP)		Dlog(GDP)	Dlog(IP)
lag of log(FIRM/MKT)			lag of log(SSE/SSM)		
1	0.002	0.006	1	0.001	0.004
	(1.14)	(1.57)		(0.42)	(1.14)
2	-0.003	-0.003	2	-0.002	-0.001
	(-1.13)	(-0.82)		(-0.92)	(-0.20)
3	-0.000	0.001	3	0.001	0.000
	(-0.02)	(0.35)		(0.27)	(0.08)
4	0.005**	0.003	4	0.004**	0.002
	(2.07)	(0.70)		(2.08)	(0.55)
Adj. R^2	0.31	0.41		0.31	0.41

1947-2011	Dlog(GDP)	Dlog(IP)	Dlog(INVEST)	Dlog(UE)		Dlog(GDP)	Dlog(IP)	Dlog(INVEST)	Dlog(UE)
lag of log(FIRM/MKT)					lag of log(SSE/SSM)				
1	0.004**	0.005**	0.004	-0.003	1	0.003***	0.004**	0.001	-0.001
	(2.59)	(2.18)	(1.20)	(-0.26)		(2.83)	(2.13)	(0.38)	(-0.08)
2	-0.001	-0.002	0.009**	-0.009	2	-0.000	-0.000	0.010***	-0.016*
	(-0.52)	(-0.73)	(2.48)	(-0.76)		(-0.27)	(-0.02)	(3.30)	(-1.76)
3	-0.000	0.005*	-0.001	-0.014	3	0.000	0.002	-0.002	-0.006
	(-0.26)	(1.92)	(-0.20)	(-1.29)		(0.03)	(0.96)	(-0.83)	(-0.67)
4	0.003**	-0.001	-0.002	0.005	4	0.003***	0.002		
	(2.05)	(-0.31)	(-0.62)	(0.51)		(2.64)	(0.83)		
Adj. R^2	0.27	0.47	0.43	0.33		0.29	0.47	0.44	0.35

Table 1.5 Stock Return Idiosyncrasy and the Business Cycle – Granger Causality Tests

This table reports selected Granger causality test p-values from the vector autoregressions (VAR). The bivariate VARs in Panel A only include one of the economic series (*dlogY)*, and one of the idiosyncrasy measure, *dlog(FIRM/MKT)* or *dlog(SSE/SSM)*. Each cell represents a regression with the variable in the column as the left-hand side variable and the variable in the row as right-hand side variable. Significant p-values reject the hypothesis that the *k* lags of the variable in the column are jointly zero in the regressions predicting the variable in the row during quarter *t*. The multivariate VARs in Panel B also include the financial variables term spread (*TERM*), default spread (*DEF*), excess market return (*MRET*), market volatility (*VOL*), dividend yield (*DIV*), and illiquidity (*ILLIQ*). Each row represents a regression with the variable in the row heading as the left-hand side variable and the variables in the column headings as the right-hand side variables. The number of lags in each VAR is chosen by AIC. "***", "**", and "*" denote the significance at 1%, 5%, and 10% level, respectively.

Panel A Bivariate VAR

1947Q1-2011Q4	$DlogY_{t-k,\,t-1}$	$Idio_{t-k,\,t-1}$	$DlogY_{t-k,\,t-1}$	$Idio_{t-k,\,t-1}$
Y = GDP, idio =log(FIRM/MKT)			*Y=GDP, idio = log(SSE/SSM)*	
$DlogY_t$		0.026**		0.104
$Idio_t$	0.132		0.131	
Y = IP, idio =log(FIRM/MKT)			*Y=IP, idio = log(SSE/SSM)*	
$DlogY_t$		0.039**		0.137
$Idio_t$	0.467		0.126	

1947Q1-2011Q4	$DlogY_{t-k,\,t-1}$	$Idio_{t-k,\,t-1}$	$DlogY_{t-k,\,t-1}$	$Idio_{t-k,\,t-1}$
Y = GDP, idio =log(FIRM/MKT)			*Y=GDP, idio = log(SSE/SSM)*	
$DlogY_t$		0.026**		0.104
$Idio_t$	0.132		0.131	
Y = IP, idio =log(FIRM/MKT)			*Y=IP, idio = log(SSE/SSM)*	
$DlogY_t$		0.039**		0.137
$Idio_t$	0.467		0.126	
Y = INVEST, idio =log(FIRM/MKT)			*Y=INVEST, idio = log(SSE/SSM)*	
$DlogY_t$		0.000***		0.000***
$Idio_t$	0.761		0.588	
Y = UE, idio =log(FIRM/MKT)			*Y=UE, idio = log(SSE/SSM)*	
$DlogY_t$		0.072*		0.030**
$Idio_t$	0.234		0.015**	

Panel B Multivariate VAR

1921Q1-2011Q4	*DlogY*$_{t-k, t-1}$	*Idio*$_{t-k, t-1}$	*TERM*$_{t-k, t-1}$	*DEF*$_{t-k, t-1}$	*MRET*$_{t-k, t-1}$	*VOL*$_{t-k, t-1}$
Y = GDP, idio =log(FIRM/MKT)						
DlogY$_t$		0.100*	0.185	0.000***	0.000***	0.007***
Idio$_t$	0.154		0.589	0.7	0.005***	0.941
Y=GDP, idio = log(SSE/SSM)						
DlogY$_t$		0.138	0.143	0.000***	0.000***	0.004***
Idio$_t$	0.134		0.751	0.802	0.016**	0.925
Y = IP, idio =log(FIRM/MKT)						
DlogY$_t$		0.308	0.154	0.003***	0.000***	0.002***
Idio$_t$	0.744		0.629	0.92	0.014**	0.584
Y = IP, Idio = log(SSE/SSM)						
DlogY$_t$		0.481	0.141	0.003***	0.000***	0.002***
Idio$_t$	0.197		0.766	0.925	0.034**	0.879

1947Q1-2011Q4	*DlogY*$_{t-k, t-1}$	*Idio*$_{t-k, t-1}$	*TERM*$_{t-k, t-1}$	*DEF*$_{t-k, t-1}$	*MRET*$_{t-k, t-1}$	*VOL*$_{t-k, t-1}$	*DIV*$_{t-k}$	*ILLIQ*$_{t-k}$
Y = GDP, idio =log(FIRM/MKT)								
DlogGDP$_t$		0.006***	0.398	0.137	0.103	0.359	0.83	0.723
Idio$_t$	0.631		0.199	0.144	0.001***	0.104	0.008***	0.385
Y=GDP, idio = log(SSE/SSM)								
DlogGDP$_t$		0.000***	0.33	0.077*	0.028**	0.777	0.679	0.679
Idio$_t$	0.143		0.591	0.67	0.018**	0.163	0.096*	0.635
Y = IP, idio =log(FIRM/MKT)								
DlogIP$_t$		0.017**	0.676	0.219	0.107	0.087*	0.702	0.59
Idio$_t$	0.259		0.157	0.123	0.002***	0.175	0.015**	0.43
Y = IP, Idio = log(SSE/SSM)								
DlogIP$_t$		0.012**	0.677	0.2	0.061*	0.339	0.78	0.464
Idio$_t$	0.158		0.587	0.711	0.035**	0.266	0.162	0.772

1947Q1-2011Q4	*DlogY*$_{t-k, t-1}$	*Idio*$_{t-k, t-1}$	*TERM*$_{t-k, t-1}$	*DEF*$_{t-k, t-1}$	*MRET*$_{t-k, t-1}$	*VOL*$_{t-k, t-1}$	*DIV*$_{t-k, t-1}$	*ILLIQ*$_{t-k, t-1}$
Y = INVEST, idio =log(FIRM/MKT)								
DlogY$_t$		0.007***	0.067*	0.24	0.424	0.032**	0.321	0.19
Idio$_t$	0.403		0.151	0.090*	0.001***	0.084*	0.005***	0.371
Y=INVEST, idio = log(SSE/SSM)								
DlogY$_t$		0.001***	0.008***	0.648	0.664	0.856	0.526	0.068*
Idio$_t$	0.535		0.468	0.41	0.030**	0.199	0.221	0.732
Y = UE, idio =log(FIRM/MKT)								
DlogY$_t$		0.308	0.601	0.559	0.122	0.646	0.832	0.758
Idio$_t$	0.309		0.135	0.182	0.004***	0.207	0.026**	0.518
Y = UE, Idio = log(SSE/SSM)								
DlogY$_t$		0.061*	0.304	0.359	0.035**	0.904	0.668	0.319
Idio$_t$	0.263		0.428	0.574	0.059*	0.274	0.32	0.899

Table 1.6 Predicting Future Economic Activity Out-of-Sample

This table reports the *MSE-F* and *ENC-NEW* test results for the out-of-sample forecasts of the macroeconomic series *dlogGDP, dlogIP, dlogINVEST,* and *dlogUE*. Significant test statistics mean that forecasting accuracy is significantly improved after stock return idiosyncrasy is added in the models forecasting future economic conditions with the variables in the first column of each panel. The in-sample regressions are estimated starting with the first half of the sample period, i.e. 1921Q1-1966Q4 for the full sample period and 1947Q1-1979Q4 for the post-war period. The forecast horizon is set to 1, 2 or 4 quarters. The critical values of both tests are the Monte Carlo simulation value produced by Clark and McCracken (2001). Panels A, B, and C list the out-of-sample test results with the forecast horizon of one, two, and four quarters, respectively. "***", "**", and "*" denote the significance at 1%, 5%, and 10% level, respectively.

Panel A Forecast Horizon: One Quarter

1921Q1-2011Q4	*dlogGDP*				*dlogIP*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	-15.422	13.952***	12.684***	21.434***	-4.699	30.020***	31.111***	48.928***
DEF	-11.739	16.064***	9.786***	18.724***	-6.392	32.352***	21.630***	41.641***
MRET	5.469**	5.246***	4.029**	3.027**	9.312***	7.513***	8.816***	7.216***
VOL	-5.331	11.217***	10.154***	14.980***	-1.684	24.637***	20.749***	35.444***
All of the above	4.435**	5.565***	5.220**	3.809**	11.899***	11.875***	12.644***	10.395***

1947Q1-2011Q4	*dlogGDP*				*dlogIP*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	0.908*	10.765***	8.755***	18.815***	2.611**	17.211***	6.479***	25.874***
DEF	-7.813	6.009***	-4.810	9.685***	-7.668	10.671***	-11.677	13.011***
MRET	-3.226	5.920***	2.638**	11.371***	-6.525	8.714***	-4.927	14.166***
VOL	-3.112	4.891**	3.521**	10.473***	-5.706	6.793***	-4.253	11.877***
ILLIQ	-1.381	4.918**	3.023**	10.403***	-5.784	6.257***	-5.407	12.426***
DIV	-3.923	5.132***	0.183	9.622***	-6.256	6.800***	-7.177	10.590***
All of the above	-3.032	1.825*	5.389**	7.663***	-3.780	2.983**	1.563**	9.291***

1947Q1-2011Q4	*dlogINVEST*				*dlogUE*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	1.484**	10.601***	7.394***	13.600***	5.095**	9.100***	10.142***	14.473***
DEF	-6.746	6.112***	-6.093	5.581***	-2.575	4.342**	-1.009	6.255***
MRET	-8.220	3.346**	-3.140	4.367**	-0.333	2.907**	2.334**	5.053***
VOL	-7.249	3.258**	-2.886	3.861**	-3.571	1.584*	-1.307	3.113**
ILLIQ	-4.248	3.073**	-1.157	4.216**	0.337	2.300*	1.915**	4.528**
DIV	-7.051	2.646**	-3.509	2.980**	-0.418	1.997*	0.833*	3.419**
All of the above	0.224	3.227**	3.030**	3.467**	-0.008	0.623	1.603**	1.803*

Panel B Forecast Horizon: Two Quarter

1921Q1-2011Q4	*dlogGDP*				*dlogIP*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	0.346	0.407	10.242***	6.647***	1.610**	2.237*	15.547***	14.075***
DEF	0.637*	0.412	5.058**	3.474**	5.643**	4.400**	12.686***	11.829***
MRET	-4.663	-0.953	-2.283	-0.743	-3.118	-0.917	3.435**	2.491**
VOL	-10.018	-2.955	-3.931	-1.430	-0.881	-0.342	7.451***	6.211***
All of the above	-1.592	-0.164	3.219**	1.938*	-0.480	-0.136	8.837***	6.174***

1947Q1-2011Q4	*dlogGDP*				*dlogIP*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	5.163**	6.000***	15.007***	16.312***	5.592**	9.813***	13.328***	22.613***
DEF	1.665**	4.156**	5.214**	9.761***	-0.974	6.260***	-3.356	11.572***
MRET	1.929**	1.623*	7.656***	6.421***	3.046**	4.269**	8.116***	12.359***
VOL	3.157**	3.457**	10.243***	10.453***	2.545**	5.002***	7.170***	13.337***
ILLIQ	1.158*	1.669*	5.626**	7.096***	1.649**	5.034***	3.354**	13.327***
DIV	0.947*	0.699	4.406**	3.880**	2.247**	2.316*	4.681**	7.799***
All of the above	4.129**	3.260**	17.313***	14.282***	5.219**	4.486**	15.810***	15.816***

1947Q1-2011Q4	*dlogINVEST*				*dlogUE*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	5.274**	10.391***	8.713***	17.578***	0.849*	7.107***	2.954**	16.184***
DEF	-1.005	7.132***	-8.728	8.064***	-5.509	3.932**	-13.794	6.402***
MRET	-0.366	4.366**	-1.593	7.093***	-0.205	1.518*	-0.179	5.032***
VOL	-1.795	5.206***	-5.049	8.101***	-6.794	1.687*	-14.985	4.705**
ILLIQ	-2.637	3.662**	-5.816	6.478***	-3.167	1.271	-8.231	4.688**
DIV	0.445*	3.430**	-2.098	5.183***	-0.187	0.788	-2.373	3.135**
All of the above	5.870**	7.830***	9.096***	14.538***	-1.693	3.688**	-1.163	12.603***

Panel C Forecast Horizon: Four Quarters

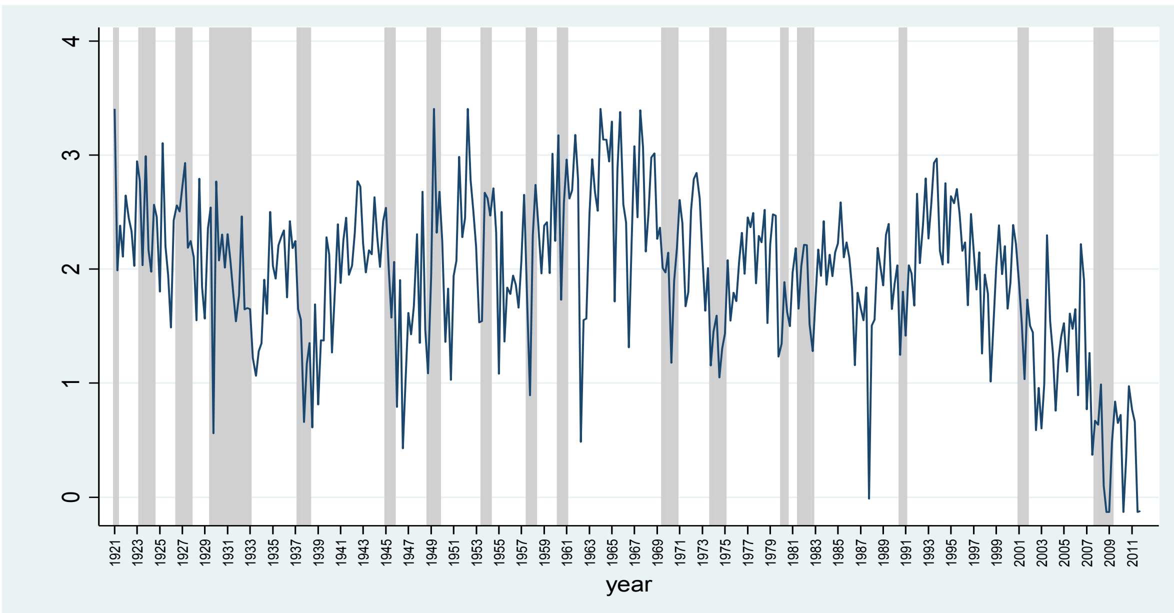
1921Q1-2011Q4	*dlogGDP*				*dlogIP*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	-17.078	-3.554	-3.850	1.828*	-6.835	0.071	6.697***	5.269***
DEF	-8.803	4.024**	0.603*	5.584***	0.784*	6.066***	7.398***	6.282***
MRET	-1.245	-0.569	-0.458	-0.034	-2.385	-0.823	-6.779	-2.337
VOL	-5.013	2.050*	-0.153	3.304**	1.407**	2.564**	3.168**	2.616**
All of the above	-7.193	1.864*	3.527**	4.747**	1.894**	1.667*	1.376**	1.008

1947Q1-2011Q4	*dlogGDP*				*dlogIP*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	2.007**	1.984*	8.666***	8.669***	3.663**	2.624**	10.47***	9.912***
DEF	0.449*	1.076	1.650**	3.975**	1.225*	1.191	1.444**	3.758**
MRET	0.422*	0.545	3.209**	3.912**	0.313	0.419	3.143**	4.260**
VOL	0.547*	0.812	3.271**	4.809**	1.239*	0.817	5.811**	5.116***
ILLIQ	0.367*	0.493	2.635**	3.942**	-0.384	-0.029	3.310**	4.214**
DIV	0.493*	0.575	3.085**	4.120**	0.229	0.422	2.576**	4.508**
All of the above	1.535**	1.399	8.059***	8.934***	1.368*	0.875	7.935***	7.617***

1947Q1-2011Q4	*dlogINVEST*				*dlogUE*			
	log(FIRM/MKT)		*log(SSE/SSM)*		*log(FIRM/MKT)*		*log(SSE/SSM)*	
Restricted model	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*	*MSE-F*	*ENC-NEW*
TERM	1.591**	1.129	6.352**	4.807**	3.032**	1.960*	5.968**	5.129***
DEF	-0.068	0.320	-0.418	1.194	0.472*	0.685	-1.373	0.906
MRET	-1.808	-0.787	-2.208	-0.496	-2.256	-1.001	-2.645	-0.993
VOL	-1.677	-0.664	-2.552	-0.462	-2.338	-0.914	-6.057	-1.956
ILLIQ	-1.699	-0.738	-1.999	-0.266	-2.695	-1.127	-2.878	-1.095
DIV	-1.925	-0.852	-2.275	-0.621	-3.027	-1.280	-3.270	-1.321
All of the above	-0.133	0.041	1.608**	1.966*	-0.335	-0.138	-4.134	-0.830

Figure 1.1 Stock Return Idiosyncrasy over the Business Cycle

Figure 1.1 plots the two stock return idiosyncrasy measures over the U.S. business cycles from 1921Q1 to 2011Q4. *Log(SSE/SSM)* and *log(FIRM/MKT)*, the ratio of firm-specific to systematic return variation, are defined according to the methodologies of Morck et al. (2000) and Campbell et al. (2001), respectively. The shaded areas denote the recession quarters defined by NBER.



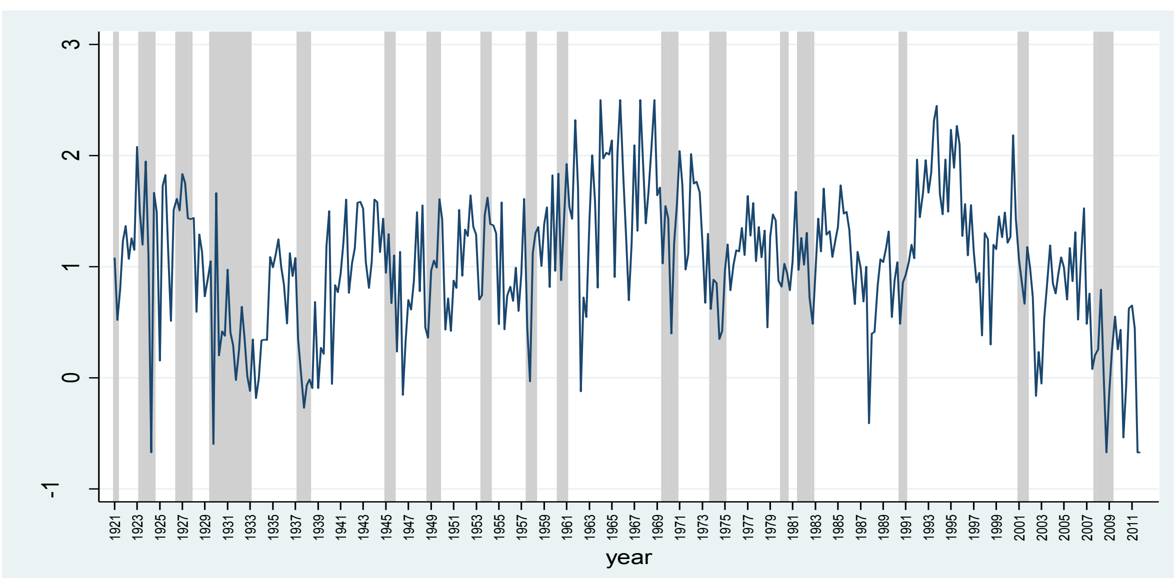
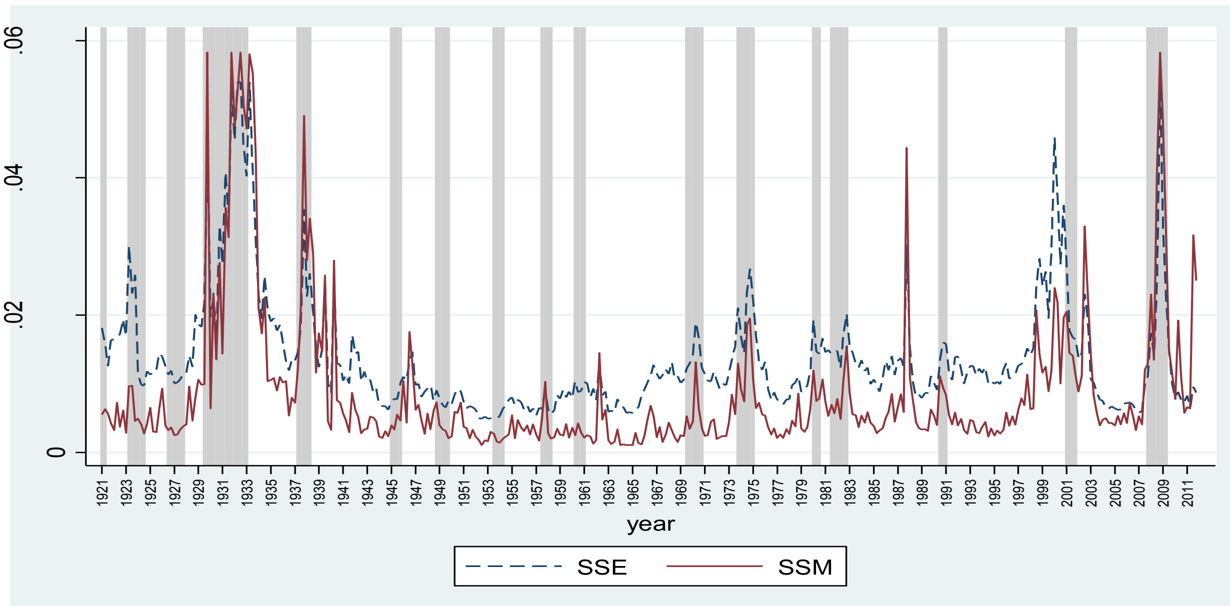


Figure 1.2 Firm-Specific and Market Volatility over the Business Cycle

Figure 1.2 plots the firm-specific and systematic components of stock return volatility over the U.S. business cycles from 1921Q1 to 2011Q4. The firm-specific volatility, measured by *SSE* or *FIRM,* and the systematic volatility, measured by *SSM* or *MKT*, are defined according to the methodologies of Morck et al. (2000) and Campbell et al. (2001), respectively. The shaded areas denote the NBER recessions quarters.



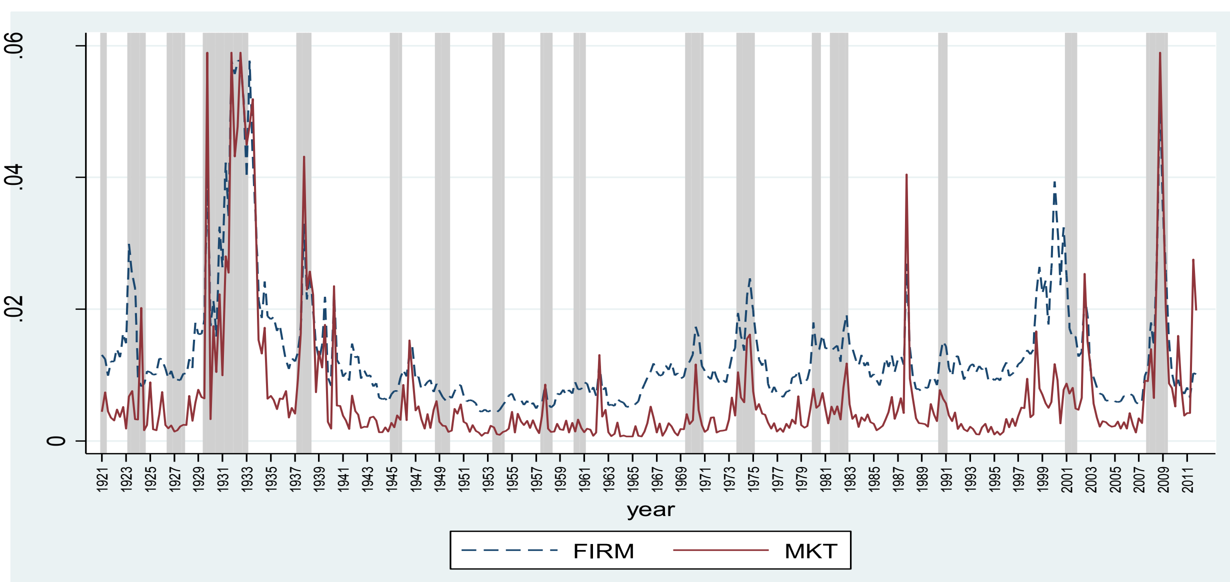
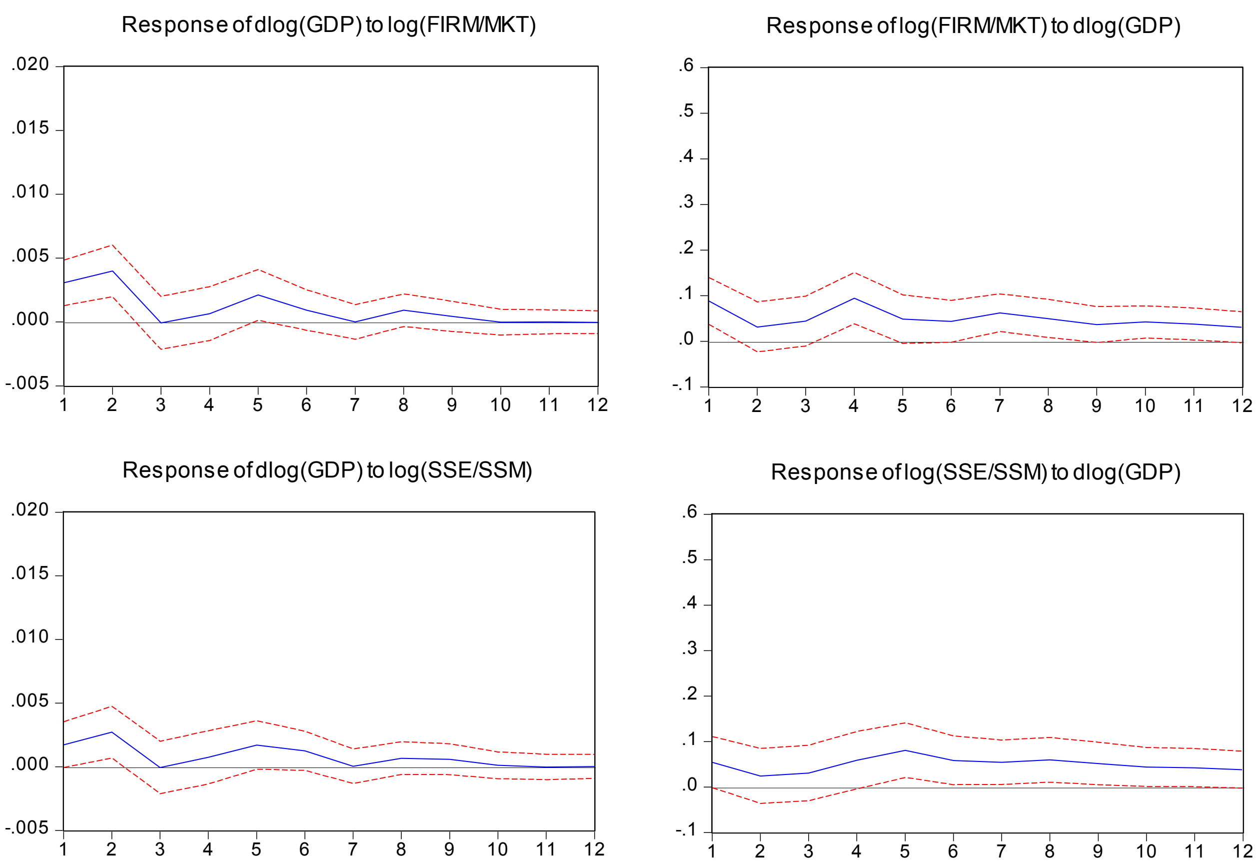
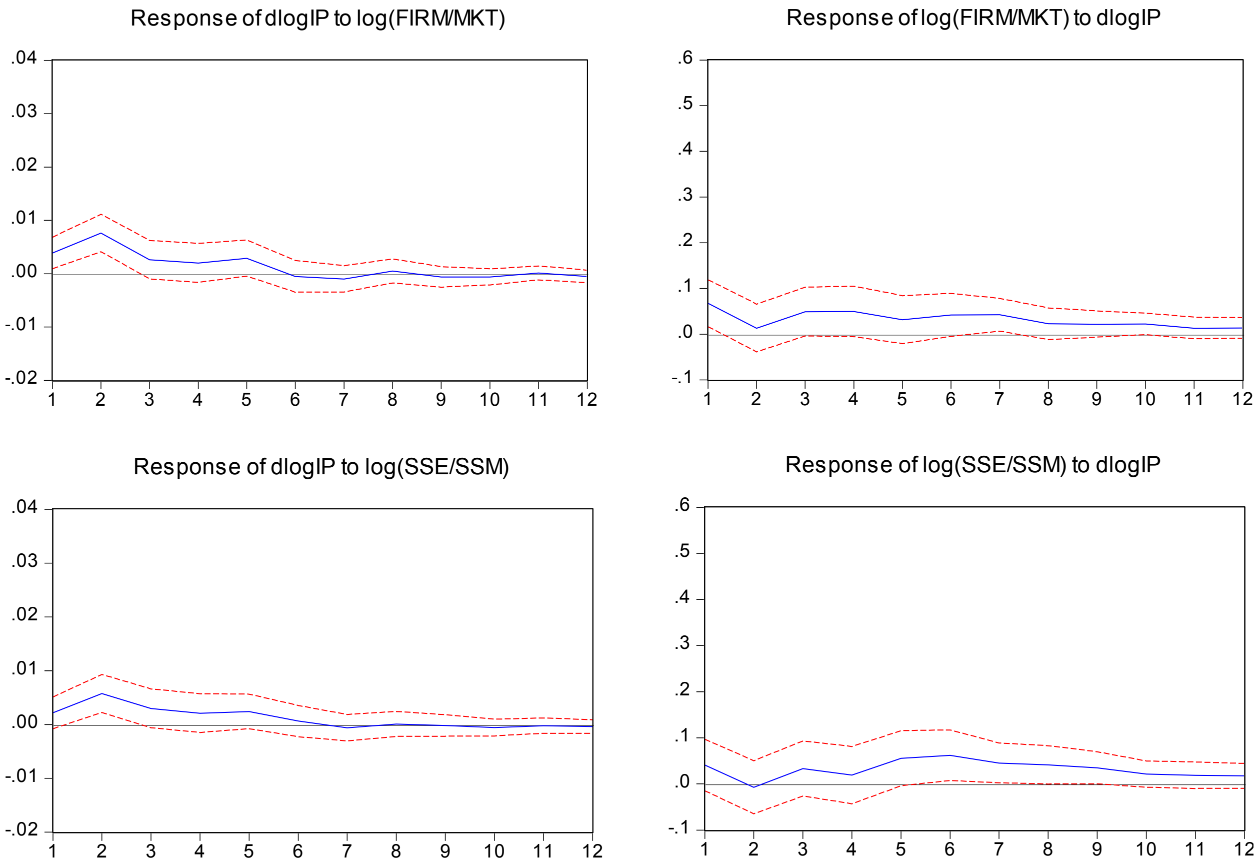


Figure 1.3 Impulse Response Functions from the Multivariate VAR

The solid line in each graph plots the impulse responses to a generalized one standard deviation innovation of the variable in question from the VAR specifications. The dotted line plots the two standard deviation confidence bands of the impulse response functions. Each VAR includes one of the economic condition indicators, one of the stock return idiosyncrasy measures, and the following control variables term spread (*TERM*), default spread (*DEF*), excess market return (*MRET*), market volatility (*VOL*), dividend yield (*DIV*), and illiquidity (*ILLIQ*). The economic indicators are the growth rate of real GDP, industrial production (IP), real fixed invesment (INVEST), and unemployment (UE). The stock return idiosyncrasy is defined as the ratio of firm-specific to systematic return variation. The two measures log(FIRM/MKT) and log(SSE/SSM) are defined according to the methedologies in Campbell et al. (2001) and Morck et al. (2000), respectively. Panel A and B show the impluse response functions over the 1921-2011 and 1947-2011 sample periods, respectively.

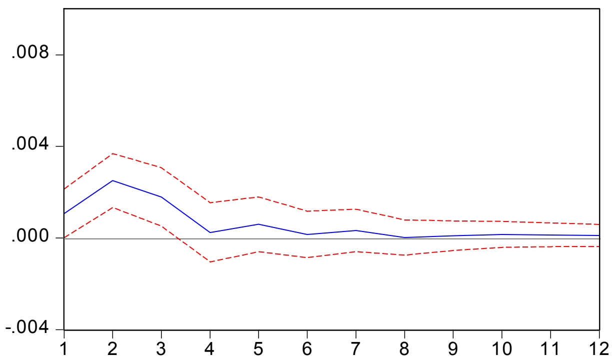
Panel A 1921- 2011



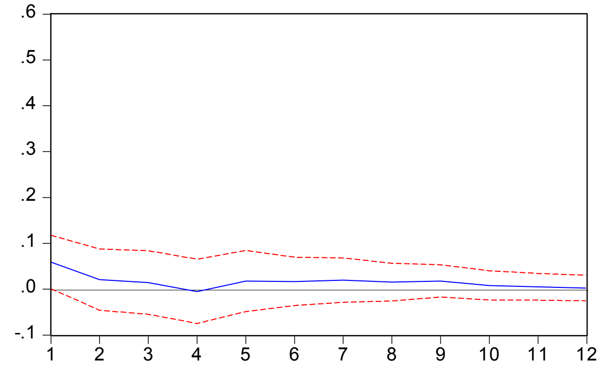

Response of dlogIP to log(FIRM/MKT)
Response of log(FIRM/MKT) to dlogIP
Response of dlogIP to log(SSE/SSM)
Response of log(SSE/SSM) to dlogIP
.04
.03
.02
.01
.00
-.01
-.02
.6
.5
.4
.3
.2
.1
.0
-.1
1 2 3 4 5 6 7 8 9 10 11 12

Panel B 1947-2011

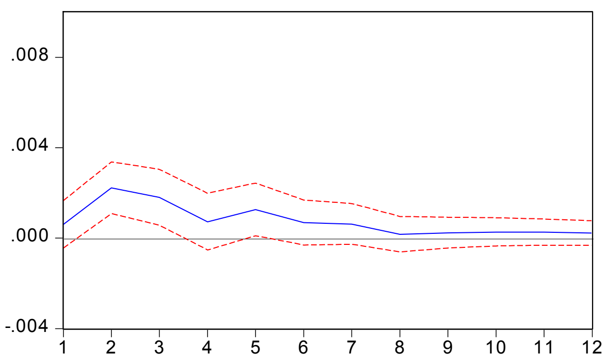
Response of dlogGDP to log(FIRM/MKT)



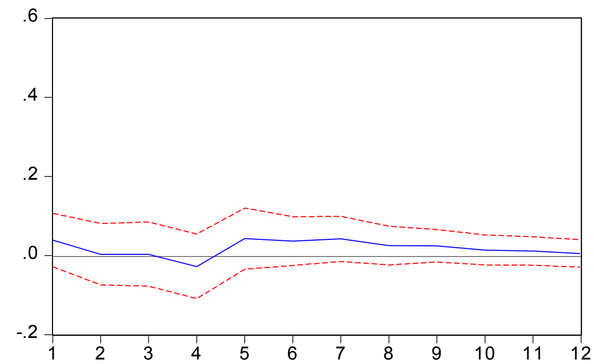
Response of log(FIRM/MKT) to dlogGDP



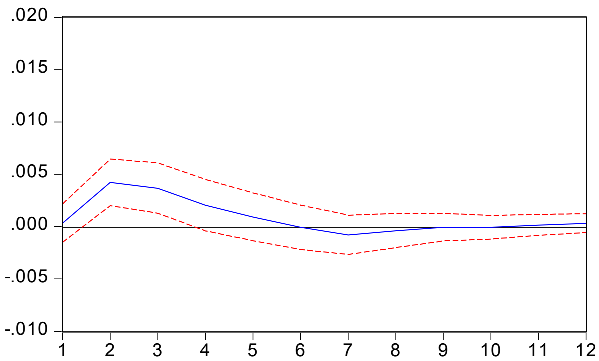
Response of dlogGDP to log(SSE/SSM)



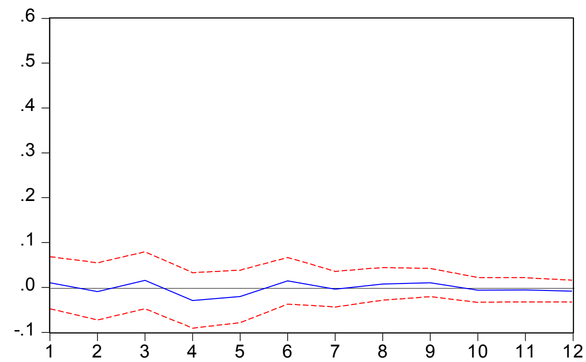
Response of log(SSE/SSM) to dlogGDP

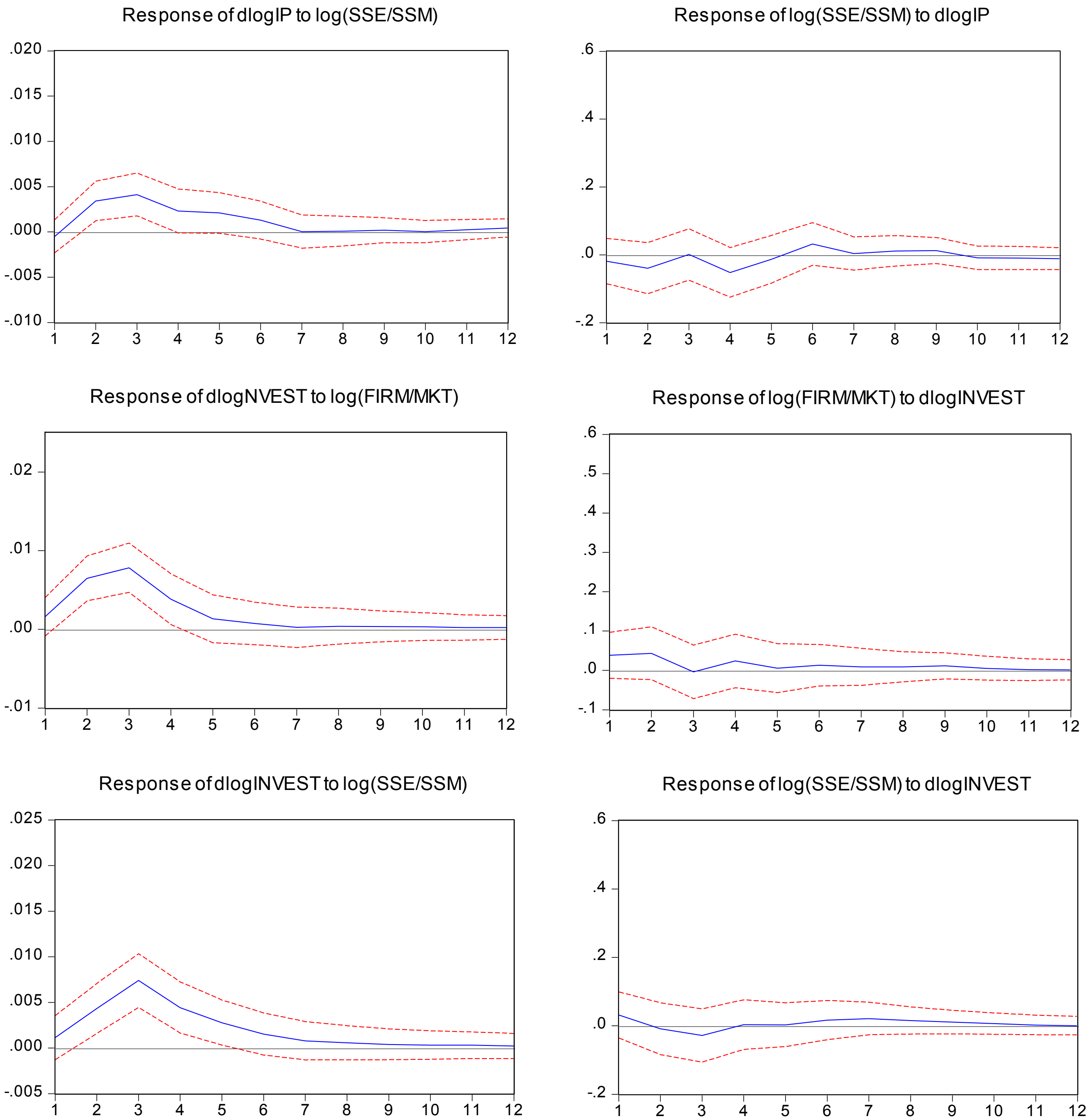


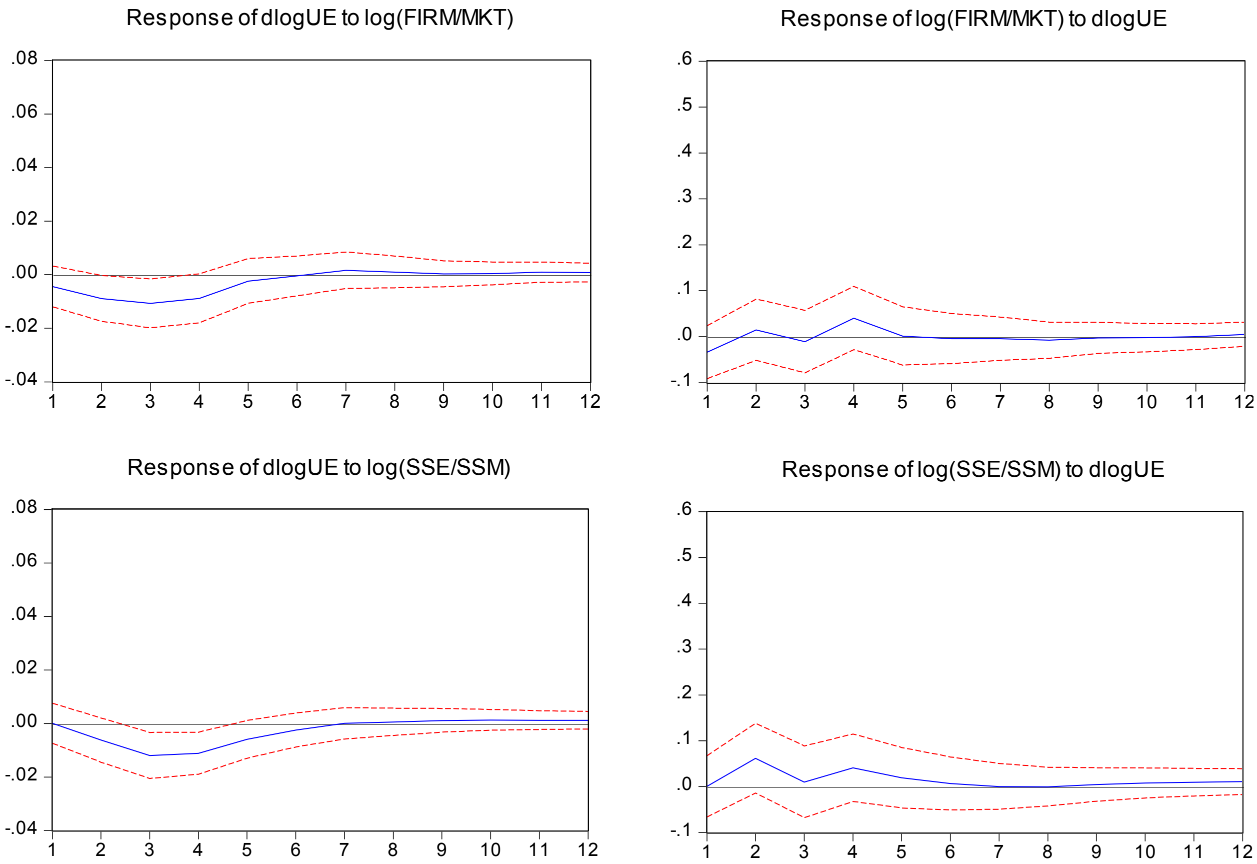
Response of dlogIP to log(FIRM/MKT)



Response of log(FIRM/MKT) to dlogIP




Response of dlogIP to log(SSE/SSM)
.020
.015
.010
.005
.000
-.005
-.010
1 2 3 4 5 6 7 8 9 10 11 12
Response of log(SSE/SSM) to dlogIP
.6
.4
.2
.0
-.2
1 2 3 4 5 6 7 8 9 10 11 12
Response of dlogNVEST to log(FIRM/MKT)
.02
.01
.00
-.01
1 2 3 4 5 6 7 8 9 10 11 12
Response of log(FIRM/MKT) to dlogINVEST
.6
.5
.4
.3
.2
.1
.0
-.1
1 2 3 4 5 6 7 8 9 10 11 12
Response of dlogINVEST to log(SSE/SSM)
.025
.020
.015
.010
.005
.000
-.005
1 2 3 4 5 6 7 8 9 10 11 12
Response of log(SSE/SSM) to dlogINVEST
.6
.4
.2
.0
-.2
1 2 3 4 5 6 7 8 9 10 11 12


Response of dlogUE to log(FIRM/MKT)
Response of log(FIRM/MKT) to dlogUE
Response of dlogUE to log(SSE/SSM)
Response of log(SSE/SSM) to dlogUE

Chapter 2 Creative Destruction and Firm-Specific Return Variation: Evidence from the 1920s and 1930s

2.1 Introduction

Technological change is associated with Schumpeter's process of creative destruction: new and creative firms arise to successfully apply new innovations, causing the complete or partial destruction of old non-innovative firms (Schumpeter, 1912). This paper studies technological innovation and creative destruction during the 1920s and 1930s, one of the most innovative periods in the 20th century (Jovanovic and Rousseau, 2003). Using a comprehensive sample of public U.S. firms, we find that elevated firm-specific return variation is associated with more intense technological innovation.

Firm-specific return variation, defined as the residual return variation that cannot be explained by market or industry related variation, reflects heterogeneity in firm performance (Roll, 1988; Morck, Yeung, and Yu, 2000, 2013; Campbell, Lettau, Malkiel, and Xu, 2001). We argue that elevated firm-specific return variation reflects a widened performance gap between successful and failing firms, and that this can measure the intensity of creative destruction. We hypothesize that innovative firms exhibit higher firm-specific return variation than less innovative firms, and test this hypothesis with both industry-level and firm-level data.

At the industry-level, we find elevated firm-specific return variation in industries obtaining more patents in the 1930s. Despite the sharp economic downturn, multifactor productivity data show that the 1930s was a decade of unmatched technological progress (Field, 2003, 2011). Innovation in chemistry, long-distance communications, electrical machinery, structural engineering, and aviation proceeded largely independently of the depression (Szostak, 1995;

Mowery and Rosenberg, 1989; Nicholas, 2011). The correlation between patents and firm-specific return variation at the industry-level is insignificant in the 1920s bull market. This may in part reflects patent races, when investors are not yet sure which firms will win and which will lose, and the prices of all the firms in the race move together on information about the value of the potential prize (Loury, 1979; Harris and Vickers, 1985, 1987). Thus, the prices of stocks in patent-intensive industries move together in the 1920s. Galbraith (1961) and others see a market mania in the 1920s, however this view is not necessarily inconsistent with patent races. If information cascades lift prices above fundamentals (Bikhchandani, Hishleifer, and Welch, 1992, 1998), patent races could set off episodes of sustained overvaluation. These could, of course, be strengthened by behaviorally-driven return chasing. However, industry-level measures of innovation other than patents – labor productivity, number of research staff, and electricity usage – point to higher firm-specific return variation in more innovate industries in both the 1920s and 1930s, if only tentatively.

At the firm-level, we find elevated firm-specific stock return variation in firms receiving more patents in high-tech sectors such as chemistry, mechanical, and electrical sectors in both the 1920s and 1930s, consistent with informed traders distinguishing innovative firms from their non-innovative or unsuccessful peers in both the boom and recession decades. These innovative firms tend to be small and young, consistent with Schumpeter's assertion (1912) that older and larger firms tend to be more bureaucratic and conservative, and therefore less attractive to creative innovators.

This paper contributes to the literature on technological innovation and economic growth by studying a time period of unprecedented technological advance. Schumpeter (1939) labels the 1920s a decade of industrial revolution. Electricity became the main source of power during this decade. By 1929, electricity provided 79% of the power supply to the U.S. manufacturing industries (Devine, 1983). In little over a decade, automobiles went from rarities to necessities owned by most households. In 1930s, technological advances continued in chemistry, electrical machinery, and aviation. The National Research Council survey of innovation reports that the number of scientists and engineers employed in industrial research laboratories increased from 6,274 in 1927 to 10,981 in 1933 and 27,777 in 1940 (Mowery and Rosenberg, 1998). Our evidence of creative destruction continuing into the Depression decade is also in line with the liquidation theory of recessions asserted by Hayek (1944) and Schumpeter (1939): recessions liquidate the unproductive firms and reallocate resources to the more productive ones. Our empirical findings support the general validity of the endogenous economic growth theory (Romer 1990), but especially support those in which technological progress occurs through creative destruction (Aghion and Howitt, 1992, 1998; Acemoglu, Aghion, and Zilibotti, 2006) and in which elevated firm performance heterogeneity follows technology shocks (Pastor and Veronesi, 2009). Elevated firm-specific return heterogeneity, similar to that evident in the 1920s and 1930s, is apparent as the pace of creative destruction accelerates in the U.S. in the latter half of the twentieth century (Liang, McLean, and Zhao, 2011), and in the high-tech IT boom of the 1990s in particular (Chun, Kim, Morck, and Yeung, 2008). This is also consistent with other findings (Fama and French, 2004; Pastor and Veronesi, 2009; Brown and Kapadia, 2007; Fink, Fink, Grullon, and Weston, 2010) that small and young firms in the 1990s also had highly idiosyncratically moving share prices.

This paper further finds parallels between the late 20th century and the 1920s and 1930s. Morck, Yeung, and Yu (2000) and Campbell, Lettau, Malkiel, and Xu (2001) observe rising firm-specific stock return variation in latter half of the twentieth century, and Wei and Zhang

(2006) find a similar trend in fundamentals variation. Pastor and Veronesi (2005), Fama and French (2004), Brown and Kapadia (2007), and Fink, Fink, Grullon, and Weston (2010) relate this rising heterogeneity to increased numbers of small and young listed firms. Comin and Philippon (2005) and Irvine and Pontiff (2009) link it to stiffening competition in deregulated industries. Wurgler (2000) links higher firm-specific stock return variation to the stock markets better allocating capital to value-creating uses. Morck, Yeung, and Yu (2000), Durnev, Morck and Yeung (2004), Li, Morck, Yang, and Yeung (2004), Jin and Myers (2006) link higher firm-specific return variation to factors associated with financial development, such as less corruption, more openness to foreign investors, and more transparency. Bris, Goetzmann, and Zhu (2007) associate higher firm-specific return variation to more active arbitrage trading. All of these findings coalesce into a unified explanation if innovation that elevates firm-specific stock return and fundamentals heterogeneity are both hastened by deregulation and competition and financed by more functionally efficient (Tobin, 1984) stock markets that more readily list and more accurately price young high-risk firms.[8] This view is consistent with Gompers and Lerner (2004), who link early-stage development of new technologies to deregulation, competition, and well-functioning equity markets, into which early stage venture capitalists can sell out and in which innovators' newly listed young firms can raise large amounts of risk-tolerant capital to grow rapidly.[9] After controlling for size, age, and industrial concentration, we find patterns in

---

[8] Tobin (1984) stresses functional efficiency (how reliably the stock market directs capital to firms with positive net present value investments) over informational efficiency (how closely stock returns follow martingales).

[9] Davis et al. (2007) finds falling firm-specific heterogeneity in the late 20th century in pooled plant-level data for listed and unlisted firms. However, this is consistent with Schumpeter's (1911) view that public equity financing as disproportionately important to the rapid growth of successful innovators and empirical evidence of the critical role of public equity markets in the venture capital cycle in the US (Gompers and Lerner 2004). The heterogeneous effects of technological change might thus be most

firm-specific return variation consistent with a period of intensive creative destruction fuelling the 1920s boom and echoing through the subsequent depression. Our findings thus likewise link elevated performance heterogeneity in the 1920s and 1930s to the adoption of new technologies by smaller and younger listed firms. Many of these firms might subsequently do poorly, as a few succeed brilliantly. All potential innovators might enjoy rising share prices until winners and losers become distinguishable to investors.

These parallels are consistent the arguments of Bresnahan and Trajtenberg (1995), Helpman and Trajtenberg (1998), Gordon (2005), Jovanovic and Rousseau (2005) and others that both electricity, a key innovation of the 1920s and 1930s, and IT, a similarly central innovation of the late 20th century, are general purpose technologies: that is, innovations that induce secondary innovations across the entire economy, and whose impact is thus not confined to one sector. Moser and Nicholas (2004) affirm this for the 1920s, showing electricity patents to be broader in scope and more original than other categories of patents, but to have lower generality and fewer forward citations, and suggest that technological change in the 1920s was much broader than previously thought.

The rest of this paper is constructed as follows. Section 2.2 describes the data and methodology. Section 2.3 examines the relationship between firm-specific return variation and patents at both the firm- and industry-levels. Section 2.4 conducts robustness checks with alternative measures of innovation at the industry-level. Section 2.5 concludes.

---

evident in listed firms, and employment growth in the far larger population of unlisted firms might easily reflect other factors.

2.2 Variable Definition

2.2.1 Firm-specific Return Variation

Firm *j*'s firm-specific return variation in year *t* is from regression

[1] $$R_{j,s} = \alpha_{1,j,t} + \beta_{m1,j,t}R_{m,j,s} + \beta_{i1,j,t}R_{i,j,s} + \epsilon_{j,s}$$

[2] $$R_{j,s} = \alpha_{2,j,t} + \beta_{m2,j,t}R_{m,j,s} + \epsilon_{j,s}$$

explaining $R_{j,s}$, firm *j*'s total (cum dividend) return on day *s*, with value-weighted market and industry returns, $R_{m,j,s}$ and $R_{i,j,s}$ respectively, where *i* is *j*'s industry. To preclude large firms unduly affecting the indexes, both exclude *j*, and so are different for each firm.

Building on Roll (1988), Morck, Yeung, and Yu (2000), and Campbell et al. (2001), we define *firm-specific return variation* in year t as

[3] $$SSE_{jt} = \frac{1}{N_{1,j,t}-1}\sum_{s\in t}\left(\hat{R}_{1,j,s} - R_{j,s}\right)^2$$

The sum of squared errors from [1]; and *j*'s *market-related return variation* as

[4] $$SSM_{jt} = \frac{1}{N_{2,j,t}-1}\sum_{s\in t}\left(\hat{R}_{2,j,s} - \bar{R}_{j,s}\right)^2$$

The model sum of squares from [2], where $N_{1,j,t}$ and $N_{2,j,t}$ are the number of observations for firm *j* used in estimating [1] and [2], respectively, in year *t*. We then define firm *j*'s *industry-related return variation* as the model sum of squares from [1] minus that from [2]

[5] $$SSI_{jt} = \frac{1}{N_{1,j,t}-1}\sum_{s\in t}\left(\hat{R}_{1,j,s} - \bar{R}_{j,s}\right)^2 - \frac{1}{N_{2,j,t}-1}\sum_{s\in t}\left(\hat{R}_{2,j,s} - \bar{R}_{j,s}\right)^2$$

where $\hat{R}_{js,1}$ and $\hat{R}_{js,2}$ are firm j's estimated day s return from (1) and (2), respectively. Firms with fewer than 30 observations and industries with fewer than 5 firms in year t are dropped.

Following Morck et al. (2000), we define firm j's *absolute* and *relative* firm-specific return variation as $ln(SSE_{j,t})$ and $\Psi_{j,t} = ln(\frac{SSE_{j,t}}{SSM_{j,t}+SSI_{j,t}})$, respectively. Log transformations generally generate approximately Gaussian distributions. Industry i's absolute and relative firm-specific return variations are value-weighted averages of the firm-level measures of all firms in each industry, and economy-level analogs are value-weighted averages of all firms in each year. Using equal weighted or trading day weighted industry-level return variations does not materially change the results.

We estimate [1] and [2] using daily cum dividend stock returns for all NYSE firms every year from 1921 to 1939, encompassing both the 1920s boom and the 1930s bust. Data are from Mehrotra et al. (2015)[10] up to 1925 and from CRSP thereafter. On zero trading days, CRSP approximates missing prices with means of bid and ask closes. We drop these observations for compatibility with pre-CRSP data. Including them does not materially change the results. The final dataset contains 1,056 stocks and 2,520,449 daily return observations.

---

[10] Daily stock prices, dividends, and end-of-week shares from 1921 to 1925 are from *The New York Times* and *The Commercial and Financial Chronicle.*

2.2.2 **Industry and Firm-level Patents**

To gauge the intensity of innovation in each firm's industry, we take industry-level annual patent data from *Historical Statistics of the United States.*[11] These use the *Yale Technology Concordance* to allocate patents into 50 *sectors of use*. For example, a fertilizer patent, received by a chemical firm but useful in agricultural is allocated to the agricultural sector. There are 50 sectors of use in this historical patent database based on the 1980 SIC-E codes. We assign each firm a sector of use based on descriptions of its business from *Moody's Industrial Manuals*, the firm's website (if this provides a corporate history), and/or Wikipedia. Because some sectors of use contain few firms, we combine some closely related ones to obtain the 29 industry classifications listed in Table 2.1. These include both the new "high-tech" sectors of the era, such as automobiles, radio, and communications; as well as the well-established industries, such as mining and food. These data are collected each year from 1910 to 1939.

To gauge the intensity of innovation by each individual firm, we obtain firm-level patent data for 131 U.S. manufacturing firms from 1910 to 1939 from Nicholas (2008), who shows these firms to be representative of listed firms at the time.[12] Of these, 107 are matched with firms for which we can estimate the daily stock return volatility measures described in section 2.1. Note that our firm-level patents do not add up to industry-level patents because the industry-level data include patents by unlisted firms, listed firms for which we do not have data (such as those

---

[11] See http://hsus.cambridge.org/HSUSWeb/toc/hsusHome.do, Table Cg69–107, Patents granted, by sector of use (SOU): 1840-1996.

[12] See "Nicholas, Tom, Does Innovation Cause Stock Market Runups? Evidence from the Great Crash, *American Economic Review*, 2008, 98:4, 1370-1396" for the detailed description of data source. We are grateful to Tom Nicholas for sharing the data.

trading on regional exchanges), and firms in other industries that obtain patents valuable in this industry.

We define three patent measures at the industry- and firm-level as measures of exposure to innovation. The first patent measure is the number of patents granted to a firm or an industry in the prior 10 years:

[6] $$log(PT10y)_{i,t} = log\left(\sum_{k=1}^{10} Patents_{i,t-k}\right),$$

where i is an industry, and t is the year. An analogously defined firm-level measure, $log(PT10y)_{j,t}$ is defined using only patents taken out by firm j. A ten-year cut-off is necessitated because the firm-level patent data begins in 1910 and the return data begins in 1921. This constraint is not unreasonable: while current patents likely reflect past innovation, patents granted further in the past are likely less useful indicators of the innovation activity today.

To account for older patents losing relevance as new technologies come on stream, we follow Hall (1990); Hall, Jaffe, and Trajtenberg (2005); and others in assuming a 15% depreciation rate for knowledge based assets in calculating our second patent measure:

[7] $$log(PTstock)_{i,t} = log\left(\sum_{k=1}^{10}(1-0.15)^{k-1} \times Patents_{i,t-k}\right)$$

The analogous firm-level variable is $log(PTstock)_{j,t}$. Alternative measures using depreciation rates of 10%, 20%, or 30% does not qualitatively change the results.

Because Jovanovic and Rousseau (2003) argue that the 1920s was an era of extremely fast-paced creative destruction, we supplement [6] and [7] with a third measure, the number of patents for industry i the prior year:

[8] $$log(PT1y)_{i,t} = log(Patent_{i,t-1}) \; .$$

With $log(PT1y)_{j,t}$ the firm-level analog, the number of patents firm j takes out the prior year. These are essentially equivalent to [7], but impose a 100% depreciation rate, assuming that only newly granted patents reflect the intensity of relevant innovation.

2.2.3 **Summary Statistics**

Figure 2.1 summarizes the distributions of the patent intensity measures from section 2.2 across industries (Panel A) and across firms (Panel B). The typical industry has a ten-year accumulation ($PT10y_{i,t}$) of over 10,000 patents, a 15% depreciation-adjusted ten-year patent stock ($PTstock_{i,t}$) of over 1,000 in a typical year, and 100 - 2,500 patents one year old or newer ($PT1y_{i,t}$). At the firm-level, the prior ten-year patent count ($PT10_{j,t}$) and depreciation-adjusted patent stock ($PTstock_{j,t}$) echo industry-level patterns, but at much lower levels. In a typical year, the typical firm has a ten-year patent stock of 1 to 50 patents, whether this is adjusted for depreciation ($PTstock_{,t}$) or not ($PT10_{j,t}$). Fewer than one in five firms has no patents over the prior decade; but many have no patents the prior year ($PT1y_{i,t}$). Only about 10% of the firm-year observations record more than 50 patents granted the prior year. Because all three patent measures are skewed to the right at both the industry- and firm-levels, we use their logs in regressions.

Table 2.2 breaks these data down into industries, listing the total numbers of patents in each industry from our industry-level (Panel A) and firm-level (Panel B) data from 1910 to 1939. Panel A ranks at the top several high-tech industries in the 1920s, including electrical equipment and motor vehicle, with large numbers of patents per firm. Others high-tech industries, including radio, electrical appliance and drugs, fall into the bottom half of the panel. This patent distribution suggests that some innovative industries might have a small number of important patents, rather than a large number of patents.

Panel B classifies the 107 firms with firm-level patent data into four sectors: chemical, electrical, mechanical, and “other” sectors. This broadened industry classification still distinguishes the innovating sectors from the others but also allows a statistically meaningful number of firms in each sector. Among these four sectors, electrical sector has the largest number of patents per firm and “others” sector has the smallest. Examples of the sample firms in electrical sector are American Telephone & Telegraph, Westinghouse Electric and Manufacturing Co., and Radio Corporation of America, all of which are well-known innovative companies in the 1920s. Automobile companies such as General Motors and Peirce-Arrow Motor Car are classified into the mechanical sector. Examples of chemical firms include Du Pont, American Agricultural Chemical Company, and Coca Cola. National Lead Company, Liggett and Myers Tobacco are classified into “others” sector.

Figure 2.2 plots the market wide time series of the stock return variation components from 1921 to 1939. The levels of both firm-specific return variation and systematic return variation rise during the depression. The ratio of firm-specific return variation to systematic variation is low in the economic downturns of 1929-1933 and 1937-1938, higher between the downturns, and is the highest from 1921 to 1928; consistent with higher firm specific return variation when creative destruction is more intensive.

Tables 2.3 and 2.4 report summary statistics and the correlation matrix at both the industry- and firm-level, respectively. On average, firm-specific variation, *log(SSE),* exceeds systematic variation, *log(SSM)* and *log(SSI),* at both the industry- and firm-level. These three components of stock return variation are positively and significantly correlated, consistent with their time-series pattern: low in the boom decade and high in the subsequent depressions.

We include industry age, market capitalization and Herfindahl Index as controlling variables in the industry-level regressions. Industry age, the average years that firms in the industries are listed on NYSE, averages about 13 years and varies from 11 to 30 years. Industry market value, the total market capitalization of all listed firms in an industry, averages $1,178 million and varies from $522 million to $4,205 million. Consequently, we use the logs of both variables in the regressions. To gauge industry competition, we use year-end firm market capitalizations to construct an annual industry-level Herfindahl Index[13]. Industry age, size, and competition are negatively and significantly correlated with firm-specific return variation. The simple correlation between industry-level patents and industry-level firm-specific return variation is insignificant in general. The accumulated patent measures are highly correlated with the new patent measure, suggesting that industries tend to have very persistent levels of patent activity during the sample years (Nicholas, 2011).

At the firm-level, we include firm age, market capitalization, and industry rival patents as controlling variables in all regressions. The industry-rival patent measures are defined as the sums of the firm-level measures, excluding those for the firm in question. This prevents intense innovators from affecting their own industry benchmarks.

[13] Herfindahl index is usually defined as $\sum_j S_j^2$, where S_j is firm j' market share in its industry. In the 1920s, public companies did not have an obligation to disclose standardized financial reports. There were no mandatory accounting standards to make financial numbers comparable across firms. We therefore use year-end firm market capitalizations as a substitute to construct the Herfindahl Index.

The firm-level patent measures are negatively and significantly correlated with both firm-specific return variation measures. However, the correlation between industry-rival patents and firm-specific return variation is consistently positive and significant, suggesting that firms share prices being positively affected by their rivals' innovation. Consistent with the industry-level results, smaller and younger firms exhibit higher firm-specific return variation.

2.3 Patents and Firm-Specific Return Variation

This section examines whether or not firms and industries with more patents exhibit higher firm-specific return variation. We conduct the analysis at the industry-level and then at the firm-level with a subsample of firm-level patent data.

2.3.1 Industry-level Patents and Firm-Specific Return Variation

Using the industry-level panel data, we estimate the following regressions to test the relationship between firm-specific return variation and innovation.

[9]
$$ln(SSE_{i,t}) = \alpha + \beta_1 Patent_{i,t} + \beta_2\, log(MV_{i,t}) + \beta_3\, log(Age_{i,t}) + \beta_4 HHI_{i,t} + \beta_5 \ln(SSM_{i,t}) + \beta_6 \ln(SSI_{i,t}) + Ind_i + Year_t + \epsilon_{i,t},$$

[10]
$$\Psi_{i,t} = \alpha + \beta_1 Patent_{i,t} + \beta_2\, log(MV_{i,t}) + \beta_3\, log(Age_{i,t}) + \beta_4 HHI_{i,t} + Ind_i + Year_t + \epsilon_{i,t},$$

where the $ln(SSE_{i,t})$ and $\Psi_{i,t}$ are measures of the absolute and relative firm-specific return variation of industry *i* in year *t*, respectively. $Patent_{i,t}$ denotes one of the three patent stock measures $log(PT10y)_{i,t}$, $log(PTstock)_{i,t}$, and $log(PT1y)_{i,t}$. $Log(MV_{i,t})$, $log(Age_{i,t})$, and $HHI_{i,t}$ are included to control for industry size, age, and competition.

Regressions on the relative return variation $\Psi_{i,t}$ essentially have the market-wide and industry-wide return variation on the right-hand side of the equation and constrain their coefficients to one. Eq. [9] thus is a more general regression specification with $ln(SSM_{i,t})$ and $ln(SSI_{i,t})$ included as the right-hand-side variables. For all regressions, industry fixed effects are included to control for potential time-invariant factors that may affect firm-specific return variation; and time fixed effects are included to control for common factors among all industries over time. . Standard errors clustered by industry.

Panel A of Table 2.5 reports the regression results of industry-level firm-specific return variation on industry-level patents from 1921 to 1939. Industries with more patents exhibit higher *relative* firm-specific return variation. The coefficients of the patent measures are positive and significant across all three patent measures, suggesting that stock returns of innovative industries capitalize more firm-specific events *relative* to systematic events. The coefficients of the patent measures are positive but insignificant in regression explaining absolute firm-specific return variation. Industries with higher firm-specific return variation tend to be smaller, younger, and more competitive, consistent with the findings of Fama and French (2004) and Irvine and Pontiff (2009).

We next test how the relationship between patents and firm-specific return variation varies during economic boom and the subsequent recessions. Panels B and C of Table 2.5 report the regression results during the boom decade, 1921-1928, and the depression decade, 1930-1939, respectively. The year 1929 is dropped out because the market peaked and then crashed during this year.

In the boom decade (Panel B), the correlation between patents and firm-specific return variation is insignificant at the industry-level. We suggest two potential explanations. First, investors may not need to distinguish successful innovators from their peers during a bull market. They can profit by betting on the rising industries rather than picking out successful innovators in each industry. Industries with more patents should then exhibit higher industry-specific return variation. To test if this is the case, we regress industry-specific return variation, *log(SSI)* or *log(SSI/(SSM+SSE))*, on industry-level patents. The top half of Panel D reports the regression results. In the 1920s, the coefficients of all three patent measures are positive, and the coefficient of patent stock, *log(PTstock)*, and prior year's patents, *log(PT1y),* are significant at 10% level, consistent with investors intensively revaluing industries with more innovation outcomes. Second, innovation has been shown to be an important driver of the stock market run-ups (Hobijn and Jovanovic, 2001; Nicholas, 2008; Pastor and Veronesi, 2006 and 2009; Frehen, Goetzmann, and Rouwenhorst, 2013). If speculators buy high-tech stocks that others are buying (Shleifer and Summers, 1990; Bikhchandani, Hishleifer, and Welch, 1992), the stock prices may not be informative of either firm-specific or systematic events.

In contrast, industries granted with more patents exhibit significantly higher firm-specific return variation during the depression decade (Panel C). The coefficients of the patent measures are positive and significant across all specifications. These findings support Friedrich Hayek's argument about cathartic recessions (1944): economic downturns are cleaning and cathartic phase of business cycles that expose the loser who manage to get by during booms. The successful and the unsuccessful innovators pooled together during booms are thus separated by recessions, lifting firm-specific return variation during recessions. The regressions explaining industry-specific information during the depression decade are also consistent with this explanation. Panel D shows a significant negative correlation between patents and industry-specific return variation, suggesting that stocks in industries granted with more patents move less with their industry trends during the 1930s[14].

2.3.2 **Firm-level Patents and Firm-Specific Return Variation**

Using a subsample of 107 manufacturing firms[15] with firm-level patent data, we test the robustness of the industry-level results above. The firm-level patent data from 1910 to 1939 are hand collected by Nicholas (2008)[16]. The firm-level patent measures are the analogs of the industry-level measures. Due to the restriction of the sample size, these 107 firms are classified into four broader sectors: chemical, electrical, mechanical, and "other" industries. We then merge the firm-level and the industry-level patent dataset by aggregating the 29 industries in the previous section into these four sectors[17]. We define the *industry-rival* patent measures for each firm as the industry-level patents minus the patents granted to the firm in concern. While the breadth of these sectors might seem excessive for some purposes, very broad industry measures are arguably defensible when investigating general purpose technologies.

---

[14] In unreported results, regressions using market wide return variation *log(SSM)* or *log(SSM/SSE)* as the left-hand side variable also show significant and negative correlation between market-wide return variation and patents in the 1930s.

[15] As shown in Nicholas (2008), this subsample is representative of a broader set of U.S. public firms.

[16] See "Nicholas, Tom, Does Innovation Cause Stock Market Runups? Evidence from the great Crash", *American Economic Review*, 2008, 98:4, 1370-1396, for a detailed description of the data source and data collecting procedure. The original Nicholas (2008) sample contains 131 manufacturing firms. A subsample of 107 firms can be matched with our annual firm-specific return variation data.

[17] The industries in the industry-level patents dataset are combined into four industry groups as follows: chemical and drug are combined as the chemical industry; electrical appliance, electrical lighting, radio and television, electrical industrial equipment, and communication are combined as the electrical industry; Transport, other office machinery, other machinery, and instruments are combined as the mechanical industry; the rest of the industries are combined as the other industry.

Table 2.6 reports the regression results at the firm-level. At first glance, the correlation between firm-specific return variation and the firm-level patents is insignificant from 1921 to 1939. The coefficients of industry-rival patents are negative and significant in regressions explaining absolute firm-specific return variation. When splitting the sample period into two sub-periods 1921-1928 (Panel B) and 1930-1939 (Panel C), we find that the coefficients of industry-rival patents become significantly negative in the 1920s and significantly positive in the 1930s in regressions explaining both absolute and relative firm-specific return variation.

The negative coefficients of the industry-rival patents in the 1920s could be explained by patent race and speculative trading lifting all possibly affected stocks in the high-tech industries. In a patent race, firms competing for an invention and the prize of obtaining patents. It is unclear which firm will win the race. Rational investors may thus invest in the industries with greater innovation outcomes in the past. These industries also tend to be chased by speculative traders (Nicholas, 2008). This industry-wide investment strategy reduces the amount of firm-specific information capitalized into stock prices of firms in the innovative industries. As the winning firms distinguish themselves from the losing ones during the depression, stock prices of innovative industries becomes more informative about firm-specific events. Firms in innovative industries exhibit higher firm-specific return variation.

These results are consistent with the previous findings about performance heterogeneity during the Great Depression. For example, Bresnahan and Raff (1991) show that automobile plants that do not adopt mass-production lines remain in operation during the 1920s but were more likely to be shut down between 1929 and 1935. Mowery (1989) examines the industrial research during the interwar years and find that innovation contributes to firms' survival.

Galambos (2000) shows that companies investing heavily in R&D, such as RCA, AT&T, and Du Pont, return to profitability far before the onset of the World War II.

We next examine whether the correlation between firm-level patents and firm-specific return variation in Table 2.6 varies across sections, as Table 2.6 shows that the correlation is insignificant in industry averages. As shown in Table 2.2, firms in chemical, electrical, and mechanical sectors are granted a lot more patents than firms in other industries. We therefore add the interactions between patents and the three "high-tech" sector dummies in the regressions.

Table 2.7 reports the regression results. In the full sample period (Panel A), the interaction terms between patents and the three high-tech sector dummies are positive and significant in the regressions explaining relative firm-specific return variation. The interactions between the patent measures and electrical and mechanical sectors dummies are also significantly positive and significant in the regressions explaining absolute firm-specific return variation. These results suggest that the stock prices of innovative firms in high-tech sectors are more informative about firm-specific events than that of their industry peers. Panel B and Panel C show consistent results in the two sub-sample periods, consistent with informed traders distinguishing innovative firms from their non-innovative peers in the high-tech sectors in both the boom and depression decades. The interaction between patents and electrical sector dummy becomes insignificant in the 1930s. This is probably because electricity, as a general purposed technology, has been widely adopted across the country and the electrification era has ended by 1930 (Jovanovic and Rousseau, 2005).

2.4 Alternative Measures of Technological Innovation

Patent count is a useful, if sometimes controversial, indicator of technological innovation (Schankerman and Pakes, 1986; Hall, Jaffe, and Trajtenberg, 2002; Schmookler, 1966; Griliches, 1984). We construct alternative industry-level measures of innovation and explore whether these measures contribute to the firm-specific return divergence during the 1920s boom and the subsequent recessions.

2.4.1 **Labor Productivity**

An alternative approach to measuring technological innovation is to search for its effects on labor productivity. Technological progress affects the economy by boosting overall productivity growth (David and Wright, 2003). Growth in labor productivity thus is an outcome of technological innovation. We construct two measures of labor productivity at the industry-level: *VA/Wage* and *VA/Employ,* where *VA* is value-added, *wage* is the total payrolls, and *employ* is the number of employees. All three series are from the *Statistical Abstracts of the United States* (1921-1942)[18]. The growth in labor productivity is defined as the log of labor productivity in year *t* minus that in *t-1*. The U.S. Census Department only collects these data for 13 manufacturing industries every two years[19], so we classify our sample firms to match the census

[18] The annual *Statistics Abstracts of the United States* is downloaded from http://www.census.gov/prod/www/statistical_abstract.html

[19] We exclude the 1937 and 1939 survey data for the textiles, chemical, petroleum and coal, iron and steel, nonferrous metals, machinery, and transportation equipment industries because the U.S. census states that their post-1937 labor series are not comparable with those in the pre-1937 years.

industries[20] and use two-year averages of the annual series to match the biannual labor productivity growth.

Table 2.8 reports these regressions. The labor productivity measures' coefficients are positive and mostly significant in regressions explaining absolute and relative firm-specific return variation, suggesting that industries with elevated firm-specific return variation experience faster labor productivity growth. The coefficients of the control variables indicate that smaller and younger industries tend to have higher firm-specific return variation. These findings are consistent with those using patents as measures of innovation: innovative industries exhibit more high firm-specific return variation, and especially in recessions.

2.4.2 **Research Staff**

The second alternative measure of technological innovation is research staff. The U.S. National Research Council surveys research personnel employed by companies and research laboratories in 16 manufacturing industries[21] in 1921, 1927, and 1933. We define the log of the number of research staff, *log(RDstaff)*, and the percentage of research staff in total employment, *RDstaff %*, as proxies for innovation input. The sample firms are reclassified into the 15 manufacturing sectors and the averages of the years in between two survey year (including the current survey year) are matched with the survey years.

---

[20] These are: stone, clay, and glass products; iron and steel products; nonferrous metals; machinery; and transportation equipment. The following three census industries have no firms in our sample: forest product, railroad repair shops, and miscellaneous.

[21] There are 19 industries sectors in this survey dataset: food; tobacco; textiles; lumber products; furniture; paper; publishing; chemicals; petroleum; robber products; leather; stone, clay, and glass; primary metals; fabricated metals; nonelectrical machinery; electrical machinery; transportation equipment; and instruments. We combined textiles with apparels, and dropped lumber products and furniture due to the small number of firms in these industries.

Table 2.9 summarizes regression explaining firm-specific return variation with the research staff measures. The coefficients of the research staff measures are positive across all specifications. The coefficient of *log(RDstaff)* is significant in regression explaining absolute firm-specific return variation, and coefficient of *RDstaff %* is also significant in the regression explaining relative firm-specific return variation, despite the small sample size (40 observations), which may limit the power of the statistical test. The industry size and age control variables again attract negative sign. These tests using research staff as a proxy of innovation input thus align with those in the previous sections in being consistent with innovative industries exhibiting more heterogeneous firm-specific returns.

2.4.3 **Electrification**

Our third alternative measure of innovation is industry-level electrification, which Jovanovic and Rousseau (2005, see Figures 4 and 8) identify as an economically important general purpose technology (an innovation that opens allows new production possibilities and improvements in efficiency in many economic sectors) in the early 19th century. We can thus use the log of electricity consumption in horsepower by each industry (*electrification*) to proxy for industry-level technological progress associated with electrification.

Table 2.10 summarizes regressions explaining measures of firm-specific return variation with electrification. The coefficients of electrification are positive and significant in regressions explaining absolute firm-specific return variation, consistent with stock returns in industries using more electrical power capitalizing more firm-specific information events, but are negative and insignificant in regressions explaining the relative firm-specific return variation.

Because Jovanovic and Rousseau (2005) argue that electrification was largely complete by the end of the 1920s, we consider the 1920s and 1930s separately by including an interaction term between the electrification variable and a 1920s dummy. The coefficient of the interaction term are now also positive and significant in the regression explaining relative firm-specific return variation, suggesting that stock returns in industries consuming more electricity incorporate more firm-specific information relative to market information. The statistical power of the test again is limited by the sample size (28 observations).

2.5 Concluding Remarks

We study creative destruction, and firm and industry characteristics that are potentially related to it, during the 1920s boom and the subsequent downturn. This is an era of intense innovation that includes the diffusion of at least one general purpose technology - electric power.

Industries applying more patents are characterized by higher firm-specific stock return variation in the 1930s, but not in the 1920s. We argue that the 1920s results could be explained by innovation races or information cascades lifting the stock prices of the whole high-tech industries. The depression in the 1930s distinguishes successful innovators from the unsuccessful ones in the affected industries. Stock prices in industries with more patents thus become more informative.

Individual firm's patents are related to elevated firm-specific return variation only in the era's high-tech industries: chemical, electrical, and chemical sectors, and in both the 1920s and the 1930s. This is consistent with informed traders being able to distinguish successful innovators within the high-tech industries, even as stocks across the entire industries rise in the 1920s.

At both the industry- and firm-levels, small and young firms/industries tend to exhibit higher firm-specific return variation, consistent with the prior findings using more recent data (Fama and French, 2004; Irvine and Pontiff, 2009). Using labor productivity, research staff, and electricity usage as alternative measures of innovation intensity at industry-level, we find evidence in support of our main findings above.

Our findings are consistent with Schumpeter's (1912) creative destruction occurring during the 1920s and 1930s. New and creative firms arise to successfully apply new technologies, thereby destroy non-innovative firm and unsuccessful innovators. A similar wave of creative destruction is observed in the later 20th century by Chun et al. (2008). They show that industries intensively invest in information technology exhibit firm-specific performance heterogeneity. We further provide evidence of the creative destruction continuing into the downturns, during which the successful innovators stands out from their unsuccessful peers.

Table 2.1 Industry Classification of the Industry Patents Dataset – Sector of Use (SOU)

This table lists the sectors of use (SOU) based on the SIC-E 1980 industry classification used by the industry-level patent dataset and the revised sector of use for the analysis of this paper. Industries with closely correlated business activities are combined. "-" indicates industries with no sample firms.

	Sector of Use (SOU)		Revised Sector of Use (SOU)
1	Agriculture-- livestock		-
2	Agriculture-- crops and combo farms		-
3	Agriculture-- fruits and vegetables	1	Food
4	Agriculture-- horticulture		-
5	Agriculture-- services to livestock		-
6	Agriculture-- services to crops		-
7	Agriculture-- other		-
30	Food-- meat, poultry and fish	1	Food
31	Food-- fruit and vegetables		
32	Food-- dairy products		
33	Food-- cereals and feed		
34	Food-- beverages		
36	Food-- other		
8	Forestry and fishing		-
9	Mining	9	Mining
39	Non-metallic minerals		
10	Electrical appliances	10	Electrical appliances
11	Electrical lighting	11	Electrical lighting
12	Radio and television	12	Radio and television
13	Electrical industrial equipment		
14	Other electrical equipment	13	Electrical industrial equipment
15	Electronic equipment		
16	Chemicals	16	Chemicals
17	Drugs	17	Drugs
18	Petroleum	18	Petroleum
19	Transport-- aerospace		-
20	Transport-- motor vehicles	20	Transport-- motor vehicles
21	Transport-- ships	21	Transport-- ships
22	Transport-- other	22	Transport-- other
23	Ferrous metals	23	Metals - ferrous and nonferrous
24	Nonferrous metals		
25	Fabricated metals	25	Fabricated metals
26	Instruments	26	Instruments
27	Computers and peripherals		-
28	Other office machinery	28	Other office machinery
29	Other machinery	29	Other machinery
35	Food-- tobacco	35	Food-- tobacco
37	Textiles	37	Textiles
38	Rubber and plastic	38	Rubber and plastic
40	Paper	40	Paper
41	Wood	41	Wood
42	Other manufacturing	42	Other manufacturing
43	Construction	43	Construction
44	Transportation and storage	44	Transportation and storage
45	Communication	45	Communication
46	Trade	46	Trade
47	Finance	47	Finance
48	Government and education		-
49	Health		-
50	Other services	50	Other services

Table 2.2 Patents at the Firm- and Industry-Level, 1910-1939

This table shows the total number of patents and the number of patents per firm from 1910 to 1939 using the firm- and industry-level patent data. Panel A lists the number of patents granted to each of the 29 sectors of use by the U.S. Patent Office from 1910 to 1939. Panel B reports the number of patents granted to the 107 U.S. manufacturing companies in the chemical, electrical, mechanical, and "other" industries. The industry-level patents are from "*Patents granted, by sector of use (SOU): 1940-1996", the Historical Statistics of the United States*. The firm-level patents are collected by Nicholas (2008).

Panel A Industry-Level Patents

SOU #	SOU name	# of patents	# of firms	# of patents per firm
29	Other machinery	2,634,993	38	69,342
13	Electrical industrial equipment	484,862	11	44,078
43	Construction	839,168	21	39,960
20	Transport-- motor vehicles	2,508,413	68	36,888
25	Fabricated metals	805,132	32	25,160
37	Textiles	1,189,887	50	23,798
1	Food	2,004,945	92	21,793
16	Chemicals	835,760	42	19,899
9	Mining	764,965	49	15,612
46	Trade	1,223,260	83	14,738
38	Rubber and plastic	184,450	13	14,188
45	Communication	137,640	11	12,513
44	Transportation and storage	1,621,480	131	12,378
26	Instruments	174,315	15	11,621
42	Other manufacturing	126,262	11	11,478
40	Paper	377,931	34	11,116
50	Other services	224,380	23	9,756
23	Metals - ferrous and nonferrous	820,362	90	9,115
22	Transport-- aircrafts	84,898	12	7,075
10	Electrical appliance	41,603	6	6,934
18	Petroleum	394,615	73	5,406
47	Finance	299,970	62	4,838
17	Drugs	58,398	13	4,492
11	Electrical lighting	110,666	31	3,570
12	Radio	23,536	7	3,362
21	Transport-- ships	9,070	3	3,023
28	Other office machinery	26,678	10	2,668
35	Food-- tobacco	47,907	22	2,178

Panel B Firm Level Patents

Sector #	Sector name	Total firm-level patents	# of firms	# of patents per firm
1	Chemistry	4,969	15	331.27
2	Electrical equipment	28,762	6	4,794
3	Mechanics	13,471	25	539
4	Others	8,155	61	134
	Total	55,357	107	

Table 2.3 Summary Statistics

Table 2.3 summarizes the variables used in the regression analysis at the industry-level and the firm-level. At the firm-level, *SSE, SSM* and *SSI* are the firm-specific, market, and industry return variation defined in eq. [1]-[5]. The methodology is based on Roll (1988) and Morck et al. (2000). The industry-level analogs are defined as the value-weighted averages of the firm-level measures. The patent measures at the industry-level are based on the number of patents granted to each of the 29 sector of use from 1910 to 1939. The patent measures at the firm-level calculated with the number of patents granted to each of the 107 manufacturing firms from 1910 to 1939, collected by Nicholas (2008). *PT10y* is the number of patents granted to a firm or an industry in the prior 10 years. *PTstock* is the patent stock of each firm or industry in the prior 10 years, estimated with a 15% annual depreciation rate. *PT1y* is the patent granted to a firm or an industry in the prior year. The firm-level *Log(MV)* is the log of the year-end market capitalizations of each firm. The industry-level market capitalization is the sum of the firm level year-end market value. Firm *Age* is the number of years since the first trading day is observed in the NYSE historical trading data published on the *New York Times since* 1851. Industry *Age* is the average age of firms in the industry. Industry *HHI* is the Herfindahl Index calculated with year-end firm market capitalizations, i.e. the sum of the squared share of firms' market capitalizations in an industry. All variables are winsorized at 1% level.

Panel A Industry-Level Data

Variable	*Obs.*	*Mean*	*Std. Dev.*	*5% Centile*	*Median*	*95% Centile*
Log(SSE)	508	-7.863	0.713	-9.058	-7.883	-6.517
Log(SSM)	508	-8.830	1.145	-10.720	-8.864	-7.033
Log(SSI)	508	-11.429	1.369	-14.018	-11.375	-9.236
Log(SSE/(SSM+SSI))	508	1.310	0.992	-0.325	1.330	2.939
PT10y	508	14,066	13,001	1,902	10,363	38,565
PTstock	508	7,600	6,985	1,014	5,578	20,617
PT1y	508	1,437	1,312	179	1,066	4,112
Log(MV)	508	13.077	1.570	10.105	13.167	15.252
Age	508	12.871	7.629	4.200	11.156	30.457
HHI	508	0.287	0.201	0.065	0.216	0.682

Panel B Firm-Level Data

Variable	*Obs.*	*Mean*	*Std. Dev.*	5% Centile	*Median*	*95% Centile*
Log(SSE)	1,670	-7.648	1.134	-9.361	-7.765	-5.597
Log(SSM)	1,670	-8.827	1.384	-11.122	-8.756	-6.745
Log(SSI)	1,670	-12.447	2.387	-17.040	-12.093	-9.256
Log(SSE/(SSM+SSI))	1,670	1.046	1.197	-0.999	1.069	3.113
PT10y	1,670	198	643	0	18	798
PTstock	1,670	113	364	0	10	468
PT1y	1,661	25	78	0	2	113
Log(MV)	1,670	10.000	1.948	6.705	9.952	13.177
Age	1,670	21.876	12.106	4	21	42

Table 2.4 Correlation Matrix

Table 2.4 lists the correlation coefficients of the variables used in the regression analysi at the industry- and firm-level. At the firm-level, *SSE, SSM* and *SSI* are the firm-specific, market, and industry return variation defined in eq. [1]-[5]. The industry-level analogs are defined as the value-weighted averages of the firm-level measures. *PT10y* is the number of patents granted to each firm or industry in the prior 10 years. *PTstock* is the patent stock of each firm or industry in the prior 10 years, estimated with a 15% annual depreciation rate. *PT1y* is the patent counts of each firm and industry in the prior year. *Log(MV)* is the log of the year-end market capitalizations of each firm. Industry market value is the sum of the firm level year-end market capitalizations. Firm *Age* is the number of years since the first trading day is observed in NYSE historical trading data published on the *New York Times since* 1851. Industry *Age* is the average age of firms in the industry. Industry *HHI* is the Herfindahl Index calculated with year-end firm market capitalizations, i.e. the sum of the squared share of firms' market capitalizations in an industry. All variables are winsorized at 1% level. *T-statisitcs* are reported in the parentheses. "*", "**", and "***" denote the significance at 10%, 5%, and 1% level, respectively.

Panel A Industry-Level Data

	Log(SSE)	*Log(SSM)*	*Log(SSI)*	*Log(SSE/(SSM+SSI))*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*	*Log(MV)*	*Age*	*HHI*
Log(SSE)	1.000									
Log(SSM)	0.608***	1.000								
	(17.23)									
Log(SSI)	0.543***	0.502***	1.000							
	(14.54)	(13.06)								
Log(SSE/(SSM+SSI))	0.047	-0.671***	-0.267***	1.000						
	(1.05)	(20.35)	(6.23)							
Log(PT10y)	0.002	0.125***	-0.164***	-0.077*	1.000					
	(0.05)	(2.83)	(3.74)	(1.73)						
Log(Ptstock)	-0.001	0.127***	-0.167***	-0.082*	0.999***	1.000				
	(0.01)	(2.89)	(3.82)	(1.84)	(586.06)					
Log(PT1y)	-0.003	0.125***	-0.179***	-0.080*	0.987***	0.991***	1.000			
	(0.08)	(2.83)	(4.09)	(1.81)	(137.33)	(169.67)				
Log(MV)	-0.578***	-0.076*	-0.141***	-0.343***	0.187***	0.196***	0.205***	1.000		
	(15.93)	(1.72)	(3.20)	(8.21)	(4.28)	(4.49)	(4.70)			
Age	-0.312***	-0.030	0.035	-0.220***	0.073	0.069	0.046	0.471***	1.000	
	(7.38)	(0.68)	(0.78)	(5.06)	(1.64)	(1.55)	(1.03)	(12.00)		
HHI	-0.103***	-0.057	-0.273***	-0.072	-0.069	-0.068	-0.059	-0.355***	-0.141***	1.000
	(2.33)	(1.29)	(6.38)	(1.62)	(1.56)	(1.53)	(1.33)	(8.55)	(3.21)	

Panel B Firm Level Data

	Log(SSE)	*Log(SSM)*	*Log(SSI)*	*Log(SSE/ SSM+SSI)*	*Firm log(PT10y)*	*Firm log(PTstock)*	*Firm log(PT1y)*	*Ind. log(PT10y)*	*Ind. log(Ptstock)*	*Ind. log(PT1y)*	*Log(MV)*	*Age*
Log(SSE)	1.000											
Log(SSM)	0.464***	1.000										
	(21.39)											
Log(SSI)	0.337***	0.312***	1.000									
	(14.62)	(13.43)										
Log(SSE/(SSM+SSI))	0.380***	-0.608***	-0.134***	1.000								
	(16.80)	(31.25)	(5.50)									
Firm log(PT10y)	-0.301***	0.099***	-0.061***	-0.367***	1.000							
	(12.88)	(4.08)	(2.50)	(16.12)								
Firm log(PTstock)	-0.306***	0.091***	-0.063***	-0.363***	0.994***	1.000						
	(13.10)	(3.72)	(2.59)	(15.89)	(368.59)							
Firm log(PT1y)	-0.311***	0.050***	-0.074***	-0.328***	0.884***	0.918***	1.000					
	(13.33)	(2.04)	(3.03)	(14.14)	(76.80)	(94.49)						
Ind. log(PT10y)	0.081***	-0.018***	0.140***	0.073***	-0.138***	-0.157***	-0.219***	1.000				
	(3.30)	(0.74)	(5.75)	(2.97)	(5.69)	(6.49)	(9.14)					
Ind. log(PTstock)	0.086***	-0.014	0.140***	0.072***	-0.141***	-0.159***	-0.221***	0.999***	1.000			
	(3.53)	(0.56)	(5.79)	(2.97)	(5.79)	(6.59)	(9.23)	(1003.67)				
Ind. log(PT1y)	0.115***	0.005	0.143***	0.078***	-0.160***	-0.178***	-0.228***	0.982***	0.988***	1.000		
	(4.72)	(0.19)	(5.90)	(3.16)	(6.62)	(7.36)	(9.55)	(214.23)	(259.34)			
Log(MV)	-0.827***	-0.234***	-0.241***	-0.489***	0.447***	0.459***	0.474***	-0.168***	-0.169***	-0.183***	1.000	
	(60.14)	(9.84)	(10.14)	(22.88)	(20.39)	(21.07)	(21.95)	(6.96)	(6.99)	(7.59)		
Age	-0.059***	0.160***	0.061***	-0.229***	0.111***	0.095***	0.075***	0.087***	0.086***	0.062***	0.082***	1.000
	(2.43)	(6.63)	(2.49)	(9.62)	(4.56)	(3.91)	(3.06)	(3.58)	(3.52)	(2.53)	(3.36)	

Table 2.5 Patents and Firm-specific Return Variation, Industry-level Regressions

This table reports the OLS regression results of industry-level firm-specific return variation on industry-level patent measures:

$$ln(SSE_{i,t}) = \alpha + \beta_1 Patent_{i,t} + \beta_2 \, log(MV_{i,t}) + \beta_3 \, log(Age_{i,t}) + \beta_4 HHI_{i,t} + \beta_5 \, ln(SSM_{i,t}) + \beta_6 \, ln(SSI_{i,t}) + Ind_i + Year_t + \epsilon_{i,t},$$

$$\Psi_{i,t} = \alpha + \beta_1 Patent_{i,t} + \beta_2 \, log(MV_{i,t}) + \beta_3 \, log(Age_{i,t}) + \beta_4 HHI_{i,t} + Ind_i + Year_t + \epsilon_{i,t},$$

Where industry-level $ln(SSE_{i,t})$ and $\Psi_{i,t}$ are the weighted average firm-level measures of absolute and relative firm-specific return variation of industry *i* in year *t*, as defined in eq. [1]-[5]. $Patent_{i,t}$ denotes one of the three patent measures: $log(PT10y_{i,t})$ is the log number of patents granted to industry *i* form year *t-1* to *t-10*, $log(PTstock_{i,t})$ is the log of patent stock of industry *i* in from year *t-1* to *t-10* estimated with 15% annual depreciation rate, and $log(PT1y_{i,t})$ is the log of the number of patent granted to industry *i* in year *t-1*. $Log(MV_{i,t})$ is the log of the industry market capitalizations of industry *i* in year *t*. $log(Age_{i,t})$ is the log of the average firm age of firms in industry *i* in year *t*, where firm age is defined as the number of years since the first stock price data of each firm is observed in the *New York Times*. $HHI_{i,t}$ is the Herfindahl index defined with industry MV. Panel A, B, and C show the regression results for the periods of 1921-1939, 1921-1928, and 1930-1939, respectively. Panel D shows the regressions using industry-specific return variation. Industry and year fixed effects are included in all the regressions. Standard errors are clustered by industry. "***", "**", and "*" denote significance at 1%, 5%, and 10% level.

Panel A 1921-1939

	Log(SSE)			Ψ=*Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Ind. patents	0.234	0.199	0.097	0.870**	0.751**	0.515**
	(0.96)	(0.86)	(0.48)	(2.56)	(2.47)	(2.13)
Log(MV)	-0.279***	-0.279***	-0.281***	-0.333***	-0.335***	-0.338***
	(-7.50)	(-7.56)	(-7.83)	(-4.36)	(-4.37)	(-4.34)
Log(Age)	-0.196**	-0.197**	-0.199**	-0.214	-0.216	-0.218
	(-2.11)	(-2.12)	(-2.18)	(-1.44)	(-1.46)	(-1.46)
HHI	-0.546*	-0.545*	-0.550*	-1.455***	-1.450***	-1.445***
	(-2.03)	(-2.04)	(-2.03)	(-3.69)	(-3.70)	(-3.64)
Log(SSM)	0.174***	0.174***	0.173***			
	(3.95)	(4.01)	(4.09)			
Log(SSI)	0.089***	0.089***	0.088***			
	(5.02)	(5.01)	(4.97)			
Adj. R^2	0.886	0.886	0.886	0.744	0.743	0.742
N	508	508	508	510	510	510

Panel B 1921-1928

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Ind. patents	-1.078	-0.834	-0.303	-1.619	-1.973	-1.179
	(-1.08)	(-1.05)	(-0.96)	(-0.75)	(-1.04)	(-1.24)
Log(MV)	-0.198***	-0.198***	-0.192***	-0.168	-0.175	-0.159
	(-2.80)	(-2.80)	(-2.80)	(-1.10)	(-1.11)	(-1.08)
Log(Age)	-0.111	-0.108	-0.108	-0.601*	-0.610*	-0.623*
	(-0.63)	(-0.61)	(-0.60)	(-1.85)	(-1.88)	(-1.95)
HHI	0.297	0.274	0.259	-0.677	-0.696	-0.756
	(0.78)	(0.71)	(0.67)	(-1.35)	(-1.38)	(-1.45)
Log(SSM)	0.173***	0.165**	0.158**			
	(2.82)	(2.57)	(2.39)			
Log(SSI)	0.078***	0.079***	0.077***			
	(2.94)	(2.97)	(2.84)			
Adj. R^2	0.839	0.838	0.837	0.545	0.550	0.554
N	201	201	201	203	203	203

Panel C 1930-1939

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Ind. patents	0.603***	0.516***	0.165	2.000***	1.827***	1.256***
	(3.22)	(2.91)	(0.83)	(3.72)	(4.05)	(3.31)
Log(MV)	-0.489***	-0.490***	-0.494***	-0.553***	-0.551***	-0.533***
	(-10.09)	(-10.22)	(-9.84)	(-4.70)	(-4.69)	(-4.92)
Log(Age)	-0.199	-0.204	-0.235	-0.363	-0.369	-0.437
	(-1.43)	(-1.45)	(-1.61)	(-0.84)	(-0.82)	(-0.92)
HHI	0.129	0.121	-0.024	-0.667	-0.639	-0.711
	(0.51)	(0.48)	(-0.09)	(-0.93)	(-0.93)	(-1.01)
Log(SSM)	0.102**	0.100*	0.082			
	(2.12)	(2.01)	(1.47)			
Log(SSI)	0.057***	0.055***	0.048**			
	(2.92)	(2.86)	(2.60)			
Adj. R^2	0.948	0.948	0.946	0.769	0.768	0.764
N	279	279	279	279	279	279

Panel D Industry-specific Return Variation

Dep. Var.	*Log(SSI)*			*Log(SSI/(SSM+SSE))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
1921-1928						
Ind. patents	4.129	4.061*	1.725*	4.015	4.013*	1.691*
	(1.42)	(1.90)	(1.74)	(1.40)	(1.87)	(1.73)
Log(MV)	0.296	0.305	0.274	0.277	0.284	0.254
	(1.26)	(1.33)	(1.21)	(1.27)	(1.32)	(1.22)
Age	0.524	0.523	0.536	0.531	0.533	0.551
	(1.24)	(1.22)	(1.20)	(1.22)	(1.21)	(1.21)
HHI	1.934*	2.002*	2.124*	1.816*	1.874*	1.970*
	(1.83)	(1.95)	(2.01)	(1.78)	(1.90)	(1.95)
Log(SSM)	0.112	0.133	0.178			
	(0.53)	(0.65)	(0.85)			
Log(SSE)	0.862***	0.863***	0.841***			
	(3.31)	(3.42)	(3.28)			
Adj. R^2	0.608	0.613	0.609	0.487	0.493	0.487
1930-1939						
Ind. patents	-2.861***	-2.579***	-1.226	-2.273**	-2.061**	-1.028
	(-3.12)	(-2.92)	(-1.55)	(-2.40)	(-2.30)	(-1.49)
Log(MV)	-0.025	-0.039	-0.114	-0.234	-0.236	-0.241
	(-0.07)	(-0.11)	(-0.33)	(-0.76)	(-0.76)	(-0.80)
Age	-0.562	-0.560	-0.475	-0.473	-0.466	-0.401
	(-1.08)	(-1.04)	(-0.80)	(-0.81)	(-0.79)	(-0.64)
HHI	0.070	0.045	0.528	0.244	0.220	0.575
	(0.08)	(0.05)	(0.44)	(0.25)	(0.22)	(0.50)
Log(SSM)	0.060	0.060	0.135			
	(0.42)	(0.42)	(0.82)			
Log(SSE)	0.931***	0.908***	0.783**			
	(3.27)	(3.18)	(2.74)			
Adj. R^2	0.734	0.733	0.722	0.582	0.581	0.570

Table 2.6 Patents and Firm-specific Return Variation, Firm Level Regressions

This table reports the OLS regression results of firm-level firm-specific return variation on firm and industry-rival patent measures using a subsample of 107 manufacturing firms. The firm-level regression equations are:

$$ln(SSE_{j,t}) = \alpha + \beta_1 Firm\ patent_{j,t} + \beta_2 Ind\ patents_{j,t} + \beta_3 log(MV_{j,t}) + \beta_4 \log(Age_{j,t}) + \beta_5\, ln(SSM_{j,t}) + \beta_6\, ln(SSI_{j,t}) + Ind_i + Year_t + \epsilon_{j,t},$$

$$\Psi_{j,t} = \alpha + \beta_1 Firm\ patent_{j,t} + \beta_2 Ind\ patents_{j,t} + \beta_3 log(MV_{j,t}) + \beta_4 \log(Age_{j,t}) + Ind_i + Year_t + \epsilon_{j,t}\ ,$$

where the firm-level $ln(SSE_{i,t})$ and $\Psi_{i,t}$ are the measures of absolute and relative firm-specific return variation of firm j in year t, as defined in eq. [3]-[5]. *Firm patent*$_{j,t}$ and *Ind Patent*$_{j,t}$ are the firm-level and industry-rival patents proxied by one of the three measures: $log(PT10y_{i,t})$ is the log number of patents granted to industry i form year $t-1$ to $t-10$, $log(PTstock_{i,t})$ is the log of patent stock of industry i in from year $t-1$ to $t-10$, estimated with a 15% annual depreciation rate, and $log(PT1y_{i,t})$ is the log of the number of patent granted to industry i in year $t-1$. *Firm patent*$_{j,t}$ is calculated with the patents of firm j in year t, and *Ind Patent*$_{j,t}$ is calculated with the industry-level patents minus the patents of firm j in year t. $Log(MV_{j,t})$ is the log of the year-end market capitalizations of firm j in year t. $log(Age_{j,t})$ is the log number of years since the first stock price data of each firm is observed in the *New York Times*. Panel A, B, and C show the regression results for the periods of 1921-1939, 1921-1928, and 1930-1939, respectively. Industry and year fixed effects are included in all the regressions. Standard errors are clustered by industry. "***", "**", and "*" denote significance at 1%, 5%, and 10% level.

Panel A 1921-1939

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Firm patents	0.008	0.013	0.031	-0.056	-0.055	-0.035
	(0.87)	(1.05)	(1.46)	(-1.33)	(-1.16)	(-0.62)
Ind. Patents	-0.242*	-0.241**	-0.196*	0.285	0.275	0.185
	(-3.05)	(-3.60)	(-2.97)	(0.94)	(1.00)	(0.76)
Log(MV)	-0.457***	-0.459***	-0.465***	-0.301***	-0.302***	-0.314***
	(-17.70)	(-18.81)	(-22.48)	(-13.15)	(-13.06)	(-14.36)
Age	-0.065	-0.064	-0.060	-0.149	-0.152	-0.153
	(-1.89)	(-1.92)	(-1.93)	(-1.36)	(-1.38)	(-1.48)
Log(SSM)	0.031***	0.031***	0.030***			
	(6.59)	(6.49)	(6.49)			
Log(SSI)	0.119***	0.118***	0.117***			
	(7.32)	(7.32)	(7.02)			
Adj. R^2	0.798	0.798	0.799	0.479	0.478	0.473
N	1,695	1,695	1,686	1,691	1,691	1,682

Panel B 1921-1928

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Firm patents	-0.013	-0.011	0.002	-0.021	-0.024	-0.042
	(-1.62)	(-0.96)	(0.10)	(-0.56)	(-0.56)	(-0.67)
Ind. Patents	-0.531**	-0.527**	-0.333**	-0.591**	-0.600***	-0.657***
	(-4.87)	(-5.08)	(-4.26)	(-4.36)	(-6.08)	(-8.28)
Log(MV)	-0.444***	-0.446***	-0.453***	-0.232**	-0.232**	-0.228**
	(-12.20)	(-12.46)	(-12.66)	(-4.72)	(-4.76)	(-4.88)
Age	-0.094*	-0.096*	-0.092*	-0.161	-0.162	-0.162
	(-2.44)	(-2.50)	(-2.64)	(-2.22)	(-2.20)	(-2.24)
Log(SSM)	0.052***	0.053***	0.052***			
	(9.10)	(9.14)	(8.89)			
Log(SSI)	0.098*	0.098*	0.096*			
	(3.02)	(2.98)	(2.86)			
Adj. R^2	0.675	0.675	0.673	0.206	0.206	0.207
N	668	668	660	666	666	658

Panel C 1930-1939

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Firm patents	0.025	0.031	0.044	-0.084	-0.080	-0.036
	(1.96)	(1.94)	(1.70)	(-1.43)	(-1.21)	(-0.49)
Ind. Patents	0.408***	0.390***	0.407***	1.523***	1.519***	1.469***
	(19.53)	(27.21)	(11.47)	(13.08)	(13.37)	(39.28)
Log(MV)	-0.478***	-0.481***	-0.483***	-0.315***	-0.318***	-0.340***
	(-23.06)	(-25.33)	(-34.52)	(-11.19)	(-10.80)	(-12.32)
Age	-0.011	-0.009	-0.011	-0.207	-0.206	-0.189
	(-0.19)	(-0.16)	(-0.20)	(-1.21)	(-1.22)	(-1.27)
Log(SSM)	0.016**	0.016**	0.016**			
	(5.20)	(5.31)	(5.80)			
Log(SSI)	0.113***	0.113***	0.115***			
	(7.83)	(7.77)	(8.98)			
Adj. R^2	0.856	0.856	0.857	0.508	0.506	0.498
N	929	929	928	927	927	926

Table 2.7 Patents and Firm-specific Return Variation in High-Tech Sectors, Firm-level Regressions

This table reports the OLS regression results of firm-specific return variation on firm-level and industry-rival patents, including the interaction between firm-level patents and the following three high-tech sector dummies, chemical, electrical, and mechanical sectors. *ln($SSE_{i,t}$)* and $\Psi_{i,t}$ are the measures of absolute and relative firm-specific return variation of firm *j* in year *t*, as defined in eq. [3]-[5]. *Firm patent$_{j,t}$* and *Ind Patent$_{j,t}$* are the firm-level and industry-rival patents proxied by one of the three measures: *log($PT10y_{i,t}$)* is the log number of patents granted to industry *i* form year *t-1* to *t-10*, *log($PTstock_{i,t}$)* is the log of patent stock of industry *i* in from year *t-1* to *t-10* estimated with 15% annual depreciation rate, and *log($PT1y_{i,t}$)* is the log of the number of patent granted to industry *i* in year *t-1*. *Firm patent$_{j,t}$* is calculated with the patents of firm *j* in year *t*, and *Ind Patent$_{j,t}$* is calculated with the industry-level patents minus the patents of firm *j* in year *t*. *Log($MV_{j,t}$)* is the log of the year-end market capitalizations of firm *j* in year *t*. *log($Age_{j,t}$)* is the log number of years since the first stock price data of each firm is observed in the *New York Times*. Panel A, B, and C show the regression results for the periods of 1921-1939, 1921-1928, and 1930-1939, respectively. Industry and year fixed effects are included in all the regressions. Standard errors are clustered by industry. "***", "**", and "*" denote significance at 1%, 5%, and 10% level.

Panel A 1921-1939

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Firm patents	-0.012	-0.012	-0.007	-0.104***	-0.113***	-0.123***
	(-0.95)	(-0.89)	(-0.56)	(-8.88)	(-8.68)	(-9.20)
Firm patents Chem*	-0.001	-0.002	-0.002	0.122***	0.134***	0.162***
	(-0.82)	(-0.97)	(-0.84)	(13.06)	(14.18)	(16.06)
Firm patents Elec*	0.030**	0.034**	0.035**	0.039	0.055*	0.096**
	(3.38)	(3.95)	(4.32)	(2.15)	(2.91)	(4.53)
Firm patents Mech*	0.077***	0.088***	0.111***	0.108***	0.127***	0.170***
	(14.45)	(16.58)	(18.04)	(6.74)	(7.62)	(9.59)
Ind. Patents	-0.162	-0.162*	-0.139	0.199	0.200	0.144
	(-1.89)	(-2.39)	(-2.29)	(0.46)	(0.51)	(0.47)
Log(MV)	-0.456***	-0.458***	-0.464***	-0.300***	-0.302***	-0.317***
	(-16.44)	(-17.44)	(-20.85)	(-12.86)	(-12.90)	(-14.28)
Age	-0.077	-0.078	-0.075	-0.171	-0.176	-0.179
	(-2.10)	(-2.14)	(-2.16)	(-1.50)	(-1.53)	(-1.61)
Log(SSM)	0.031***	0.031***	0.031***			
	(6.87)	(6.83)	(7.03)			
Log(SSI)	0.121***	0.121***	0.118***			
	(7.47)	(7.61)	(7.71)			
Adj. R^2	0.800	0.801	0.803	0.485	0.485	0.481
N	1,695	1,695	1,686	1,691	1,691	1,682

Panel B 1921-1928

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Firm patents	-0.034*	-0.036*	-0.037	-0.079**	-0.093**	-0.161***
	(-2.47)	(-2.54)	(-2.23)	(-4.88)	(-5.57)	(-9.42)
Firm patents Chem*	-0.005	-0.007	-0.011	0.192***	0.213***	0.310***
	(-0.47)	(-0.56)	(-0.90)	(23.07)	(24.35)	(26.07)
Firm patents Elec*	0.077***	0.083***	0.098***	0.128***	0.149***	0.254***
	(15.82)	(19.66)	(17.36)	(9.79)	(11.72)	(16.05)
Firm patents Mech*	0.062***	0.075***	0.095***	0.094***	0.104***	0.146***
	(6.97)	(9.44)	(16.55)	(14.51)	(17.43)	(20.92)
Ind. Patents	-0.473**	-0.466**	-0.274*	-0.768**	-0.757**	-0.744***
	(-4.63)	(-5.02)	(-2.53)	(-3.88)	(-4.66)	(-7.68)
Log(MV)	-0.443***	-0.445***	-0.454***	-0.234**	-0.234**	-0.236**
	(-11.41)	(-11.73)	(-12.43)	(-4.75)	(-4.82)	(-5.05)
Age	-0.099*	-0.101*	-0.101*	-0.179*	-0.180*	-0.186*
	(-2.38)	(-2.43)	(-2.60)	(-2.65)	(-2.59)	(-2.63)
Log(SSM)	0.052***	0.051***	0.050***			
	(10.58)	(10.88)	(11.41)			
Log(SSI)	0.101*	0.101*	0.099*			
	(3.05)	(3.06)	(3.12)			
Adj. R^2	0.677	0.677	0.676	0.222	0.223	0.233
N	668	668	660	666	666	658

Panel C 1930-1939

	Log(SSE)			*Ψ=Log(SSE/(SSM+SSI))*		
	Log(PT10y)	*Log(PTstock)*	*Log(PT1y)*	*Log(PT10y)*	*Log(PTstock)*	*Log(PT1y)*
Firm patents	0.013	0.015	0.016	-0.142***	-0.150**	-0.121**
	(1.34)	(1.52)	(1.78)	(-5.97)	(-5.60)	(-4.70)
Firm patents Chem*	-0.010	-0.011	-0.011	0.120***	0.124***	0.103**
	(-1.41)	(-1.45)	(-1.14)	(9.84)	(9.51)	(5.70)
Firm patents Elec*	-0.027	-0.026	-0.023	0.003	0.018	0.021
	(-1.93)	(-1.86)	(-1.85)	(0.11)	(0.56)	(0.63)
Firm patents Mech*	0.070***	0.078***	0.099***	0.155**	0.178**	0.201***
	(7.45)	(7.68)	(9.15)	(4.97)	(5.41)	(6.34)
Ind. Patents	0.459***	0.430***	0.435***	1.516***	1.491***	1.445***
	(9.74)	(11.11)	(8.46)	(7.09)	(7.78)	(42.97)
Log(MV)	-0.476***	-0.478***	-0.481***	-0.314***	-0.317***	-0.340***
	(-21.37)	(-23.08)	(-29.54)	(-11.29)	(-11.07)	(-12.92)
Age	-0.026	-0.026	-0.029	-0.249	-0.253	-0.229
	(-0.50)	(-0.51)	(-0.57)	(-1.36)	(-1.39)	(-1.39)
Log(SSM)	0.016**	0.016**	0.017***			
	(5.21)	(5.67)	(6.68)			
Log(SSI)	0.116***	0.116***	0.118***			
	(8.50)	(8.69)	(10.01)			
Adj. R^2	0.858	0.859	0.860	0.519	0.517	0.508
N	929	929	928	927	927	926

Table 2.8 Labor Productivity and Firm-specific Return Variation

Table 2.8 lists the regression results of industry-level firm-specific return variation on labor productivity growth over 13 manufacturing industries. The two measures of labor productivity are the ratio of value-added and employment, *VA/Wage*, and the ratio of value-added and wages, *VA/Employ*. The growth of labor productivity is the log difference of labor productivity between year *t* and *t-2* (the labor series are bi-annual). The annual series firm-specific return variation, *log(SSE)* and *log(SSE/(SSM+SSI)),* and the controlling variables (*MV, HHI*, and *Age*) are matched with the biannual labor productivity series by using the average of year *t* and *t-1* of the annual series to match the survey years *t*. Industry and year fixed effects are included in all the regressions. Standard errors are clustered by industry. T-statistics are reported in the parentheses. "*", "**", and "***" denote the significance at 10%, 5%, and 1% level, respectively.

	Log(SSE)		*Log(SSE/(SSM+SSI))*	
	ΔVA/Employ	*ΔVA/Wage*	*ΔVA/Employ*	*ΔVA/Wage*
labor productivity growth	0.346*	0.275	1.485**	1.804*
	(1.82)	(0.88)	(2.94)	(2.10)
Log(MV)	-0.194*	-0.231**	-0.192	-0.303**
	(-2.08)	(-2.53)	(-1.78)	(-2.43)
Age	-0.576*	-0.390	-1.107**	-0.639
	(-2.13)	(-1.78)	(-2.84)	(-1.36)
HHI	0.178	-0.467	0.360	0.030
	(0.34)	(-0.95)	(0.58)	(0.04)
Log(SSM)	0.180	0.068		
	(1.22)	(0.95)		
Log(SSI)	0.073*	0.057		
	(2.12)	(1.76)		
Adj. R^2	0.862	0.835	0.819	0.740
N	87	108	87	108

Table 2.9 Research Staff and Firm-specific Return Variation

Table 2.9 reports the regression results of industry-level firm-specific return variation on the number or percentage of research staff over the 16 manufacturing sectors. *Log(RDstaff)* and *RDstaff%* are the log number and the percentage of research personnel, respectively. The annual series firm-specific return variation, *log(SSE)* and *log(SSE/(SSM+SSI)),* and the controlling variables (*MV, HHI*, and *Age*) are matched with the research staff series by using the average from the year after the prior survey year to the current survey year to match current survey year. The research staff survey is conducted by the National Research Council in 1921, 1927, and 1933. Industry and year fixed effects are included in all the regressions. Standard errors are clustered by industry. T-statistics are reported in the parentheses. "*", "**", and "***" denote the significance at 10%, 5%, and 1% level, respectively.

	Log(SSE)		*Log(SSE/(SSM+SSI))*	
	Log(RDstaff)	*RDstaff%*	*Log(RDstaff)*	*RDstaff%*
Research Staff	0.247**	1.548	0.412	6.581**
	(2.27)	(1.10)	(1.16)	(2.26)
Log(MV)	-0.156	-0.175	-0.399	-0.518*
	(-1.38)	(-1.12)	(-1.43)	(-1.88)
Age	-0.510**	-0.525*	-0.598	-0.971
	(-2.57)	(-2.01)	(-0.88)	(-1.49)
HHI	0.441	-0.173	3.657	2.618
	(0.43)	(-0.17)	(0.92)	(0.73)
Log(SSM)	0.164	0.149		
	(1.21)	(1.02)		
Log(SSI)	-0.024	0.026		
	(-0.35)	(0.34)		
Adj. R^2	0.910	0.886	0.781	0.806
N	40	40	40	40

Table 2.10 Electrification and Firm-specific Return Variation

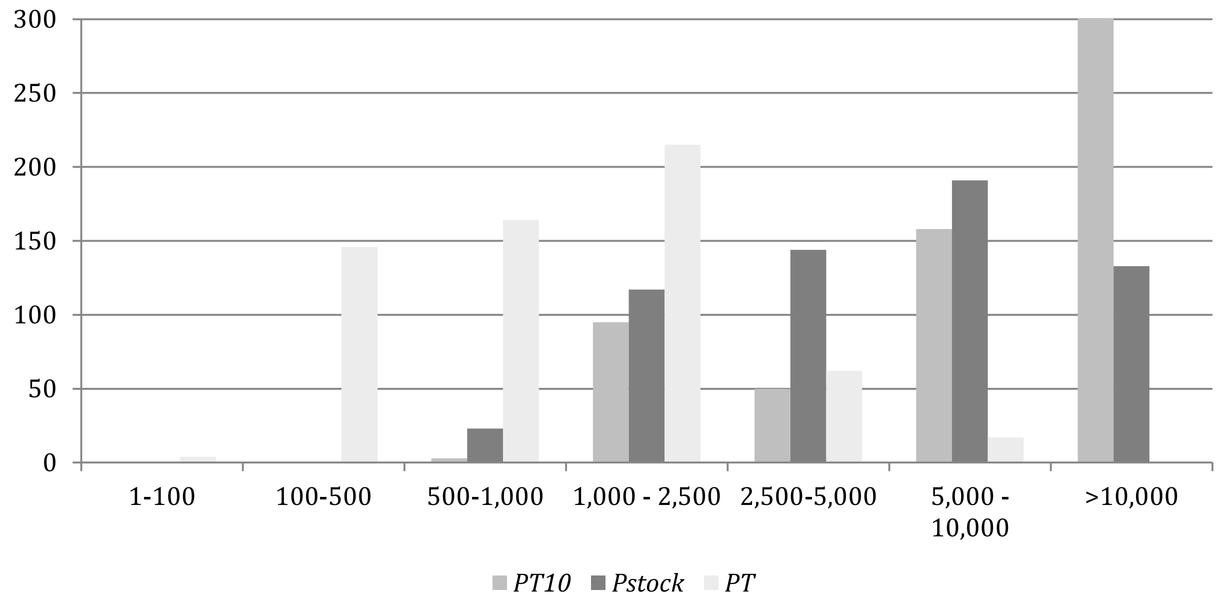
This table shows the regression results of industry-level firm-specific return variation on the log of electricity usage over 14 manufacturing sectors in 1929 and 1939. *Electrification* is defined as the log of the electricity horsepower consumed by each industry. *1920s* is a dummy variable indicate the decade the 1920s. The annual series of firm-specific return variation, *log(SSE)* and *log(SSE/(SSM+SSI)),* and the controlling variables (*MV, HHI*, and *Age*) from 1921 to 1929 are matched with the 1929 electrification, and the annual series from 1930 to 1939 are matched with the 1939 electrification. Standard errors are clustered by industry. T-statistics are reported in the parentheses. "*", "**", and "***" denote the significance at 10%, 5%, and 1% level, respectively.

	Log(SSE)		*Log(SSE/(SSM+SSI))*	
Electrification	0.134***	0.188***	-0.125	-0.202
	(-4.16)	(-3.68)	(-0.98)	(-1.38)
*Electrification*1920s*		-0.102		0.160**
		(-1.37)		(-2.37)
Log(MV)	-0.246***	-0.248***	-0.215	-0.237*
	(-4.87)	(-5.87)	(-1.39)	(-1.93)
Age	-0.043	-0.043	-0.750	-0.751
	(-0.14)	(-0.14)	(-1.13)	(-1.13)
HHI	0.073	0.222	0.758	0.846
	(-0.23)	(-0.66)	(-0.57)	(-0.74)
Log(SSM)	(0.05)	0.045		
	(-0.55)	(-0.60)		
Log(SSI)	0.113*	0.104*		
	(-2.11)	(-1.83)		
Adj. R^2	0.727	0.743	0.477	0.472
N	28	28	28	28

Figure 2.1 Firm and Industry Level Patents Distribution

This figure shows the histograms of the firm- and industry-level patent measures. At the industry-level, the patents measures are defined with the number of patents granted to each of the 29 sectors of use from 1910 to 1939. At the firm-level, the patent measures are calculated with the patents granted to a subsample of 107 public firms. *PT10y* is the log of the total patents in the prior 10 years. *PTstock* is the log of the patent stock in the prior 10 years estimated with a 15% annual depreciation rate. *PT1y* is the log of the patents granted in the prior year.

Panel A Industry Level Patents



Panel B Firm Level Patents

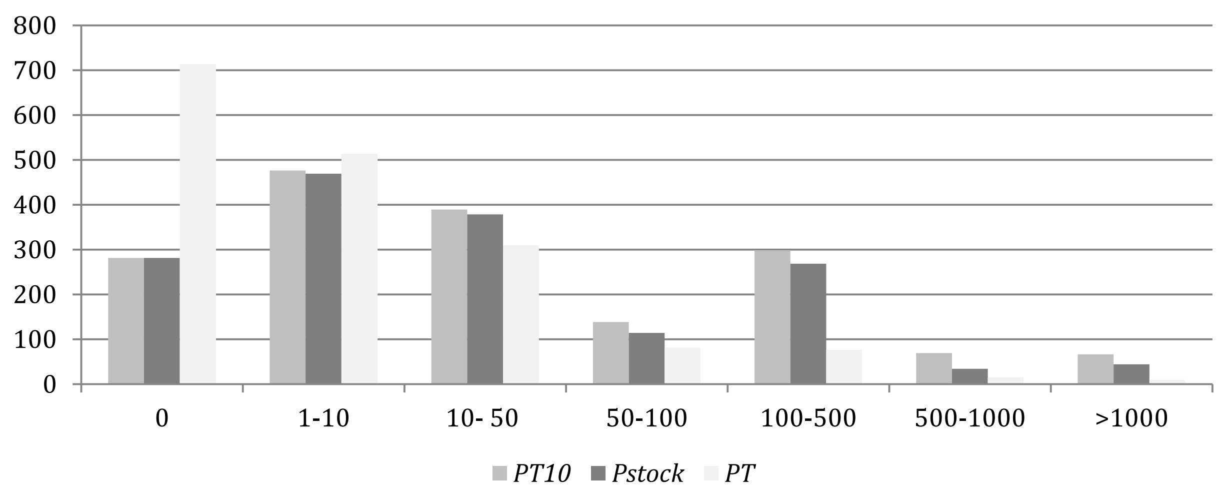
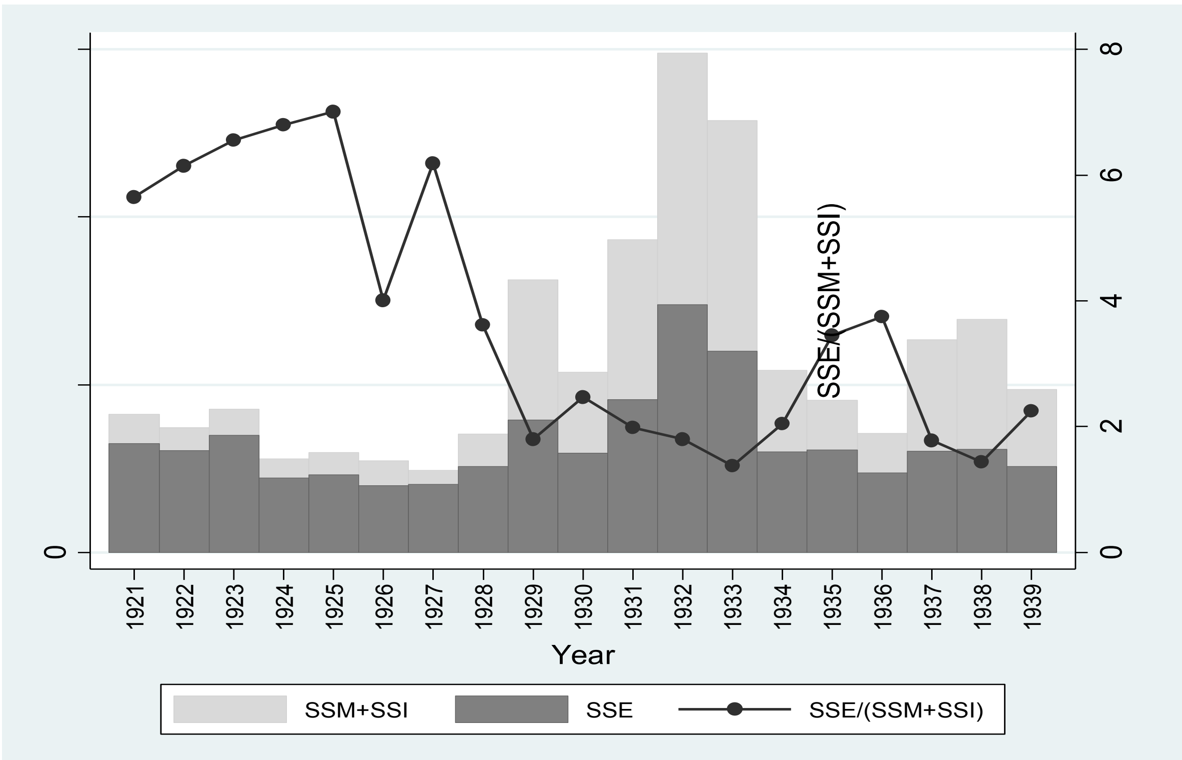


Figure 2.2 Economy Level Mean Firm-specific Return Variation, 1921-1939

This figure plots the value-weighted averages of firm-specific return variation *SSE*, systematic return variation *SSM*, and their ratio *SSE/SSM* across all NYSE firms from 1921 to 1939. At the-firm level, *SSE* and *SSM* are the sum of the squared residuals and the model sum of squares as defined by eq. [1] to [4], respectively.



Chapter 3 Stock Return Comovement and Market Run-ups

3.1 Introduction

We construct a profitable trading strategy using information about past market-wide stock return comovement. Based on Morck, Yeung, and Yu (2000), we define a directional comovement measure as the proportion of moving stocks that move up. A prior high upward comovement dummy is set to one if the upward comovement in a prior window (e.g. one week) is above a comovement threshold (e.g. the 4^{th} decile of the historical upward comovement distribution). A trading strategy of investing in the market index when prior upward comovement is high and in three-month U.S. Treasury bills otherwise earns economically and statistically significant positive abnormal returns relative to simply holding the market.

In a frictionless market with rational investors, the source of comovement in asset returns is correlated fundamentals. The proportion of stocks moving up together should fluctuate randomly if fundamental news arrives in a random manner. In markets with frictions or irrational investors, comovement in asset returns could be driven by correlated trading behavior. Shiller (1984), Delong, Shleifer, Summers and Waldman (1990), and others analyze how noise traders might move asset prices. In their models, the correlated demand of noise trades elevates systematic risk (comovement), which that deters arbitrage, keeping prices from converging to fundamentals. In the information cascades model of Bikhchandani, Hirshleifer, and Welch (1992, 1998), investors imitate the trades of other investors regardless of their own private signals. It is rational to imitate when investors believe that their private information is uncertain and that others are better informed. The synchronous demand of these investors could then cause comovement in asset returns. In their style investment model, Barberis and Shleifer (2003) argue that many investors allocate funds across groups of assets (styles), rather than individual assets, and that their synchronous trading behavior could then cause comovement in asset returns within the style. If stocks in general sometimes gain salience as a style of assets in this sense, the synchronized demand of return-chasing investors might induce stock market-wide episodes of this sort, which would be apparent as periods of unusually high stock price comovement.

These theories all suggest a similar pattern of comovement in stock returns. At the start of a cycle, high returns across a definable asset class create a degree of comovement, though this need not be large. In the next stage, these high returns attract return-chasing investors, whose buying pushing prices higher, drawing in yet more return-chasing investors. Investors might chase returns for either behavioral reasons (Shiller 1984; Delong, Shleifer, Summers and Waldman 1990; Barberis and Shleifer 2003) or rational reasons in a world with costly information (Bikhchandani, Hirshleifer, and Welch 1992, 1998). This upward positive feedback cycle might persist, generating a prolonged period of superior performance across the asset class. We hypothesize that this run-up stage is characterized by high upward stock return comovement driven at least partly, and perhaps largely, by the correlated trading of return-chasing investors. Upward comovement ends when bad news interrupts this cycle, and investors begin selling. Stocks adjust idiosyncratically as informed arbitrageurs regain dominance in price setting (Bris, Goetzmann, and Zhu. 2007).

We expect that a comovement-based investment strategy of buying a market index when the upward comovement becomes high and selling the market index when the upward comovement ends might profitably exploit such periods of large fund inflows. We find that the strategy of using upward comovement in the prior week as the signal to be in the NYSE value-weighted index, as opposed to being in three-month T-bills, generates an annualized alpha of 6.42% from 1954 to 2014. This strategy thus outperforms the market portfolio.

The strategy also outperforms an index return-based benchmark strategy. The idea behind this alternative benchmark strategy is that daily market index return should fluctuate around zero if moved only by fundamentals news, and that a period of clustered positive daily index returns might reflect large fund inflows lifting the market. This is thus an approach to detecting fund inflows without using stock-level co-movement data. On each trading day *t*, we count the number of positive returns in a prior period (such as one week) and use binomial tests to see if the number of trading days with positive index returns is higher than expected. This strategy invests in the market index after statistically significant runs of positive returns (periods of large fund inflows), and invests in the T-bills otherwise. We find that our comovement-based strategy also significantly outperforms this index return-based benchmark strategy by a statistically and economically significant 2.27% annualized alpha.

Our findings are highly robust. Alternative historical comovement thresholds, such as the 3rd or 6th historical comovement deciles, generate qualitatively similar results. So do investment strategies using an equal-weighted market index rather than a value-weighted index. The technique is also valid outside the NYSE sample we initially investigate. Repeating the exercise using stocks listed on NASDAQ or on the Tokyo Stock Exchange also yields positive and significant alphas.

The rest of the paper is organized as the follows. Section 3.2 reviews the literature. Section 3.3 describes the data describes the data and comovement-based strategy. Section 3.4 compares the risk-adjusted performance between the comovement-based strategy and the value-weighted

NYSE index using the Sharpe ratio and the asset pricing models. Section 3.5 conducts robustness checks. Section 3.6 investigates the performance of the comovement-strategy in NASDAQ and TSE. Section 3.7 concludes.

3.2 Literature Review

Our study relates to the literature on correlated trading behaviors. The first view attributes correlated trading to correlated fundamentals. Price only reflects the present value of the future cash flows (Grossman and Stiglitz, 1976; Diamond and Verrecchia, 1981). If fundamental news arrives randomly, market-wide upward comovement would fluctuate around a half.

An alternative view argues that correlated trading can be caused by non-fundamental reasons. Shleifer and Summers (1990) argue that noise traders can affect asset prices because of limited arbitrage. Arbitrageurs face the risk that mispricing worsens in the short run, yielding temporary negative returns (De Long et al., 1990a; Shleifer and Vishny, 1997). This risk matters especially for professional fund managers, because temporary losses may lead fund investors to withdraw. Arbitrageurs also face other type of risks such as potential high transaction costs, difficulty to set up a perfect hedge, and incorrect information. Empirical studies find evidence in support of correlated noise trading affecting asset prices. Dorn, Huberman, and Sengmueller (2008) document correlated trading among over 37,000 clients of a large German discount broker. Barber, Odean, and Zhu (2009) report that trading of clients at a U.S. retail broker are highly correlated and persistent. Using a large sample of retail investor transaction data, Kumar and Lee (2006) provide direct evidence of the association between correlated trading of individual investors and stock return comovement for stocks with high retail concentration.

The information cascades model by Bikhchandani, Hirshleifer, and Welch (1992, 1998) analyze correlated trading in a rational framework. Information cascades arise when a rational investor follows the trades of his predecessors despite of his private information. This could occur because investors have imperfect private information or believe that others are better quality information. Recent empirical studies find evidence of institutional investors following each other's trades over periods of time. Wermers (1999) finds higher level of herding in of small stocks and in growth-oriented mutual funds. Sias (2004) finds that institutional demand is strongly correlated over time. Using daily trade data, Christoffersen and Tang (2010) find patterns of institutional trading consistent with the information cascades models at short horizons. As new information becomes public or highly informed traders start to trade against the herd, information cascades may only be evidence in short term. Their study also helps to explain the less significant correlated trading in pension funds documented by Lakonishok, Shleifer, and Vishny (1992) using quarterly data.

Our findings that high upward comovement predicts high future returns in short windows are consistent with both the limited arbitrage theory and the information cascades model. These theories suggest that the coordinated behavior among investors can affect asset prices. The upward comovement measure detects periods of investors following each other into stock markets, due to rational or behavioral reasons, and the profit opportunity is not fully explored by arbitrageurs.

The style investment model by Barberis and Shleifer (2003) explains correlated trading by investor habitats. They argue that investors allocate funds across asset groups (or styles) rather than individual assets. Styles can be broad groupings of assets, such as stocks versus bonds, or narrower categories, such as income, growth, industry, and size stocks. Their model predicts that style investing generates excess comovement in returns unrelated to comovement in cash flows among stocks within a style. Excess comovement across assets has been shown to be related to index affiliation (Barberis, Shleifer, and Wurgler, 2005; Greenwood, 2008), correlated trading by retail investors (Kumar and Lee, 2006) and institutional investors (Sun, 2008; Anton and Polk, 2014), location of trade (Froot and Dabora, 1999), and common banking networks (Grullon, Underwood, and Weston, 2014).

Investment strategies based on past style returns are found profitable. For example, Moskowitz and Grinblatt (1999) and Asness et al. (1997) find that momentum strategies based on industry and country portfolios are profitable. Chen and De Bondt (2004) find evidence of style momentum in the S&P 500 index. Teo and Woo (2004) find evidence of momentum and reversals with mutual fund investment styles based on characteristics such as size, value/growth, and industries. Wahal and Yavuz (2013) find short-term return momentum and long-term reversal in multiple investment styles. This paper focus on “stocks”, an asset class that has repeatedly gone in and out of style, and investigates the time series change of (instead of cross-sectional difference in) comovement and its ability to predict future returns. Our finding that the comovement-based strategy outperforms an index return-based strategy suggests that the degree to which individual assets within a style comoves together provide additional information about future style returns.

Our finding that investors are able to ride the waves of market run-ups and profit from doing so is consistent with the theories on the limits of arbitrage. First, arbitrageurs may be reluctant to trade against mispricing. Shiller, Fisher, and Friedman (1984) and Campbell and Kyle (1987) show that aversion to fundamental risk can limit arbitrage. In De Long, Shleifer, Summers, and Waldmann (1990a), short horizons deter arbitrageurs from betting against noise traders because mispricing may deepen further before being corrected. Shleifer and Vishny (1997) argue that professional fund managers may forgo profitable long-run arbitrage opportunities because temporary losses would lead fund investors to withdraw.

Second, under certain circumstances, it can be optimal for arbitrageurs to follow the herd. Kindleberger (1978) notices that stock prices are driven up by insiders who sell to outsiders at market peak. In the model of De Long, Shleifer, Summers, and Waldmann (1990b), arbitrageurs push prices higher than fundamental values, anticipating that positive feedback traders will buy in response to the price rise. Though eventually arbitrages sell out and bring prices back to fundamentals, in the short run they profit by feeding the trend. Abreu and Brunnermeier (2002) model arbitrageurs who cannot individually correct mispricing and show that their lack of coordination causes them to trade with rather than against the noise traders. Empirical studies provide evidence in support of these theories. For example, using intra-day and daily data, Griffin, Harris, and Topaloglu (2003) document the trend-chasing behavior of institutional investors in NASDAQ 100 stocks. Goetzmann and Massa (2003) find a strong contemporaneous correlation between fund flows and S&P 500 index returns. Brunnermeier and Nagel (2004) find evidence of hedge funds catching the upturn and avoiding much of the downturn during the technology bubble. Temin and Voth (2004) conduct a case study of a well-informed investor during the South Sea Bubble and find that this investor actively invested in and profited from the bubble. Our study is consistent with the prior studies by suggesting that market run-ups are not corrected by arbitragers right away. Instead, profits can be made by trading with the herd. We propose upward stock return comovement as an indicator of large fund inflows and can be used to study trading behaviors in market episodes.

3.3 Data Sources and Investment Strategies

3.3.1 Comovement-based Investment Strategies

We use all NYSE common stocks (share code 10 to 12) to define the market-wide upward comovement measure. Daily stock returns over the period of 1926-2014 are from the CRSP daily stock file. We choose the NYSE because it has the longest history among the U.S. stock exchange and has a large number of actively traded stocks. Based on the method in Morck, Yeung, and Yu (2002), we define daily upward comovement on day *t* as:

$$[1] \qquad comove_{k,t} = \frac{N^{up}_{t,t-k}}{N^{up}_{t,t-k} + N^{down}_{t,t-k}},$$

where $N^{up}_{t,t-k}$ and $N^{down}_{t,t-k}$ denote, respectively, the number of stocks with positive and negative cumulative returns in the prior *k* trading days. We set *k* to 5, 10, 15, or 20 trading days (1, 2, 3, or 4 calendar weeks). The short window allows $comove_{k,t}$ to react to recent changes in market conditions. Using weekly or longer period cumulative returns also avoids non-synchronous trading issue in daily returns.

We then define a high upward comovement dummy

$$[2] \qquad HIcomove_{k,t} = \begin{cases} 1 & if\ comove_{k,t-1} > comove\ threshold_{k,1\text{ to }t-1}, \\ 0 & if\ comove_{k,t-1} \leq comove\ threshold_{k,1\ to\ t-1}. \end{cases}$$

We use the historical comovement deciles as the comovement thresholds. For each trading day t, we calculate the past $comove_{k,t-j}$ (*j = 1, ... t-1*) up to *t-1* and define the 1st, 2nd, …, and the 9th deciles of the historical comovement distribution as the thresholds. Because the CRSP dataset we use to construct the comovement measures starts on January 1st, 1926, this date is set as $t=1$. We use the long time series of CRSP data to define comovement thresholds so that the comovement thresholds are not affected by recent market trend.

Our comovement-based investment strategy holds the value-weighted NYSE index if $HIcomove_{t,k}$ is equal to 1 and invests in the three-month U.S. Treasury bill secondary market if $HIcomove_{t,k}$ is equal to 0. The portfolio is adjusted daily. Figure 3.1 illustrates the time line of this strategy. We start to implement this strategy on January 4th, 1954, when T-bill rate become available at daily frequency[22]. For example, $comove_5$ on January 2nd, 1980 is 0.33, meaning that the proportion of stocks moving up in the prior 5 trading days is 33%. This number is above the 2nd decile (30%) of the historical $comove_5$ from January 1st 1926 to January 1st, 1980. Thus $HIcomove_5$ =1 using the 2nd decile as the comovement threshold, and we hold the value-weighted index on this day based on this comovement signal.

3.3.2 **Summary Statistics**

Table 3.1 compares NYSE daily value-weighted index returns on trading days with versus without high upward comovement, where the 1st to 9th deciles of historical comovement distribution is used as the comovement thresholds to define high upward comovement.

Panel A reports the results using one-week cumulative returns to define upward comovement. We find that trading days with high upward comovement (*HIcomove*=1) in the prior week exhibit higher mean returns than trading days without (*HIcomove*=0). The return differences increase as the comovement threshold used to define high upward comovement

[22] The three-month Treasury bill secondary market rate can be downloaded at https://research.stlouisfed.org/fred2/series/DTB3

increases from the 1st decile to the 4th decile and declines as the comovement threshold exceeds the 4th decile. Using the 1th decile value of historical comovement as the threshold, trading days with high upward comovement exhibit 0.033% higher daily returns than trading days without. The return difference is statistically insignificant. When the 4th decile historical comovement is used as the comovement threshold, the return difference rises to 0.085% and become significant at 1% level. This is an annualized return difference of 35.78%. The difference drops to 0.069% when the 5th decile return threshold is used and goes down further to 0.016% when the 9th decile is used.

Panels B, C, and D use the cumulative return in the prior two, three, and four weeks to define upward comovement, respectively, and find similar results. We call the windows used to define prior comovement reference windows. The return differences are smaller than those reported in Panel A, but remain positive and significant when the 3rd to 8th decile comovement thresholds are used to define high upward comovement. Across all four panels, using the 4th decile value as the comovement threshold and a one reference week window exhibits the largest difference in mean returns, suggesting that a short reference window combined with a moderate upward comovement reference point may describe the most profitable investment strategy on offer. A lower comovement threshold keeps our comovement-based strategy in the market longer, and thus incurs more risk of not getting out in time and losing money in market downturns. A higher threshold keeps our strategy in the risk-free asset longer and thus may miss out part of market upturns.

3.4 Risk-adjusted Performance of the Comovement-based Strategy

In this section, we calculate the monthly return of the comovement-based strategies and test whether they outperform the market index using risk-adjusted performance measures. Monthly cumulative returns for month *t* and strategy *p* are defined as

[3] $$R_{p,t} = \prod_{\tau=1}^{t}(R_{p,\tau} + 1)\text{-}1,$$

[4] $$R_{p,\tau} = \begin{cases} R_{m,\tau} & if\ HIcomove_{\tau} = 1 \\ R_{f,\tau} & if\ HIcomove_{\tau} = 0 \end{cases}$$

for the comovement-based strategy, and

[5] $$R_{p,\tau} = R_{m,\tau}$$

for the passive strategy; where $R_{m,\tau}$ is the value-weighted NYSE index return on day τ and $R_{f,\tau}$ is the three-month Treasury bill secondary market rate on day τ. $HIcomove_{\tau}$ is the high upward comovement dummy defined in eq. [2].

3.4.1 Sharpe Ratio

The Sharpe ratio is a measure of reward to total risk. It is defined as the ratio of the mean and the standard deviation of portfolio excess returns:

[6] $$Sharpe_p = \frac{\mu_p}{\sigma_p},$$

where $\mu_p = E(R_{p,t} - R_{f,t})$, and $\sigma_p = \sqrt{Var(R_{p,t} - R_{f,t})}$. $R_{p,t}$ and $R_{f,t}$ are the monthly returns of portfolio *p* and the risk-free asset in month *t*, respectively.

Table 3.2 reports Sharpe ratios for the comovement-based strategies and the passive strategy. The comovement-based strategies using the 2nd to the 6th deciles comovement thresholds exhibit higher Sharpe ratio than the market index regardless of the length of the reference window used to calculate prior comovement. The strategies using a one-week reference window dominate the market index, with both a higher mean return and a lower return standard deviation.

The Sharpe ratios exhibit a U-shape as the comovement threshold increases. The comovement-based strategy based on the 4^{th} decile comovement threshold and one-week reference window again performs the best. It Sharpe ratio is 1.8 times that of the market index. The strategies based on the 9^{th} decile comovement threshold generate the lowest Shape ratios. When the reference window is extended beyond one week, the Sharpe ratio declines. This finding is consistent with the prediction of information cascade model by Bikchandani, Hirshlefier, and Welch (1992), in which the arrival of public information or highly informed trader would stop an "incorrect" information cascade quickly. Information cascade thus should be more evident in short term. Christoffersen and Tang (2009) provide supporting empirical evidence using high frequency institutional trading data.

The last column of each panel lists the number of trades incurred by the comovement-based strategies. The number of trades decreases as the comovement threshold goes up. From January 1954 to December 2014, the comovement-based strategies using the 4^{th} decile comovement threshold require from 1,008 to 2,287 trades. This works out to an average of 1.38 to 3.13 trades per month. Even though the allocation decision is made daily, the comovement-based strategies require trades much less frequently.

3.4.2 Alphas from the Asset Pricing Models

We next use the alphas from the CAPM model and the four-factor model to test whether or not our comovement-based investment strategies generate returns above the passive strategy. These models are

$$[7] \qquad R_{p,t} - R_{f,t} = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \epsilon_{p,t},$$

$$[8] \qquad R_{p,t} - R_{f,t} = \alpha_p + \beta_{1p}(R_{m,t} - R_{f,t}) + \beta_{2p}SMB_t + \beta_{3p}HML_t + \beta_{4p}UMD_t + \epsilon_{p,t}\ ,$$

where $R_{p,t}$ denote the return of portfolio p in month t. $R_{m,t} - R_{f,t}$, SMB_t, HML_t, and UMD_t denote the excess market return and the returns of the size, growth, and the momentum factor in month *t*, respectively[23].

Table 3.3 reports the alphas of the comovement-based strategies. The alphas are positive and highly significant using the 3rd to 6th comovement thresholds across all reference windows. The highest alphas occur when high upward comovement in the prior one week is used as the signal to be in the market index. The alphas drops around a half when the window used to define comovement signal extend above one week. The strategy using the 4th decile comovement threshold exhibits the highest alphas among all the strategies. The four-factor model alpha of the strategy using the 4th decile comovement threshold and one-week window is 0.52% per month or 6.42% per year. The alphas of the strategies using the 1st, 8th and 9th decile comovement threshold are in general insignificant or negative.

The positive and significant alphas of our comovement-based strategies cannot be fully explained by transaction costs. Wermers (2000) estimates 0.28% of expense ratio and 0.07% of

[23] The monthly factor returns are downloaded from Kenneth French's website http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

transaction costs per year for Vanguard Index 500 fund from 1977 to 1994. Karceski, Livingston, and O'Neal (2004) estimate 0.071% of brokerage commission per year for index mutual funds in 2002. Kacperczyk, Sialm, and Zheng (2008) calculate the average trading costs of actively managed mutual funds to be 0.58% per year and expense ratio of 1.24% per year from 1984 to 2003. The large alphas of the comovement-based strategies using the 3rd to the 6th decile of comovement thresholds and the one-week reference window exceed the estimated transaction costs.

To summarize, Tables 3.3 and 3.4 show the comovement-based strategy that use the 3rd to the 6th deciles of historical comovement as thresholds outperform the market index in both Sharpe ratios and alphas based on asset pricing models. These findings are consistent with high upward comovement in the prior one to four weeks indicating periods of large fund inflows and investing in market index being abnormally profitable during at these times.

3.5 Robustness checks

3.5.1 **Equal-weighted Market Index**

Market-wide upward comovement is defined as the proportion of stocks moving up in the prior one to four weeks. This definition assigns equal weight to each stock. We thus expect the comovement-based strategies to work particularly well in market run-ups characterized by many small stocks moving up together. We therefore test whether the comovement-based strategies might yield even more superior performance using the equal-weighted market index as the investment vehicle.

Table 3.4 reports the alphas for the alternative comovement strategies. We use the equal-weighted NYSE index return to calculate the risk premium in the asset pricing models. We find that the comovement-based strategies investing in equal-weighted market index exhibit stronger performance. Both the magnitudes and significance levels of the alphas exceed those in Table 3.3. The comovement-based strategy using the 4^{th} decile comovement threshold and a one-week reference window still generates the highest alpha. The four-factor model alpha of this strategy is 0.80% per month or 10.08% per year, and the CAPM alpha is 0.72% per month or 8.98% per year. The strategies using the 1^{st} and the 2^{nd} decile comovement threshold also exhibit positive and significant alphas across all reference windows. The scale of the alphas is still smaller than that of the alphas using the 4^{th} decile comovement threshold, suggesting that these positive alphas are due to the superior performance at times when prior comovement is exceeds the higher comovement thresholds. The alphas of the comovement-based strategies using the 9^{th} decile comovement threshold remain insignificant or negative.

3.5.2 **An Index Return-based Benchmark Strategy**

We next test whether or not the past comovement provides additional information about large fund inflows that is not discernable in past index returns. To do this, we construct a benchmark strategy to identify periods of large fund inflows using abnormal sequences of positive market return alone, without resorting to stock-level comovement data.

If good and bad market-wide fundamentals news arrives randomly, a run of positive market returns is likely to be caused by large fund inflows. This suggests a profitable trading strategy of holding the market index after observing a sequential cluster of positive returns and holding the risk-free asset otherwise. However, positive market returns can reflect either a few stocks with very high returns or many stocks rising together, and a sequential cluster of the latter accords better with "stocks" becoming a favored style of investment. The upward comovement measure thus provides additional information to distinguish these two possibilities, and thus better identify episodes of positive feedback investors moving money into the market and lifting the prices of all stocks.

To identify a period of clustered positive market returns in a short window, such as 5 or 10 trading days, we define the positive daily market return dummy on day *t* as

$$[9] \qquad ret_t^+ = \begin{cases} 1, & if\ Index\ return_t > 0 \\ 0, & if\ Index\ return_t \leq 0 \end{cases}$$

and use binomial tests to identify sequential clusters of positive market returns. Assuming ret_t^+ binomially distributed with a probability parameter of 0.5, we define a clustered sequence of positive market returns dummy using binomial tests as

$$[10] \qquad CPreturn_{t,k} = \begin{cases} 1 & if\ Prob\left(Nret_{t-1,t-k}^+ > Nret_k^+\right) \leq 0.5 \\ 0 & if\ Prob\left(Nret_{t-1,t-k}^+ > Nret_k^+\right) > 0.5 \end{cases},$$

where $Nret_{t-1,t-k}^+$ is the number of positive return days from *t-k* to *t-1*, and $Nret_k^+$is the expected number of positive return days in a window of *k* trading days, where k = 5, 10, 15, or 20 – that is, 1, 2, 3 or 4 weeks long calendar window. For example, in a given 5 trading days, the binomial distribution puts the probability of observing at least 3 positive returns at 0.5. Thus $CPreturn_{t,k}$ = 1 if the market return actually is positive on 3 or more days in the window of *t-1* to *t-5*, and otherwise $CPreturn_{t,k}$ = 0. Probability thresholds other than 0.5 produce similar, but weaker, performance for the index return-based strategy.

As with the comovement-based strategies, the index return-based strategies invest in the value-weighted NYSE index if $CPreturn_{t,k}$ is 1 and the three-month U.S. Treasury bill if $CPreturn_{t,k}$ is 0. The allocation decision is made on a daily basis.

[11] $$R_{p,\tau} = \begin{cases} R_{m,\tau} & if\ CPreturn_\tau = 1 \\ R_{f,\tau} & if\ CPreturn_\tau = 0 \end{cases}$$

The second bottom row of Table 3.4 lists the alphas for the index return-based benchmark strategies. The alphas are in general positive and significant, but the magnitude is smaller than that of the comovement-based strategies in Table 3.3. The index return-based strategy using a one-week reference window yields the highest alpha, an annualized return of 4.20%. However, the positive alphas of the other index return-based strategies largely disappear after taking transaction costs into consideration. For example, annual transaction costs of 0.7% per year would reduce the alpha of the strategy using a longer than one-week reference window to near zero.

We next test whether the difference between the alphas of the comovement-based strategy and index return-based strategy is statitically significant. To do this, we replace the monthly portfolio excess returns $R_{p,t} - R_{f,t}$ in asset pricing models with the difference in returns between the comovement-based portfolio *p* and return-based portfolio *b*, $R_{p\text{-}b,t}$, and rerun the regressions with the same right-hand side variables.

Table 3.4 reports the incremental alphas of the comovement-based strategies compared to the index return-based strategies. We find positive and significant alphas for the comovement-based strategies using the 4th and the 5th decile comovement thresholds. For example, the comovement-based strategy using the 4th decile comovement threshold and a one-week reference window outperforms the index return-based strategy using the same reference window by 0.18% per month, or 2.23% per year. Extending the reference window up to four weeks, we still find significant difference in alphas between the comovement-based strategies and the index return-based benchmark strategy. These findings are consistent with information about the market-wide comovement in the prior one to four weeks not being discernable from market returns after adjusting for systematic risk factors and past market returns.

3.5.3 **Combined Strategies of Return and Comovement**

Both the comovement- and index return-based strategies are designed to detect and exploit episodes of large fund inflows, consistent with the previous findings associating large fund inflows with boosts to asset prices (Goetzmann and Massa 2003; Froot, O'Connell, and Seasholes, 2001). We next ask if a strategy that combines the signals provided by past returns and comovement performs better that comovement- and return-based strategies applied separately. We consider two ways of combining the two strategies: investing in the market index when observing *both* returns and comovement buying signals, and investing in the market index when observing return *or* comovement buying signals. On each trading day τ, the portfolio returns of the two combined strategies are defined as:

[12] $$R_{c1,\tau} = \begin{cases} R_{m,\tau} & if\ HIcomove_{\tau} = 1\ and\ CPreturn_{\tau} = 1 \\ R_{f,\tau} & if\ else \end{cases}$$

[13] $$R_{c2,\tau} = \begin{cases} R_{m,\tau} & if\ HIcomove_{\tau} = 1\ or\ CPreturn_{\tau} = 1 \\ R_{f,\tau} & if\ else \end{cases},$$

where $HIcomove_{\tau}$ and $CPreturn_{\tau}$ are the high upward comovement and clustered positive return dummies defined in eq. [2] and [10], respectively. To test the performance of the combined strategies, we use the comovement-based (4th decile comovement thresholds) and index return-based portfolio returns as benchmarks and estimate the incremental alphas of the combined strategies.

Table 3.6 reports the incremental alphas for the combined strategies from the CAPM and the four-factor models. The combined strategies in Panel A use both past return and past comovement signals in deciding whether to be in the market index or T-bills. These strategies in general underperform the comovement-based strategies. Their incremental alphas are negative, and significantly so in short reference windows where the comovement-based strategies perform best. The combined strategy using a one-week reference window underperforms the comovement-based benchmark strategy 2.22% per year. The performance differences between the combined strategies and the index return-based strategies are mostly insignificant. A marginal significant alpha occurs when the strategies use a four-week reference window, where the performance of the index return-based strategies is the weakest.

Panel B reports incremental alphas of the combined strategies using either return or comovement signal. According to the two asset pricing models, this set of strategies consistently outperforms the index return-based strategies but underperforms the comovement-based strategies. The incremental alphas using the comovement-strategies as the benchmark are negative and significantly so in most cases, while the incremental alphas using the return-based strategies as the benchmark range from -0.02% to 0.21% per month. These findings are consistent with those in Section 3.4.2. High upward comovement identifies periods of market run-ups more accurately than can prior index returns. Adding past comovement information to the return-based strategies significantly improve portfolio performance.

3.6 Does it work in Other Stock Markets?

The previous sections develop comovement-based strategies that perform well in the NYSE. We next investigate whether analogous strategies work in other stock markets. We choose NASDAQ and Tokyo Stock Exchange, markets that differ from the NYSE in many ways but that also are thought to have experiences episodes of large fund inflows.

3.6.1 **NASDAQ**

NASDAQ stocks have a wider size range and tend to be smaller and younger than NYSE stocks. As shown in section 3.5.1, our comovement strategy seems better at exploiting market run-ups driven by smaller stocks than by large stocks. We therefore go in and out of the equal-weighted NASDAQ market index in our comovement-based strategies. Our NASDAQ comovement measure begins in January 1993 because very few NASDAQ stocks traded actively before then, and our comovement measure is of uncertain informational value if we employ the bid-ask means NASDAQ substitutes for prices if volume is zero. We require at least five years of data to calculate the deciles of historical comovement distribution. Our investment strategies thus begin in January 1999.

Panel A of Table 3.7 reports incremental alphas of the comovement-based strategies compared to the equal-weighted NASDAQ index. The incremental alpha is defined as the intercept from the regression of the difference in returns between the two strategies on the systematic risk factor(s). We do this because the equal-weighted NASDAQ index yields positive and significant alphas over the sample period. Incremental alphas allow us to test whether or not our comovement-based strategies significantly outperform the passive strategy holding the equal-weighted NASDAQ index.

The incremental alphas of the comovement-based strategies are in general positive when the 1st to the 6th deciles comovement thresholds are used to define high upward comovement. The highest alphas occur when the comovement-based strategy is defined with the 3rd decile comovement threshold and a one-week reference window. This strategy outperforms the equal-weighted NASDAQ index by 0.86% per month or 10.85% per year. The alpha declines as the comovement threshold goes above the 3rd decile, but remains significant up to the 6th decile. As the reference window used to calculate prior comovement extends from one week to four weeks, the incremental alphas decline. These findings are consistent with those using NYSE stocks. A short reference window combined with a moderate upward comovement reference point detects profitable periods to invest in market index.

Panels B of Table 3.7 uses the return-based strategies as the benchmark. The return-based strategies invests in the equal-weighted NASDAQ index after observing statistically significant runs of positive returns in the prior one to four weeks, and invests in the T-bills otherwise. The point estimates for the alphas become smaller, but the best performed comovement-based strategies (using the 3rd and 4th decile threshold and a one week reference window) nonetheless outperform both benchmarks The annualized incremental alphas for these two strategies are 4.43% and 4.74%, respectively.

3.6.2 **Tokyo Stock Exchange**

The size distribution of stocks listed on Tokyo Stock Exchange (TSE) resembles that on NASDAQ more than that of the NYSE. The TSE has high growth start-ups as well as large and well-established companies. We again use the equal-weighted market index as the investment tool for comovement-based strategies. Our comovement-based strategy in TSE starts in July 1990, when Fama and French factors for the Japanese market becomes available. The historical comovement distribution is calculated with all of the available Japanese stock return data in DataStream, starting in January 1973. We define an equal-weighted TSE market return using all common stocks, and our comovement-based strategies invest in this equal-weighted index when prior upward comovement is high and invest in the Gensaki one-month Treasury bill otherwise. The Gensaki rate, an interest rate applied to bond repurchase agreements, is used as proxy for equivalent Treasury bill rate in Japan (Chan, Hamao, and Lakonishok, 1991; Campbell and Hamao, 1992). Monthly portfolio returns are converted to U.S. dollars to match the factor returns provided be Kenneth French.

Table 3.8 reports incremental alphas for comovement-based strategies compared to the equal-weighted TSE index and the index-return based strategies. The results are consistent with those in NASDAQ and NYSE. The comovement-based strategies consistently outperform the passive strategy. The incremental alphas of the comovement-based strategies using the 3rd to the 8th decile comovement threshold are all positive and significant. The highest alphas occur when the 4th decile of the historical distribution is used as the comovement threshold. This strategy outperforms the equal-weighted market index by 7.46% to 12.33% per year. The strategies using the 3rd to 5th decile comovement thresholds and a short reference window also outperform the index-return based strategies that switch between the equal-weighted TSE index and the T-bill. The incremental alpha of the strategy using the 4th decile comovement threshold and a one-week reference window is 3.98% per year.

3.7 Concluding Remarks

We design a profitable investment strategy using past comovement information to detect and profit from market run-ups. This comovement-based strategy invests in the market index if past market-wide upward comovement exceeds a threshold decile of its historical distribution, and otherwise invests in the risk-free asset. This simply strategy generates large positive Shape ratio and large significantly positive alphas from asset pricing models. This strategy outperforms the value-weighted market index and an index return-based strategy designed as an alternative approach to detecting periods of large fund inflows.

The findings are consistent with the style investment model of Barberis and Shleifer (2003) that style-chasing investors reallocate funds into styles with superior past performance, artificially creating comovement among assets within the style as their prices rise. Asset prices so affected thus keep going up in the short run, but reverse in the longer run when the fad is over. We find that high upward comovement identifies period of high short-term style returns, and posit that these are related to large fund inflows. Using the NYSE stocks from 1954 to 2014, we find that the comovement-based strategies generate an incremental alpha of 6.42% per year compared with the value-weighted market index and an incremental alpha of 2.23% per year compared with an index return-based benchmark strategy. We test the performance of the comovement-based strategy in two other stock markets and find consistent results.

Our findings support the noise trading models of De Long, et al. (1990b) and Abreu and Brunnermeier (2002), which both arguing that arbitrageurs can profit by taking advantage of such market run-ups, rather than by correcting the mispricing and trading against the herd. They are also consistent with limited arbitrage (De Long, et al. 1990a, Shleifer and Vishny, 1997) making it more profitable to trade with the herd for a time, so that their trading actually contributes to the persistence of abnormal returns.

Our findings are also consistent with the prediction of the information cascades model where rational investors imitate the trades of others and completely disregard their own private information. Our comovement-based investment strategy detects large fund inflows driven by investors’ correlated demands during the information cascades and profits before the cascades been ended by the arrival of new public information or a highly informed trader.

Figure 3.1 Time line of the Comovement-based Investment Strategy

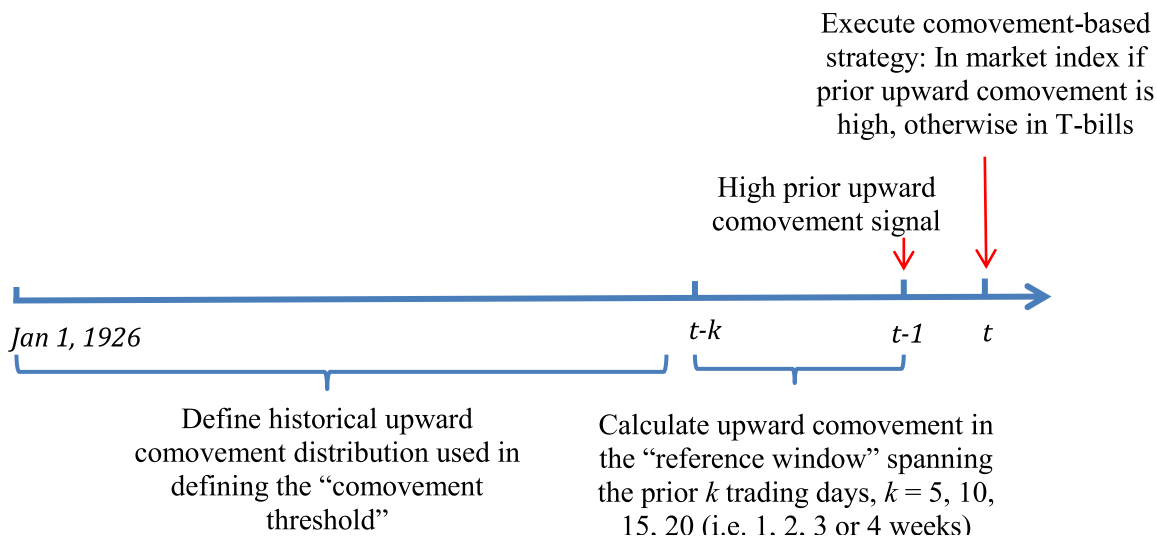


Table 3.1 Summary of the NYSE Value-Weighted Index Return, 1954-2014

This table contrasts the daily NYSE value-weighted market index return on trading days with versus without higher than threshold upward comovement. For each trading day *t*, the comovement threshold is defined as the nine deciles value of the distribution up to day *t-1*. *HIcomove* is a dummy variable set to 1 if the market-wide upward return comovement is above the comovement threshold. Panel A, B, C, and D uses the 1, 2, 3, and 4 weeks cumulative return to define the market-wide upward comovement. P-values for the differences in mean return between the two sets of trading days in each panel are in parentheses. Numbers in bold indicate statistical significance at least at the 5% level.

Comovement threshold decile	*HIcomove*	*Mean (%)*	*Std. Dev.*	*N*	*Mean Diff.*	*Mean (%)*	*Std. Dev.*	*N*	*Mean Diff.*
		Panel A 1-week Upward Comovement				**Panel B 2- week Upward Comovement**			
1	1	0.048	0.839	14,300	0.033	0.046	0.833	14,456	-0.002
	0	0.015	1.763	1,056	(0.27)	0.048	1.918	900	(0.95)
2	1	0.055	0.812	12,907	**0.056**	0.051	0.801	13,015	0.036
	0	-0.002	1.405	2,449	(0.01)	0.015	1.460	2,341	(0.09)
3	1	0.064	0.804	11,232	**0.067**	0.054	0.784	11,411	0.032
	0	-0.004	1.213	4,124	(0.00)	0.022	1.266	3,945	(0.06)
4	1	0.078	0.788	9,537	**0.085**	0.056	0.772	9,655	0.028
	0	-0.007	1.127	5,819	(0.00)	0.028	1.153	5,701	(0.07)
5	1	0.079	0.793	7,914	**0.069**	0.063	0.771	7,907	**0.036**
	0	0.010	1.059	7,442	(0.00)	0.027	1.077	7,449	(0.02)
6	1	0.088	0.796	6,208	**0.071**	0.066	0.767	6,136	**0.033**
	0	0.017	1.013	9,148	(0.00)	0.032	1.027	9,220	(0.03)
7	1	0.080	0.814	4,491	**0.049**	0.066	0.782	4,362	0.029
	0	0.032	0.977	10,865	(0.00)	0.038	0.985	10,994	(0.08)
8	1	0.072	0.868	2,793	0.032	0.060	0.824	2,646	0.018
	0	0.040	0.946	12,563	(0.10)	0.043	0.953	12,710	(0.37)
9	1	0.060	0.991	1,314	0.016	0.048	0.938	1,161	0.003
	0	0.044	0.927	14,042	(0.55)	0.046	0.932	14,195	(0.93)

Comovement threshold decile	*HIcomove*	*Mean (%)*	*Std. Dev.*	*N*	*Mean Diff.*	*Mean (%)*	*Std. Dev.*	*N*	*Mean Diff.*
		Panel C 3-week Upward Comovement				**Panel D 4-week Upward Comovement**			
1	1	0.046	0.832	14,478	-0.003	0.047	0.833	14,552	0.017
	0	0.048	1.949	878	(0.93)	0.030	2.008	804	(0.62)
2	1	0.050	0.792	13,124	0.032	0.049	0.787	13,212	0.021
	0	0.019	1.513	2,232	(0.13)	0.028	1.551	2,144	(0.34)
3	1	0.058	0.769	11,509	**0.049**	0.056	0.763	11,528	**0.039**
	0	0.009	1.303	3,847	(0.00)	0.017	1.316	3,828	(0.02)
4	1	0.057	0.759	9,730	0.029	0.058	0.746	9,712	**0.034**
	0	0.027	1.173	5,626	(0.06)	0.024	1.186	5,644	(0.03)
5	1	0.060	0.746	7,990	**0.030**	0.067	0.726	7,877	**0.043**
	0	0.030	1.099	7,366	(0.04)	0.024	1.108	7,479	(0.00)
6	1	0.064	0.742	6,132	**0.030**	0.069	0.722	6,054	**0.038**
	0	0.034	1.039	9,224	(0.05)	0.031	1.046	9,302	(0.01)
7	1	0.078	0.757	4,280	**0.045**	0.076	0.751	4,271	**0.042**
	0	0.033	0.991	11,076	(0.01)	0.034	0.993	11,085	(0.01)
8	1	0.086	0.781	2,554	**0.048**	0.088	0.779	2,520	**0.050**
	0	0.038	0.959	12,802	(0.02)	0.038	0.959	12,836	(0.01)
9	1	0.087	0.835	1,052	0.044	0.090	0.869	1,028	0.048
	0	0.043	0.939	14,304	(0.14)	0.043	0.936	14,328	(0.11)

Table 3.2 Sharpe Ratios of the Comovement-Based Strategies

This table reports the Sharpe ratios of the comovement-based strategies and the NYSE value-weighted market index. The mean and standard deviation are calculated with monthly portfolio excess returns. The risk-free rate is the one-month Treasury bill rate. The comovement strategies invests in the valued-weighed NYSE market index if the comovement in prior 1 to 4 weeks is above the comovement threshold, and invests in the three-month Treasury bill secondary market otherwise. For each trading day *t*, the comovement threshold is defined as the 1st to the 9th decile value of the distribution up to day *t-1*. Panel A, B, C, and D use the cumulative returns in the prior 1, 2, 3, and 4 weeks to define upward comovement, respectively.

Comovement threshold decile	*Mean*	*Std. Dev.*	*Sharpe Ratio.*	*# of Trades*	*Mean*	*Std. Dev.*	*Sharpe Ratio.*	*# of Trades*
	Panel A 1-week Upward Comovement				**Panel B 2-week Upward Comovement**			
1	0.593	3.891	0.152	745	0.560	3.998	0.140	463
2	0.645	3.563	0.181	1,447	0.591	3.580	0.165	911
3	0.692	3.351	0.207	1,947	0.553	3.344	0.165	1,253
4	0.758	3.033	0.250	2,287	0.480	3.014	0.159	1,525
5	0.626	2.727	0.230	2,501	0.449	2.687	0.167	1,698
6	0.544	2.447	0.222	2,587	0.347	2.296	0.151	1,692
7	0.316	2.126	0.149	2,380	0.221	1.981	0.111	1,600
8	0.126	1.767	0.072	1,860	0.072	1.646	0.044	1,252
9	-0.016	1.327	-0.012	1,055	-0.046	1.188	-0.038	702

Comovement threshold decile	*Mean*	*Std. Dev.*	*Sharpe Ratio.*	*# of Trades*	*Mean*	*Std. Dev.*	*Sharpe Ratio.*	*# of Trades*
	Panel C 3-week Upward Comovement				**Panel D 4-week Upward Comovement**			
1	0.558	3.938	0.142	377	0.580	3.920	0.148	283
2	0.588	3.716	0.158	675	0.559	3.666	0.153	591
3	0.624	3.360	0.186	967	0.582	3.224	0.180	857
4	0.493	3.061	0.161	1,225	0.512	2.847	0.180	1,008
5	0.424	2.603	0.163	1,318	0.488	2.519	0.194	1,110
6	0.330	2.300	0.143	1,432	0.365	2.165	0.169	1,164
7	0.285	1.982	0.144	1,328	0.270	1.919	0.141	1,070
8	0.156	1.570	0.100	991	0.156	1.426	0.109	797
9	0.003	1.090	0.003	511	0.004	1.060	0.004	417
Market Index	0.589	4.199	0.140	-				

Table 3.3 Alphas of the Comovement-based Strategies

This table lists the alphas (in percentage) from the time-series four-factor model and the CAPM model of the comovement-based strategies. The regression equation is $R_{p,t} - R_{f,t} = \alpha_p + \beta_{1p}(R_{m,t} - R_{f,t}) + \beta_{2p}SMB_t + \beta_{3p}HML_t + \beta_{4p}UMD_t + \epsilon_{p,t}$ for the four-factor model and $R_{p,t} - R_{f,t} = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \epsilon_{p,t}$ for the CAPM model, where $R_{p,t}$ denote the excess return of portfolio *p* in month *t*, and $R_{m,t}$, SMB_t, HML_t, and UMD_t denote the excess market return, the returns of the size, growth, and the momentum factor in month *t*, respectively. For each trading day *t*, the comovement thresholds are the 1st to the 9th deciles of the historical upward comovement from January 1st 1926 to trading day *t-1*. Each column indicates the calendar window used to calculate the cumulative return and the upward comovement. The comovement-based strategies invest in the valued-weighed NYSE market index if the upward comovement in prior 1 to 4 weeks is higher than the comovement thresholds, and invests in the three-month Treasury bill secondary market otherwise. P-values are reported in parentheses. Numbers in bold indicate statistical significance at least at the 5% level.

	Four-factor Model				*CAPM Model*			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	**0.123**	0.068	0.068	0.048	**0.101**	0.055	0.069	0.086
	(0.02)	(0.23)	(0.27)	(0.39)	(0.07)	(0.34)	(0.27)	(0.13)
2	**0.276**	**0.187**	**0.168**	0.096	**0.227**	**0.174**	**0.156**	0.126
	(0.00)	(0.01)	(0.02)	(0.17)	(0.00)	(0.01)	(0.03)	(0.07)
3	**0.382**	**0.234**	**0.278**	**0.226**	**0.329**	**0.190**	**0.259**	**0.229**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
4	**0.520**	**0.182**	**0.212**	**0.257**	**0.464**	**0.170**	**0.185**	**0.225**
	(0.00)	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.02)	(0.00)
5	**0.425**	**0.196**	**0.223**	**0.301**	**0.376**	**0.195**	**0.186**	**0.265**
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
6	**0.391**	**0.142**	**0.160**	**0.217**	**0.339**	**0.148**	**0.152**	**0.198**
	(0.00)	(0.04)	(0.03)	(0.00)	(0.00)	(0.03)	(0.03)	(0.00)
7	**0.165**	0.049	**0.157**	**0.153**	**0.164**	0.069	**0.149**	**0.138**
	(0.02)	(0.44)	(0.02)	(0.02)	(0.02)	(0.26)	(0.02)	(0.03)
8	0.013	-0.061	0.065	0.069	0.015	-0.030	0.076	0.076
	(0.84)	(0.28)	(0.24)	(0.17)	(0.80)	(0.59)	(0.16)	(0.12)
9	-0.067	**-0.097**	-0.054	-0.064	-0.070	**-0.094**	-0.037	-0.035
	(0.17)	(0.03)	(0.18)	(0.10)	(0.14)	(0.03)	(0.34)	(0.36)

Table 3.4 Alphas of the Comovement-based Strategies, Investing in the Equal-Weighted Market Index

This table lists the alphas from the time-series four-factor model and the CAPM model of the comovement-based strategies using the *equal-weighted* NYSE index return as the investment vehicle. The market return in the asset pricing models is also calculated with the equal-weighted NYSE index. For each trading day *t*, the comovement thresholds are the 1st to the 9th deciles of the historical upward comovement from January 1st 1926 to trading day *t-1*. Each column indicates the calendar window used to calculate the prior upward comovement. The comovement-based strategies invest in the equal-weighed NYSE market index if the upward comovement in prior 1 to 4 weeks is higher than the comovement thresholds, and invests in the three-month Treasury bill secondary market otherwise. P-values are reported in parentheses. Numbers in bold indicate statistical significance at least at the 5% level.

	Four-factor Model				*CAPM model*			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	**0.228**	**0.174**	**0.158**	**0.128**	**0.160**	**0.112**	0.114	**0.132**
	(0.00)	(0.00)	(0.02)	(0.04)	(0.00)	(0.05)	(0.08)	(0.02)
2	**0.488**	**0.379**	**0.330**	**0.225**	**0.401**	**0.312**	**0.275**	**0.217**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
3	**0.638**	**0.455**	**0.500**	**0.416**	**0.552**	**0.358**	**0.434**	**0.381**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
4	**0.804**	**0.436**	**0.435**	**0.480**	**0.719**	**0.373**	**0.362**	**0.401**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
5	**0.720**	**0.428**	**0.443**	**0.534**	**0.646**	**0.383**	**0.363**	**0.454**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
6	**0.665**	**0.378**	**0.359**	**0.419**	**0.588**	**0.345**	**0.316**	**0.364**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
7	**0.377**	**0.234**	**0.303**	**0.312**	**0.357**	**0.217**	**0.261**	**0.276**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
8	**0.162**	0.060	**0.169**	**0.171**	**0.155**	0.072	**0.158**	**0.168**
	(0.02)	(0.36)	(0.01)	(0.00)	(0.02)	(0.26)	(0.02)	(0.00)
9	0.019	-0.042	-0.002	-0.025	0.002	-0.053	0.003	-0.004
	(0.74)	(0.41)	(0.96)	(0.61)	(0.97)	(0.28)	(0.95)	(0.93)

Table 3.5 Incremental Alphas of the Comovement-based Strategies, Compared with the Index Return-based Strategies

This table lists the incremental alphas for the comovement-based strategies compared with the index return-based strategies. The left-hand side variable in the asset pricing models is the monthly *return difference* between the comovement-based strategy and the index-return based strategy. The comovement-based strategy invests in the valued-weighed NYSE market index if the upward comovement in prior 1 to 4 weeks is higher than the comovement thresholds, and invests in the three-month Treasury bill secondary market otherwise. The return-based strategy invests in the valued-weighed NYSE market index if the number of positive return days in prior 1 to 4 weeks is higher than expected value of a binomial distribution, and invests in the three-month Treasury bill secondary market otherwise. The last two rows of the table reports the alphas of the index return-based strategies based on the asset pricing models. P-values are reported in parentheses. Numbers in bold indicate statistical significance at least at the 5% level.

	Four-factor Model				*CAPM model*			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	**-0.216**	**0.102**	0.018	0.014	**-0.236**	0.024	**-0.034**	-0.008
	(0.00)	(0.00)	(0.29)	(0.40)	(0.00)	(0.16)	(0.04)	(0.63)
2	**-0.060**	**0.223**	**0.114**	**0.061**	**-0.106**	**0.146**	**0.048**	0.031
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)
3	**0.046**	**0.273**	**0.221**	**0.189**	-0.004	**0.164**	**0.149**	**0.132**
	(0.01)	(0.00)	(0.00)	(0.00)	(0.80)	(0.00)	(0.00)	(0.00)
4	**0.184**	**0.222**	**0.153**	**0.214**	**0.128**	**0.144**	**0.073**	**0.125**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
5	**0.087**	**0.228**	**0.160**	**0.255**	**0.040**	**0.163**	**0.074**	**0.162**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
6	**0.051**	**0.176**	**0.092**	**0.173**	0.001	**0.119**	**0.035**	**0.095**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.94)	(0.00)	(0.04)	(0.00)
7	**-0.173**	**0.083**	**0.092**	**0.113**	**-0.173**	**0.038**	0.033	**0.039**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.06)	(0.02)
8	**-0.324**	-0.030	-0.001	**0.031**	**-0.320**	**-0.063**	**-0.039**	-0.021
	(0.00)	(0.06)	(0.97)	(0.07)	(0.00)	(0.00)	(0.02)	(0.20)
9	**-0.409**	**-0.064**	**-0.118**	**-0.100**	**-0.409**	**-0.125**	**-0.151**	**-0.132**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Index return-based strategies	**0.344**	**-0.031**	**0.062**	**0.037**	**0.340**	**0.033**	**0.113**	**0.097**
	(0.00)	(0.06)	(0.00)	(0.03)	(0.00)	(0.04)	(0.00)	(0.00)

Table 3.6 Incremental Alphas of the Combined Strategies

Table 3.6 reports the incremental alphas for the combined strategies compared with the comovement-based and index return-based strategies. In Panel A, the combined strategy invests in the valued-weighed NYSE market index if *both* the comovement- and the index return-based strategies indicate buying, and invests in the three-month Treasury bill secondary market otherwise. In Panel B, the combined strategy invests in the market index if either the comovement- *or* index return-based strategy indicates buying. The left-hand side variable in the asset pricing models is the monthly *return difference* between the combined strategy and the comovement- or index return-based strategy. The comovement-based strategy invests in the valued-weighed NYSE market index if the upward comovement in prior 1 to 4 weeks is higher than the comovement thresholds, and invests in the three-month Treasury bill secondary market otherwise. The return-based strategy invests in the valued-weighed NYSE market index if the number of positive return days in prior 1 to 4 weeks is higher than expected value of a binomial distribution, and invests in the three-month Treasury bill secondary market otherwise. P-values are reported in parentheses. Numbers in bold indicate statistical significance at least at the 5% level.

	Four-factor Model				*CAPM model*			
	1-week	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
Panel A Comovement and Return Signals								
Comovement-based Strategy as Benchmark								
	-0.180	**-0.201**	-0.044	**-0.128**	**-0.136**	**-0.123**	0.024	-0.069
	(0.00)	(0.00)	(0.42)	(0.01)	(0.01)	(0.04)	(0.65)	(0.17)
Index Return-based Strategy as Benchmark								
	0.004	0.013	0.100	**0.085**	-0.010	0.012	0.090	0.052
	(0.93)	(0.71)	(0.08)	(0.06)	(0.84)	(0.72)	(0.10)	(0.23)
Panel B Comovement or Return Signals								
Comovement-based Strategy as Benchmark								
	-0.006	**-0.016**	**-0.102**	**-0.084**	0.006	**-0.015**	**-0.091**	**-0.053**
	-(0.57)	-(0.04)	(0.00)	(0.00)	-(0.57)	-(0.04)	(0.00)	(0.00)
Index Return-based Strategy as Benchmark								
	0.177	**0.206**	**0.051**	**0.130**	**0.134**	**0.129**	-0.018	**0.072**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-(0.11)	(0.00)

Table 3.7 Alphas of the Comovement-based Strategies, NASDAQ

Table 3.7 reports the alphas for the comovement-based strategies from January 1999 to December 2014 using NASDAQ sample. Alphas are defined as the intercept from the regressions of the four-factor model and the CAPM model. The left-hand side variable in the asset pricing models is the monthly *return difference* between the comovement-based strategy and the benchmark strategy. Panels A and B use NASDAQ equal-weighted market index and the index-return based strategy as benchmarks, respectively. The comovement-based strategy invests in the *equal-weighed* NASDAQ market index if the upward comovement in prior 1 to 4 weeks is higher than the comovement thresholds, and invests in the three-month Treasury bill otherwise. For each trading day *t*, the comovement thresholds are the 1st to the 9th deciles of the historical upward comovement from January 1st 1993 to trading day *t-1*. The index return-based strategy invests in the equal-weighed NASDAQ market index if the number of positive return days in prior 1 to 4 weeks is higher than expected value from a binomial distribution, and invests in the three-month Treasury bill otherwise. Each column indicates the calendar window used to calculate the past upward comovement. P-values are reported in parentheses. Numbers in bold indicate statistical significance at least at the 10% level.

Panel A The Equal-weighted NASDAQ Index Returns as Benchmark

	Four-factor Model				*CAPM model*			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	**0.694**	**0.454**	**0.388**	**0.388**	**0.613**	**0.470**	**0.384**	**0.390**
	(0.00)	(0.03)	(0.05)	(0.08)	(0.00)	(0.03)	(0.06)	(0.09)
2	**0.739**	**0.513**	**0.632**	**0.396**	**0.648**	**0.506**	**0.606**	**0.395**
	(0.00)	(0.01)	(0.00)	(0.06)	(0.00)	(0.01)	(0.00)	(0.08)
3	**0.862**	**0.411**	**0.668**	**0.442**	**0.778**	**0.400**	**0.638**	**0.437**
	(0.00)	(0.04)	(0.00)	(0.04)	(0.00)	(0.06)	(0.00)	(0.06)
4	**0.743**	**0.375**	**0.589**	**0.493**	**0.648**	**0.336**	**0.563**	**0.465**
	(0.00)	(0.04)	(0.00)	(0.02)	(0.00)	(0.11)	(0.01)	(0.04)
5	**0.636**	0.294	**0.511**	**0.371**	**0.527**	0.241	**0.464**	0.305
	(0.00)	(0.11)	(0.01)	(0.08)	(0.01)	(0.27)	(0.02)	(0.18)
6	**0.660**	**0.432**	**0.448**	**0.392**	**0.549**	**0.367**	**0.368**	0.320
	(0.00)	(0.01)	(0.01)	(0.06)	(0.01)	(0.09)	(0.07)	(0.16)
7	**0.424**	**0.285**	0.261	0.310	0.328	0.204	0.161	0.221
	(0.02)	(0.10)	(0.15)	(0.13)	(0.13)	(0.34)	(0.46)	(0.36)
8	**0.333**	0.050	0.057	-0.063	0.237	-0.035	-0.084	-0.135
	(0.09)	(0.78)	(0.76)	(0.76)	(0.32)	(0.88)	(0.72)	(0.59)
9	-0.141	-0.218	**-0.320**	**-0.462**	-0.202	-0.340	**-0.430**	**-0.557**
	(0.43)	(0.18)	(0.08)	(0.02)	(0.39)	(0.13)	(0.07)	(0.03)

Panel B The Index Return-based Strategies as Benchmark

	Four-factor Model				CAPM model			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	-0.325	**-0.411**	-0.202	**-0.625**	-0.319	**-0.461**	-0.293	**-0.590**
	(0.08)	(0.04)	(0.27)	(0.00)	(0.07)	(0.02)	(0.10)	(0.00)
2	0.079	-0.016	**0.323**	-0.279	0.110	-0.110	0.209	**-0.308**
	(0.64)	(0.93)	(0.03)	(0.09)	(0.50)	(0.54)	(0.16)	(0.05)
3	**0.362**	0.043	**0.446**	0.156	**0.392**	-0.023	**0.377**	0.104
	(0.03)	(0.81)	(0.00)	(0.31)	(0.02)	(0.89)	(0.01)	(0.48)
4	**0.387**	0.065	0.337	0.224	**0.364**	0.071	0.247	0.144
	(0.03)	(0.69)	(0.05)	(0.14)	(0.03)	(0.65)	(0.14)	(0.32)
5	0.301	-0.050	0.298	0.172	0.284	-0.040	0.194	0.025
	(0.10)	(0.76)	(0.10)	(0.28)	(0.11)	(0.80)	(0.28)	(0.87)
6	0.276	0.132	0.283	0.234	0.270	0.109	0.151	0.081
	(0.14)	(0.47)	(0.13)	(0.16)	(0.14)	(0.54)	(0.41)	(0.62)
7	0.076	-0.008	0.134	0.147	0.111	-0.057	-0.011	0.036
	(0.71)	(0.97)	(0.51)	(0.40)	(0.57)	(0.76)	(0.96)	(0.84)
8	-0.020	-0.140	0.053	-0.128	0.044	-0.217	-0.111	-0.228
	(0.93)	(0.49)	(0.81)	(0.51)	(0.83)	(0.27)	(0.60)	(0.24)
9	-0.410	-0.302	-0.202	**-0.470**	-0.313	-0.416	-0.325	**-0.551**
	(0.07)	(0.19)	(0.37)	(0.03)	(0.15)	(0.06)	(0.14)	(0.01)

Table 3.8 Alphas of the Comovement-based Strategies, Tokyo Stock Exchange

Table 3.8 reports the alphas for the comovement-based strategies from July 1990 to December 2014 using TSE sample. Alphas are defined as the intercepts from the regressions of the four-factor model and the CAPM model. The left-hand side variable in the asset pricing models is the monthly *return difference* between the comovement-based strategy and the benchmark strategies. Panel A and Panel B uses the TSE equal-weighted market index and the index-return based strategies as benchmark, respectively. The comovement-based strategy invests in the *equal-weighed* TSE market index if the upward comovement in prior 1 to 4 weeks is higher than the comovement thresholds, and invests in the three-month Treasury bill otherwise. For each trading day *t*, the comovement thresholds are the 1st to the 9th deciles of the historical upward comovement from January 1st 1973 to trading day *t-1*. Each column indicates the calendar window used to calculate the prior upward comovement. P-values are reported in parentheses. Numbers in bold indicate statistical significance at least at the 5% level.

Panel A The Equal-weighted NASDAQ Index Returns as Benchmark

	Four-factor Model				*CAPM model*			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	**0.381**	0.200	**0.319**	0.166	**0.355**	0.193	0.277	0.141
	(0.00)	(0.05)	(0.01)	(0.24)	(0.00)	(0.06)	(0.05)	(0.35)
2	**0.805**	**0.591**	**0.544**	0.314	**0.758**	**0.544**	**0.493**	0.252
	(0.00)	(0.00)	(0.00)	(0.06)	(0.00)	(0.00)	(0.01)	(0.20)
3	**0.948**	**0.776**	**0.695**	**0.598**	**0.924**	**0.736**	**0.648**	**0.542**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
4	**0.974**	**0.866**	**0.699**	**0.637**	**0.943**	**0.810**	**0.644**	**0.602**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
5	**0.965**	**0.869**	**0.706**	**0.681**	**0.944**	**0.803**	**0.657**	**0.654**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
6	**0.877**	**0.794**	**0.543**	**0.677**	**0.865**	**0.727**	**0.475**	**0.652**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)
7	**0.776**	**0.680**	**0.595**	**0.649**	**0.745**	**0.609**	**0.539**	**0.611**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
8	**0.583**	**0.578**	**0.574**	**0.492**	**0.537**	**0.509**	**0.523**	**0.446**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
9	**0.352**	0.270	0.267	0.215	0.295	0.203	0.215	0.184
	(0.01)	(0.04)	(0.05)	(0.16)	(0.09)	(0.28)	(0.22)	(0.29)

Panel B The Index Return-based Strategies as Benchmark

	Four-factor Model				*CAPM model*			
Comovement threshold decile	*1-week*	*2-week*	*3-week*	*4-week*	*1-week*	*2-week*	*3-week*	*4-week*
1	**-0.268**	-0.279	-0.158	-0.243	**-0.252**	-0.236	-0.154	-0.243
	(0.01)	(0.05)	(0.26)	(0.09)	(0.02)	(0.13)	(0.21)	(0.05)
2	0.156	0.112	0.066	-0.094	0.152	0.115	0.062	-0.133
	(0.20)	(0.30)	(0.68)	(0.48)	(0.23)	(0.28)	(0.68)	(0.22)
3	**0.300**	**0.297**	0.217	0.189	**0.318**	**0.307**	0.216	0.158
	(0.01)	(0.02)	(0.23)	(0.11)	(0.01)	(0.01)	(0.20)	(0.12)
4	**0.326**	**0.386**	0.221	0.229	**0.337**	**0.381**	0.213	0.217
	(0.00)	(0.00)	(0.18)	(0.08)	(0.01)	(0.01)	(0.19)	(0.06)
5	**0.316**	**0.390**	0.229	**0.273**	**0.337**	**0.374**	0.225	**0.270**
	(0.01)	(0.01)	(0.12)	(0.00)	(0.02)	(0.02)	(0.14)	(0.00)
6	0.228	0.314	0.065	**0.269**	0.259	0.298	0.044	**0.267**
	(0.08)	(0.09)	(0.63)	(0.00)	(0.08)	(0.11)	(0.78)	(0.00)
7	0.128	0.200	0.117	**0.240**	0.138	0.180	0.108	**0.226**
	(0.35)	(0.18)	(0.34)	(0.02)	(0.37)	(0.25)	(0.45)	(0.02)
8	-0.066	0.099	0.096	0.084	-0.069	0.079	0.091	0.061
	(0.62)	(0.39)	(0.35)	(0.47)	(0.64)	(0.52)	(0.49)	(0.53)
9	**-0.297**	-0.210	**-0.210**	-0.193	**-0.311**	-0.226	-0.216	**-0.200**
	(0.02)	(0.10)	(0.02)	(0.09)	(0.03)	(0.07)	(0.08)	(0.03)

Bibliography

Chapter 1

Amihud, Yakov. Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets* 5.1 (2002): 31-56.

Barro, Robert J. The stock market and investment. *Review of Financial Studies* 3.1 (1990): 115-131.

Bennett, James A., Richard W. Sias, and Laura T. Starks. Greener pastures and the impact of dynamic institutional preferences. *Review of Financial Studies* 16.4 (2003): 1203-1238.

Bennett, James A., and Richard W. Sias. Why company-specific risk changes over time. *Financial Analysts Journal* 62.5 (2006): 89-100.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy* (1992): 992-1026.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives* (1998): 151-170.

Brandt, Michael W., et al. The idiosyncratic volatility puzzle: Time trend or speculative episodes?. *Review of Financial Studies* (2009): hhp087.

Bris, Arturo, William N. Goetzmann, and Ning Zhu. Efficiency and the bear: Short sales and markets around the world. *The Journal of Finance* 62.3 (2007): 1029-1079.

Brockman, Paul, Ivonne Liebenberg, and Maria Schutte. Comovement, information production, and the business cycle. *Journal of Financial Economics* 97.1 (2010): 107-129.

Brown, Gregory, and Nishad Kapadia. Firm-specific risk and equity market development. *Journal of Financial Economics* 84.2 (2007): 358-388.

Campbell, John Y., et al. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance* 56.1 (2001): 1-43.

Cao, Charles, Timothy Simin, and Jing Zhao. Can growth options explain the trend in idiosyncratic risk?. *Review of financial studies* 21.6 (2008): 2599-2633.

Chan, Kalok, and Allaudeen Hameed. Stock price synchronicity and analyst coverage in emerging markets. *Journal of Financial Economics* 80.1 (2006): 115-147.

Chun, Hyunbae, et al. Creative destruction and firm-specific performance heterogeneity. *Journal of Financial Economics* 89.1 (2008): 109-135.

Chun, Hyunbae, Jung-Wook Kim, and Randall Morck. Varying heterogeneity among US firms: facts and implications. *Review of economics and statistics*93.3 (2011): 1034-1052.

Clark, Todd E., and Michael W. McCracken. Tests of equal forecast accuracy and encompassing for nested models. *Journal of econometrics* 105.1 (2001): 85-110.

Davis, Steven J., et al. Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. *NBER Macroeconomics Annual 2006, Volume 21*. MIT Press, 2007. 107-180.

De Long, J. Bradford, et al. Noise trader risk in financial markets. *Journal of political Economy* (1990): 703-738.

Durnev, Artyom, et al. Does greater firm-specific return variation mean more or less informed stock pricing?. *Journal of Accounting Research* 41.5 (2003): 797-836.

Durnev, Art, Randall Morck, and Bernard Yeung. Value-enhancing capital budgeting and firm-specific stock return variation. *The Journal of Finance* 59.1 (2004): 65-105.

Fama, Eugene F., and Kenneth R. French. Industry costs of equity. *Journal of financial economics* 43.2 (1997): 153-193.

Fama, Eugene F., and Kenneth R. French. New lists: Fundamentals and survival rates. *Journal of Financial Economics* 73.2 (2004): 229-269.

Estrella, Arturo, and Frederic S. Mishkin. The yield curve as a predictor of US recessions. *Current Issues in Economics and Finance* 2.7 (1996).

Estrella, Arturo, and Frederic S. Mishkin. Predicting US recessions: financial variables as leading indicators. *Review of Economics and Statistics* 80.1 (1998): 45-61.

Fair, Ray C., and Robert J. Shiller. Comparing information in forecasts from econometric models. *The American Economic Review* 80.3 (1990): 375-389.

Fink, Jason, Kristin E. Fink, Gustavo Grullon and James P. Weston. What drove the increase in idiosyncratic volatility during the internet boom? *Journal of Financial and Quantitative Analysis* 45.5 (2010): 1253-1278.

Fischer, Stanley, and Robert C. Merton. Macroeconomics and finance: The role of the stock market. *Carnegie-Rochester Conference Series on Public Policy*. Vol. 21. North-Holland, 1984.

French, Kenneth R., and Richard Roll. Stock return variances: The arrival of information and the reaction of traders. *Journal of financial economics* 17.1 (1986): 5-26.

Gordon, Robert J. The American business cycle. continuity and change. *The American Business Cycle: Continuity and Change*. University of Chicago Press, 1986. 15-0.

Harvey, Campbell R. The real term structure and consumption growth. *Journal of Financial Economics* 22.2 (1988): 305-333.

Irvine, Paul J., and Jeffrey Pontiff. Idiosyncratic return volatility, cash flows, and product market competition. *Review of Financial Studies* 22.3 (2009): 1149-1177.

Jin, Li, and Stewart C. Myers. R^2 around the world: New theory and new tests. *Journal of Financial Economics* 79.2 (2006): 257-292.

Kindleberger, Charles P. Manias, Panics and crashes: A history of financial crites, New York, Basic Books, 1978.

King, Robert G., and Ross Levine. Finance and growth: Schumpeter might be right. *The quarterly journal of economics* (1993): 717-737.

Xu, Yexiao, and Burton G. Malkiel. Investigating the behavior of idiosyncratic volatility*. *The Journal of Business* 76.4 (2003): 613-645.

Morck, Randall, Bernard Yeung, and Wayne Yu. R^2 and the economy, *Annual Review of Financial Economics* 5 (2013): 143-166.

Næs, Randi, Johannes A. Skjeltorp, and Bernt Arne Ødegaard. Stock market liquidity and the business cycle. *The Journal of Finance* 66.1 (2011): 139-176.

Nelson, Daniel B. Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of Econometrics* 52.1 (1992): 61-90.

Newey, Whitney K., and Kenneth D. West. Hypothesis testing with efficient method of moments estimation. *International Economic Review* (1987): 777-787.

Pesaran, H. Hashem, and Yongcheol Shin. Generalized impulse response analysis in linear multivariate models. *Economics letters* 58.1 (1998): 17-29.

Pástor, Ľuboš, and Veronesi Pietro. Stock valuation and learning about profitability. *The Journal of Finance* 58.5 (2003): 1749-1790.

Piotroski, Joseph D., and Darren T. Roulstone. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The Accounting Review* 79.4 (2004): 1119-1151.

Rajgopal, Shiva, and Mohan Venkatachalam. Financial reporting quality and idiosyncratic return volatility. *Journal of Accounting and Economics* 51.1 (2011): 1-20.

Rogoff, Kenneth S., and Vania Stavrakeva. The continuing puzzle of short horizon exchange rate forecasting. No. w14071. National Bureau of Economic Research, 2008.

Roll, Richard. R^2, *Journal of Finance* 43.3 (1988): 541-566.

Rosenberg, Joshua V., and Samuel Maurer. Signal or noise? Implications of the term premium for recession forecasting. *Federal Reserve Bank of New York-Economic Policy Review* 14.1 (2008).

Samuelson, Paul. Science and stocks. *Newsweek, September* 19 (1966): 92.

Scheffer, Marten, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. Van Nes, Max Rietkerk, and George Sugihara Early-warning signals for critical transitions. *Nature* 461.7260 (2009): 53-59.

Schwert, G. William. Stock volatility in the new millennium: how wacky is Nasdaq?. *Journal of Monetary Economics* 49.1 (2002): 3-26.

Schiller, Robert J. Irrational exuberance. Princeton University Press (2000).

Stock, James H., and Mark W. Watson. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual 1989, Volume 4*. MIT press, 1989. 351-409.

Veldkamp, Laura L. Slow boom, sudden crash. *Journal of Economic Theory*124.2 (2005): 230-257.

Veldkamp, Laura L. Information markets and the comovement of asset prices. *The Review of Economic Studies* 73.3 (2006): 823-845.

Vogelsang, Timothy J. Trend function hypothesis testing in the presence of serial correlation. *Econometrica* 66.1 (1998): 123-148.

Wei, Steven X., and Chu Zhang. Why did individual stocks become more volatile?. *The Journal of Business* 79.1 (2006): 259-292.

Zhang, Lu. Creative destruction and firm-specific return variation: evidence from the 1920s and 1930s. working paper, 2014.

Chapter 2

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti. Distance to frontier, selection, and economic growth. *Journal of the European Economic association* 4.1 (2006): 37-74.

Aghion, Philippe, and Peter Howitt. A model of growth through creative destruction. *Econometrica* (1990): 323–351.

Aghion, Philippe, Peter Howitt, and Cecilia García-Peñalosa. Endogenous growth theory. MIT press, 1998.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy* 100.5 (1992): 992-1026.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives* 12.3 (1998): 151-170.

Bresnahan, Timothy F., and Daniel MG Raff. Intra-industry heterogeneity and the Great Depression: The American motor vehicles industry, 1929–1935. *The Journal of Economic History* 51.2 (1991): 317-331.

Bresnahan, Timothy F., and Manuel Trajtenberg. General purpose technologies "engines of growth"?. *Journal of econometrics* 65.1 (1995): 83-108.

Bris, Arturo, William Goetzmann, and Ning Zhu. Efficiency and the bear: short-sales and markets around the world. *Journal of Finance* 62.3 (2007):1029-79.

Brown, Gregory, and Nishad Kapadia. Firm-specific risk and equity market development. *Journal of Financial Economics* 84.2 (2007): 358-388.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance* 56.1 (2001): 1-43.

Chun, Hyunbae, Jung-Wook Kim, Randall Morck, and Bernard Yeung. Creative destruction and firm-specific performance heterogeneity. *Journal of Financial Economics* 89.1 (2008): 109-135.

Comin, Diego A., and Thomas Philippon. The rise in firm-level volatility: causes and consequences. *NBER Macroeconomics Annual 2005,* Volume 20. MIT Press, 2006. 167-228.

David, Paul A. and Gavin Wright. General purpose technologies and surges in productivity: historical reflections on the future of the ICT revolution. In: David, P.A., Thomas, M. (Eds.), The Economic Future in Historical Perspective. Oxford University Press for The British Academy, 1999.

Davis, Steven J., John Haltiwanger, Ron Jarmin, and Javier Miranda. Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. NBER Macroeconomics Annual 21 (2007): 107-180.

Devine, Warren D. From shafts to wires: Historical perspective on electrification. *Journal of Economic History* 43.02 (1983): 347-372.

Durnev, Art, Randall Morck, and Bernard Yeung. Value-enhancing capital budgeting and firm-specific stock return variation. *The Journal of Finance* 59.1 (2004): 65-105.

Fama, Eugene F., and Kenneth R. French. New lists: fundamentals and survival rates. *Journal of Financial Economics* 73.2 (2004): 229-269.

Field, Alexander J. The most technologically progressive decade of the century. *American Economic Review* (2003): 1399-1413.

Field, Alexander J. A great leap forward: 1930s depression and US economic growth. Yale University Press, 2011.

Fink, Jason, Kristin E. Fink, Gustavo Grullon and James P. Weston. What drove the increase in idiosyncratic volatility during the internet boom?. *Journal of Financial and Quantitative Analysis* 45 (2010): 1253-1278.

Frehen, Rik GP, William N. Goetzmann, and K. Geert Rouwenhorst. New evidence on the first financial bubble. *Journal of Financial Economics* 108.3 (2013): 585-607.

Galambos, Louis. The U.S. corporate economy in the twentieth century. in Stanley Engerman and Robert Gallman, eds., The Cambridge Economic History of the United States, vol. 3, (2000): 927-967.

Galbraith, John Kenneth. The great crash of 1929. Houghton Mifflin Harcourt, 1961.

Gompers, Paul and Joshua Lerner. The venture capital cycle, 2nd ed. MIT Press, 2004.

Gordon, Robert J. The 1920s and the 1990s in mutual reflection. No. w11778, National Bureau of Economic Research, 2005.

Griliches, Zvi. (ed.) R&D, patents, and productivity, NBER Conference Proceedings. University of Chicago Press, 1984.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg. The NBER Patent-Citations Data File: Lessons, insights, and methodological tools. in patents, citation, and innovations: a window on the knowledge economy, edited by Adam Jaffe and Manuel Trajtenberg, 403–59. Cambridge, MA: MIT Press, 2002.

Harris, Christopher, and John Vickers. Perfect equilibrium in a model of a race. *The Review of Economic Studies* 52.2 (1985): 193-209.

Harris, Christopher, and John Vickers. Racing with uncertainty. *The Review of Economic Studies* 54.1 (1987): 1-21.

Hayek, Friedrich A. The road to serfdom. University of Chicago Press, 1944.

Helpman, Elhanan, and Manuel Trajtenberg. A time to sow and a time to reap: growth based on general purpose technologies. in: E. Helpman (ed.) General purpose technologies and economic growth, Cambridge, MIT press (1998).

Hobijn, Bart, and Boyan Jovanovic. The information-technology revolution and the stock market: evidence. *American Economic Review* (2001): 1203-1220.

Irvine, Paul J., and Jeffrey Pontiff. Idiosyncratic return volatility, cash flows, and product market competition. *Review of Financial Studies* 22.3 (2009): 1149-1177.

Jin, Li, and Stewart C. Myers. R^2 around the world: new theory and new tests. *Journal of Financial Economics* 79.2 (2006): 257-292.

Jovanovic, Boyan, and Peter L. Rousseau. Two technological revolutions. *Journal of the European Economic Association* 1.2-3 (2003): 419-428.

Jovanovic, Boyan, and Peter L. Rousseau. General purpose technologies. *Handbook of economic growth* 1 (2005): 1181-1224.

Li, Kan, Randall Morck, Fan Yang, and Bernard Yeung. Firm-specific variation and openness in emerging markets. *Review of Economics and Statistics* 86.3 (2004): 658-669.

Loury, Glenn C. Market structure and innovation. *The Quarterly Journal of Economics* (1979): 395-410.

Morck, Randall, Bernard Yeung, and Wayne Yu. The information content of stock markets: why do emerging markets have synchronous stock price movements?. *Journal of financial economics* 58.1 (2000): 215-260.

Mowery, David C., and Nathan Rosenberg. Technology and the pursuit of economic growth. Cambridge University Press, 1989.

Mowery, David C. Path of innovation: technological change in the 20th century America, Cambridge University Press, 1998.

Moser, Petra, and Tom Nicholas. Was electricity a general purpose technology?. *The American Economic Review, Papers and Proceedings*. 94.2 (2004): 388-394.

Nicholas, Tom. Does innovation cause stock market runups? Evidence from the great crash. *The American Economic Review* (2008): 1370-1396.

Nicholas, Tom. Did R&D firms used to patent? Evidence from the first innovation surveys. *The Journal of Economic History* 71.04 (2011): 1032-1059.

Pastor, Lubos, and Pietro Veronesi. Was there a Nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81.1 (2006): 61-100.

Pastor, Lubos, and Pietro Veronesi. Technological revolutions and stock prices. *American Economic Review* 99.4 (2009), 1451-1483.

Richard, Du Boff. Electric power in American manufacturing, 1889-1958, Arno Press, (1979).

Roll, Richard. R^2. The *Journal of Finance* 43.3 (1988): 541-566.

Romer, Paul M. Endogenous technological change. *Journal of political Economy* 98.5 (1990): S71-S102.

Schankerman, Mark. and Ariel Pakes. Estimates of the value of patent rights in European countries during the post-1950 period, *Economic Journal*, 96.12 (1986): 1052-1077.

Schmookler, J. Invention and economic growth. Harvard University Press, 1966.

Schumpeter, Joseph A. The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle. Harvard University Press, 1921.

Schumpeter, Joseph A. Business cycles. Vol. 1. McGraw-Hill, 1939.

Shleifer, Andrei, and Lawrence H. Summers. The noise trader approach to finance. *The Journal of Economic Perspectives* 4.2 (1990): 19-33.

Szostak, Rick. Technological innovation and the Great Depression. Westview Press, 1995.

Tobin, James. On the efficiency of the financial system. *Lloyd's Banking Review* 1 (1984): 1-15.

Wei, Steven X., and Chu Zhang. Why did individual stocks become more volatile?. *The Journal of Business* 79.1 (2006): 259-292.

Chapter 3

Abreu, Dilip, and Markus K. Brunnermeier. Synchronization risk and delayed arbitrage. *Journal of Financial Economics* 66.2 (2002): 341-360.

Asness, Clifford S., John M. Liew, and Ross L. Stevens. Parallels between the cross-sectional predictability of stock and country returns. *The Journal of Portfolio Management* 23.3 (1997): 79-87.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler. Comovement. *Journal of Financial Economics* 75.2 (2005): 283-317.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy* (1992): 992-1026.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives* (1998): 151-170.

Bris, Arturo, William Goetzmann and Ning Zhu. Efficiency and the Bear: Short-Sales and Markets around the World. *The Journal of Finance* 62.3 (2007):1029-79.

Brunnermeier, Markus K., and Stefan Nagel. Hedge funds and the technology bubble. *The Journal of Finance* 59.5 (2004): 2013-2040.

Campbell, John Y., and Yasushi Hamao. Predictable stock returns in the United States and Japan: A study of long-term capital market integration. *The Journal of Finance* 47.1 (1992): 43-69.

Campbell, John Y., and Albert S. Kyle. Smart money, noise trading and stock price behaviour. *Review of Economic Studies* 60 (1993):1-34.

Chan, Louis KC, Yasushi Hamao, and Josef Lakonishok. Fundamentals and stock returns in Japan. *The Journal of Finance* 46.5 (1991): 1739-1764.

Chen, Hsiu-Lang, and Werner De Bondt. Style momentum within the S&P-500 index. *Journal of Empirical Finance* 11.4 (2004): 483-507.

Christoffersen, Susan Kerr, and Ya Tang. Institutional herding and information cascades: Evidence from daily trades. Working paper (2010).

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers and Robert J.Waldmann. Noise trader risk in financial markets. *Journal of political Economy* 98.4 (1990a): 703-738.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers and Robert J.Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance* 45.2 (1990b): 379-395.

Diamond, Douglas W., and Robert E. Verrecchia. Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics* 9.3 (1981): 221-235.

Fama, Eugene F., and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics* 33.1 (1993): 3-56.

Froot, Kenneth A., Paul GJ O'connell, and Mark S. Seasholes. The portfolio flows of international investors. *Journal of Financial Economics* 59.2 (2001): 151-193.

Griffin, John M., Jeffrey H. Harris, and Selim Topaloglu. The dynamics of institutional and individual trading. *The Journal of Finance* 58.6 (2003): 2285-2320.

Grossman, Sanford J., and Joseph E. Stiglitz. Information and competitive price systems. The American Economic Review 66.2 (1976): 246-253.

Goetzmann, W. N., and M. Massa. Index Funds and Stock Market Growth. *The Journal of Business* 76.1 (2003):1–28.

Jegadeesh, Narasimhan, and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance* 48.1 (1993): 65-91.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. Unobserved actions of mutual funds. *Review of Financial Studies* 21.6 (2008): 2379-2416.

Karceski, Jason, Miles Livingston, and Edward S. O'Neal. Portfolio transactions costs at US equity mutual funds. Zero Alpha Group, (2004).

Kindleberger, Charles P. Manias, panics and crashes: a history of financial crises. New York: Basic Books, 1978.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. The impact of institutional trading on stock prices. *Journal of financial economics* 32.1 (1992): 23-43.

Morck, Randall, Bernard Yeung, and Wayne Yu. The information content of stock markets: why do emerging markets have synchronous stock price movements?. *Journal of financial economics* 58.1 (2000): 215-260.

Moskowitz, Tobias J., and Mark Grinblatt. Do industries explain momentum?. *The Journal of Finance* 54.4 (1999): 1249-1290.

Shiller, Robert J., Stanley Fischer, and Benjamin M. Friedman. Stock prices and social dynamics. *Brookings Papers on Economic Activity*, (1984) 457-510.

Shleifer, Andrei, and Robert W. Vishny. The limits of arbitrage. *The Journal of Finance* 52.1 (1997): 35-55.

Shleifer, Andrei, and Lawrence H. Summers. The noise trader approach to finance. *Journal of Economic Perspectives* 4.1 (1990): 19-33.

Sias, Richard W. Institutional herding. *Review of Financial Studies* 17.1 (2004): 165-206.

Temin, Peter, and Hans-Joachim Voth. Riding the south sea bubble. *American Economic Review* 94.5 (2004).

Teo, Melvyn, and Sung-Jun Woo. Style effects in the cross-section of stock returns. *Journal of Financial Economics* 74.2 (2004): 367-398.

Wahal, Sunil, and M. Deniz Yavuz. Style investing, comovement and return predictability. *Journal of Financial Economics* 107.1 (2013): 136-154.

Wermers, Russ. Mutual fund herding and the impact on stock prices. *The Journal of Finance* 54.2 (1999): 581-622.

Wermers, Russ. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance* 55.4 (2000): 1655-1703.