Title
Essays in Empirical Asset Pricing

Author
Sabbatucci, Riccardo

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays in Empirical Asset Pricing

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Management

by

Riccardo Sabbatucci

Committee in charge:

Professor Allan Timmermann, Chair
Professor Snehal Banerjee
Professor Graham Elliott
Professor Joseph E. Engelberg
Professor Christopher A. Parsons
Professor Rossen Valkanov

2016

The dissertation of Riccardo Sabbatucci is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2016

DEDICATION

To my family and friends.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Allan Timmermann, my primary advisor and committee chair. I am eternally grateful for his excellent guidance and patience throughout these years. Looking back, I realize how much I have grown through numerous emails, conversations, and meetings with him.

I am also very grateful to Christopher A. Parsons and Rossen Valkanov, my other important sources of learning throughout my graduate school years. Chris has always been genuinely interested in my research and dedicated a lot of time and effort to improving my writing and presentation skills. He has also been a great mentor, suggesting research topics and discussing research ideas. Ross, my office neighbor, has always given me great support and constructive feedback, and really helped me preparing for the job market with practice interviews and insightful comments on my presentation. I really appreciate the time he took to read my papers and the many discussions we had about academia.

I would like to thank two other committee members, Snehal Banerjee and Joey Engelberg for the valuable comments on my job market paper received during my presentations and for the many informal talks. I would also like to thank Sheridan Titman, one of my coauthors, for being extremely proactive and insightful, and Graham Elliott for the econometrics class he taught me and for accepting to serve on my committee.

I also want to thank my fellow Rady students and friends, in particular Wei, Yinchu, Silvia and Kristen. Wei, a great friend and one of the best people I met here in San Diego, I will never forget our daily coffee breaks at Peets, and our discussions about credit cards, travels and investments. Yinchu, my “de facto” office mate, I will always remember our discussions about research and the many lunches we had together at Short Stop, The Bistro and Bella Vista. Silvia, thank you for some great dinners and conversations about life preferences and priorities. Kristen, my extremely valuable copy-editor, thank you for all the time you spent helping me out and, most importantly, for being fun.

I also owe a debt of gratitude to Marisol Nierva-Magnano and Taylor Robinette for their support along the way.

I would also like to thank my Italian buddies in San Diego, Claudio, Eugenia, Giacomo, Lorenzo, Marco, Simone, Simone, Stefano for the many dinners, trips, and adventures experienced together.

I also thank Alberto, for the multiple, daily phone calls over the last few years and Antonio, for his philosophical insights about life.

I am obviously thankful to my best, lifetime friends, Pietro and Luca, for constant support and encouragement over these years.

Finally, I owe a huge debt of gratitude to my family, in particular my mother, father and aunt. They have always been supporting me unconditionally, during the many ups and downs of the Ph.D.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Parsons, Christopher A.; Sabbatucci, Riccardo; Titman, Sheridan. The dissertation author was the primary investigator and author of this paper.

VITA

2007	B. S. in Business Administration *cum laude*, Bocconi University, Italy
2010	M. S. in Finance *cum laude*, Bocconi University, Italy
2016	Ph. D. in Management, University of California, San Diego, USA

ABSTRACT OF THE DISSERTATION

Essays in Empirical Asset Pricing

by

Riccardo Sabbatucci

Doctor of Philosophy in Management

University of California, San Diego, 2016

Professor Allan Timmermann, Chair

The focus of my dissertation is the study of stock market predictability. More precisely, I use econometric tools to understand, explain, and predict aggregate and cross-sectional patterns in stock prices. Predictability of aggregate stock market returns and dividend growth is a widely studied topic, of great interest to both academics and practitioners. It is related to theories of market efficiency and information diffusion, both rational and behavioral. It also allows us to determine which types of information generate the movements in stock prices that we observe. Understanding why stock prices move and what factors drive their variation is critical from theoretical and policy-making perspectives. Chapter 1 of my dissertation revisits one of the main findings of

the predictability literature, namely that all variation in aggregate stock prices is explained by changes in aggregate risk through discount rates and none by news about firms' expected cash flows. I propose a more comprehensive measure of dividends that includes M&A cash flows and show that dividend growth is predictable and that cash flow news explains around 60% of the observed variation in prices, while the remaining 40% is accounted for by discount rate news. Chapter 2 shows that information about fundamentals of the aggregate economy derived from closely held firms help predict stock returns of public firms. A common feature of most stock market predictors is that they are constructed using financial data of public firms. I construct a new economy-wide dividend-price ratio that takes into account dividends and market capitalization of both listed (public) and non-listed (private) U.S. companies and show that it strongly predicts stock returns both in-sample and out-of-sample. I also find that changes in dividends of private firms lead those of public firms and that the economy-wide dividend-price ratio subsumes the standard dividend-price ratio in predictive regressions. Chapter 3, co-authored with Christopher A. Parsons and Sheridan Titman, explores geographic momentum: a positive lead-lag stock return relation between neighboring firms operating in different sectors. It shows that a portfolio of firms headquartered in the same area, but operating in different sectors, strongly forecasts individual stock returns up to one year ahead. The economic significance of a city-momentum trading strategy is of similar magnitude to that observed with industry momentum. However, while industry momentum is strongest among thinly traded, small firms, and/or those with scant analyst following, geographic momentum is unrelated to these proxies for information processing. We propose an explanation linking this to the structure of the investment analyst business, which is organized by sector, rather than by geographic region.

Chapter 1

Are Dividends and Stock Returns Predictable? New Evidence Using M&A Cash Flows

Abstract

The lack of predictability of aggregate dividends has long been considered a puzzle - "the dog that did not bark", Cochrane (2008). I show that this empirical finding is related to the measurement of dividends. If M&A cash flows are taken into account, the adjusted R^2 from a regression of dividend growth on the dividend-price ratio goes from being negative (-1.18%) to being positive (17.54%) and coefficients become highly statistically significant. Strong improvements are also found for consumption growth (2.10% to 11.76%) and returns (1.86% to 4.40%). Out-of-sample R^2 for dividend growth and returns are large and statistically significant. I also show that dividend-price variation is fundamentally linked to cash flows news and not only to discount rate news. Lastly, I find stronger predictability in industries with the largest M&A activity.

1.1 Introduction

Conventional wisdom holds that aggregate dividend growth is largely unpredictable. Following the present value decomposition in Campbell and Shiller (1988b), the dividend-price ratio is used as predictor of expected stock returns and dividend growth. Standard asset pricing models imply that time variation in the dividend-price ratio must be accompanied by either time-varying risk premia or time-varying expected dividend growth. As a consequence, either returns or dividend growth (or both) must be predictable if the conventional present value model holds.

Predictability of dividend growth is economically important for several reasons.[1] Showing that dividend growth is predictable is equivalent to stating that cash flow news affects stock prices (Cochrane (2008)). Moreover, testing asset pricing and present value

[1] Shiller recently highlighted the importance of dividend growth while discussing stock price volatility during the Nobel Prize Lecture on the 8th December 2013.

models with better measures of dividend growth is important in order to assess their validity.[2] Market makers and equity investors require good dividend estimates when pricing equity options, futures and forwards, while the recently created market for dividend options and futures requires forecasts of future dividends and pricing of their term structure since it enables investors to trade dividends independently from the underlying asset. Asset managers focusing on industry allocations or dividend/cash flow strategies require a good estimate of future cash flows in order to improve their portfolio allocation. Last but not least, we care about dividend growth predictability because it could be possible to infer companies' future earnings from their past dividends, given that managers may observe industry practice selecting a target constant payout ratio (see Lintner (1956) or Marsh and Merton (1987)).

As discussed in Koijen and Nieuwerburgh (2011), most of the studies on stock market predictability focus on the return component.[3] Only recently the literature has shifted attention to dividend growth predictability[4] which is both important on its own - since dividend growth plays an essential role in asset pricing models - and because it is an implied outcome of the present value models that jointly analyze return and dividend dynamics (Lettau and Nieuwerburgh (2008), Cochrane (2008), Bollerslev et al. (2013), Piatti and Trojani (2013)).

The general finding of many empirical studies, sometimes accepted as a stylized fact, is that aggregate stock returns are predictable by the dividend-price ratio but

[2] Recently many authors have used present value models to analyze the joint dynamics of expected returns and dividend growth. See van Binsbergen and Koijen (2010), Kelly and Pruitt (2013), Piatti and Trojani (2013), Bollerslev et al. (2013) and Golez (2014).

[3] Pesaran and Timmermann (1995), Kothari and Shanken (1997), Lettau and Ludvigson (2001), Lewellen (2004), Robertson and Wright (2006), Campbell and Yogo (2006), Boudoukh et al. (2007), Goyal and Welch (2008), Campbell and Thompson (2008), Lettau and Nieuwerburgh (2008), Ferreira and Santa-Clara (2011), Shanken and Tamayo (2012), Li et al. (2013) and Johannes et al. (2014) amongst others.

[4] See, for example, Campbell and Shiller (1988a), Cochrane (1992), Goyal and Welch (2003), Menzly et al. (2004), Lettau and Ludvigson (2005), Ang and Bekaert (2007), Cochrane (2008), Chen (2009), van Binsbergen and Koijen (2010), Engsted and Pedersen (2010), Golez (2014) and Rangvid et al. (2014).

dividend growth is not (Cochrane (1992), Cochrane (2008), Lettau and Nieuwerburgh (2008), Cochrane (2011)). I will refer to this empirical evidence as the dividend growth predictability puzzle, since claiming that stock prices react only to variation in discount rates and not to news about future cash flows is, to say the least, counterintuitive.

More precisely, quoting Cochrane (2008), this lack of dividend growth predictability is puzzling because "our lives would be much easier if we could trace price movements back to visible news about dividends or cash flows...if market price-dividend ratio variation comes from varying expected returns and none from varying expected growth in dividends or earnings, much of the rest of finance still needs to be rewritten." I find that, luckily, this is not necessary. I introduce a more precise and comprehensive measure of cash flows containing M&A cash dividends and show that dividend growth, consumption growth and excess returns are strongly predictable, at both annual and quarterly frequencies, using standard predictive regressions. Most theories underlying the dividend-price ratio's usefulness in predicting either stock returns or dividend growth do not specify how cash is transferred between firms and their shareholders and this is probably why researchers have not focused on this important cash flow component.

Figure 1.1 shows the aggregate cash flows received by equity investors decomposed into four different categories, namely ordinary and liquidation dividends, M&A cash dividends, stock repurchases and new equity issues. It is immediately clear that cash flows from M&A activity together with repurchases are a very important component of shareholders' total payout starting from the 1970s, as highlighted by Allen and Michaely (2003). In fact, in many years (e.g., during the dot-com bubble period) these cash flows account for the great majority of the total cash flows received by shareholders. M&A cash flows include, in general, both cash and stock distributions originating from a merger or acquisition. I only add M&A cash dividends to adjust the standard dividend measures. This M&A cash component is not included in the measures of dividends com-

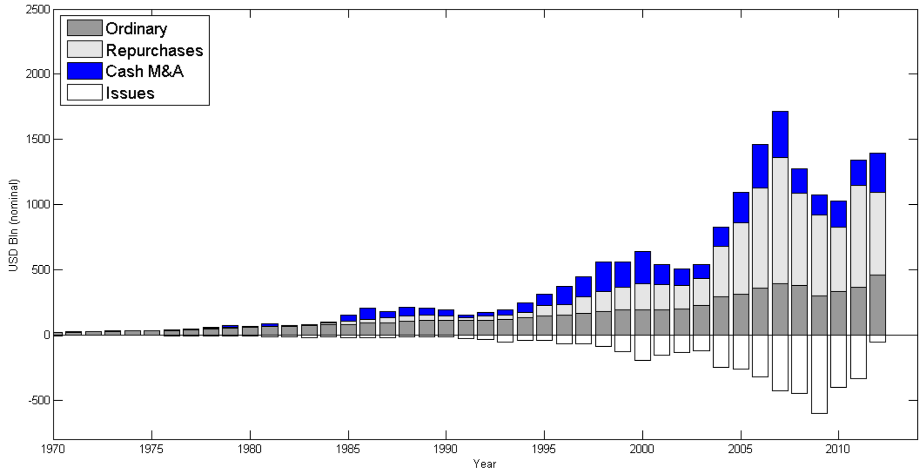

Figure 1.1: Total cash flows received by shareholders at the aggregate macro level. Ordinary dividends are from CRSP. M&A cash dividends are from CRSP pre-1980 and SDC post-1980. Stock repurchases and equity issues are from Compustat (buybacks: entry 115 plus reduction in 56; equity issues: entry 108 minus increase in 56). A positive (negative) bar represent a positive (negative) flow of funds to shareholders. All figures are nominal amount in bln dollars. Annual data, 1970-2012.

monly found in the literature for a trivial reason: only ordinary and liquidation dividends can be extracted "top-down" from the CRSP index.[5]

I motivate this chapter with an example highlighting the economic importance of M&A cash flows. At $t = 0$ an investor buys 1 share in firm X for \$40 each. At $t = 1$ this investor receives a \$2 dividend per share, resulting in a 5% dividend yield. At $t = 2$ firm X gets acquired by firm Y in an all cash takeover for \$100 per share, and the investor tenders his share. As a consequence, he receives \$100 as a "liquidation" dividend in order to give up his ownership of firm X. This liquidation dividend does not appear in the standard dividend measures used in the literature. But, as this example shows, the majority of the cash flows realized by the investor are in the form of M&A dividends at $t = 2$. This chapter revisits the predictability of the dividend-price ratio, taking into account both cash flows realized in the form of ordinary dividends ($t = 1$), as

[5] See Section 1.3.3 for a detailed discussion.

well as cash flows realized in the form of cash dividends from M&A ($t = 2$). As I show, this second consideration dramatically changes the ability of the dividend-price ratio to forecast future dividend growth, consumption growth and excess returns.

To the best of my knowledge, this chapter is the first to explicitly analyze the impact of M&A cash dividends for asset pricing. In so doing, I provide several contributions to the predictability literature.

First, I describe the problem of using dividends extracted from the CRSP value-weighted Index for the construction of both dividend growth and dividend-price ratios and discuss why a measure of cash flows should include M&A cash dividends.

Second, I construct a new dividend-price ratio (e.g., adjusting the numerator of the ratio) and dividend growth measure (e.g., adjusting both the numerator and denominator) that take into account M&A cash dividends. I highlight the importance of M&A cash dividends by showing their impact on dividend growth, consumption growth and excess return predictability.

I find adjusted R^2 values of 17.54% (8.22%) over the full sample period, 77.08% (42.63%) pre-1945 and 4.34% (1.84%) post-war in predictive regressions of dividend growth at annual (quarterly) frequency. Out-of-sample I find a statistically significant R^2 of 6.82% (1.73%). I also find strong consumption growth predictability, with an adjusted R^2 of 11.77% over the full sample and 6.75% (10.03%) post-war at the annual (quarterly) frequency. Dividend growth predictability means that cash flow news explain stock price variation and M&A cash flows are arguably the main cash flow news available in the market. I also show that dividend-price variation is fundamentally linked to cash flows news and not only to discount rate news within the long-run predictability framework of Cochrane (2008).

As far as return predictability is concerned, I report adjusted R^2 values of 4.40% (1.59%) over the full sample period and 10.62% (2.83%) post-war at annual (quarterly) frequency. Out-of-sample I find a statistically significant R^2 of 7.73% (2.16%). My predictor performs better, post-war and out-of-sample, than both the net payout yield of Boudoukh et al. (2007) and the total payout price ratio of Robertson and Wright (2006). I also gauge the importance of each individual cash flow component for return predictability and show that M&A cash dividends are the only component with explanatory power for future returns, while net equity issues are irrelevant if M&A cash dividends are accounted for. Moreover, I find strong out-of-sample return predictability only during periods of low M&A activity, which are related to recessions.

Lastly, I analyze the impact of M&A cash dividends on dividend growth and return predictability at the industry level. I find return predictability mainly in those industries that have the largest M&A activity (e.g., consumer non-durables, manufacturing, wholesale and retail), while dividend growth predictability is strong for every industry over the full sample, pre- and post-war. Using industry data has the potential of highlighting economic cross-dependencies amongst related industries and observe industry specific patterns absent at the aggregate level (Hong et al. (2007), Cohen and Frazzini (2008)).

A few recent papers discuss evidence on US dividend growth predictability. Chen (2009) shows that dividend growth predictability from the dividend-price ratio is present only pre-1945 if dividends are measured without reinvestment. This means that dividends paid out during the year by companies should be reinvested at the risk free rate or not reinvested at all. However, he does not find evidence of dividend growth predictability post-war or over the full sample. van Binsbergen and Koijen (2010) model expected returns and dividend growth rates as latent processes and use filtering techniques to show that both of them are predictable within a present value model, but they do not find any dividend growth predictability using OLS predictive regressions. Chen et al. (2013) find that there is a significant component of cash flow news in stock returns and that its importance increases with the investment horizon. They do not use predictive regressions but back out cash flow news by using a return decomposition based on the implied cost of equity capital and do not discuss dividend growth predictability. van Binsbergen et al. (2013) show that equity yields are able to predict dividend growth, but this does not address the dividend growth predictability puzzle.[6] Golez (2014) adjusts the standard dividend-price ratio for changes in expected dividend growth using estimates implied by the derivatives market and finds that dividend growth is predictable by an implied dividend growth rate, but not by the unadjusted dividend-price ratio. Maio and Santa-Clara (2014) find that dividend growth is predictable using the standard dividend-price ratio in a predictive regression framework only for small and value stocks, but not for large and growth stocks. International evidence on dividend growth predictability is discussed in Engsted and Pedersen (2010) and Rangvid et al. (2014).

This chapter is also closely related to the work of Robertson and Wright (2006) and Boudoukh et al. (2007). Robertson and Wright (2006) use aggregate total payout data from the Federal Reserve Financial Accounts (Flow of Funds) tables, but only look at return predictability. Moreover, differently from this chapter, they do not provide a split of the total payout measure into each individual component, which allows us to understand the impact of M&A cash dividends on predictability. Boudoukh et al. (2007) show that taking into account repurchases and issues results in stronger return predictability. However, they do not discuss dividend growth predictability nor take into account M&A cash dividends in their analysis.

The remainder of the chapter is organized as follows. Section 1.2 presents a cash flow measure inclusive of M&A cash dividends and reports descriptive statistics. Section 1.3 provides details of M&A cash flows, repurchases and equity issues and sum-

[6] The equity yields are a linear function of expected dividend growth and maturity-specific risk premia.

marizes possible concerns related to the use of CRSP-extracted dividends. Section 1.4 reports empirical results of dividend growth, consumption growth and return predictive regressions at the aggregate level, long run predictability, predictability during M&A waves and predictability of individual cash flow components. It also compares my dividend measure against other payout measures that have been proposed in the literature. Section 1.5 discusses predictability at the industry level. Section 1.6 presents a trading strategy that exploits my dividend-price ratio inclusive of M&A cash dividends to improve out-of-sample portfolio returns and Sharpe ratios. Section 1.7 concludes.

1.2 A New Dividend Measure that Includes Cash from M&A

Given the shortcomings of the existing dividend measures, I construct a novel measure of dividends that, in addition to ordinary and liquidation dividends, takes into account M&A cash flows received by shareholders. At the aggregate level, I get data on ordinary and liquidation dividends[7] by looking at the distributions paid by the individual issues in the CRSP dataset.[8] Data for M&A cash dividends are from CRSP before 1980 and from the SDC Platinum dataset[9] after 1980. CRSP cash M&A data include all dividends with distributions codes between 3000 and 3400. Using SDC Platinum data, I aggregate by calendar year or quarter[10] all cash dividends generated by M&A transactions financed completely or partially with cash (e.g., both cash-only and cash plus stock) having US public buyer and US target.[11] The SDC dataset allows us to take also into account the cash reserves that US listed companies utilize to acquire non-listed companies, which are omitted from CRSP, but still constitute a substantial chunk of the US aggregate economy and shareholders' cash flows. Finally, using this new measure of dividends I re-construct the dividend growth rates and dividend-price ratio[12] that are used in the following sections of the chapter. Real data are constructed using the Consumer Price Index-All Urban Consumers (CPI-U) found in the Federal Reserve Economics Dataset (FRED) to deflate nominal dividends. The market excess return is calculated by subtracting the risk free rate in the Goyal and Welch (2008) dataset from the CRSP value-weighted cum dividend index return (e.g., WVRETD) that

[7] CRSP distribution codes 1xxx, 2xx2/2xx8.

[8] Results using only common stocks (share codes 10 and 11) instead of the whole CRSP universe are qualitatively the same.

[9] SDC is the standard data source on mergers. It is used, for example, in the Payout Policy handbook chapter by Allen and Michaely (2003). Results using CRSP-only M&A data are qualitatively the same.

[10] I select the cash flows distributed when the deal status is completed and not at announcement date.

[11] This is to ensure consistency with the CRSP database that only includes listed US companies.

[12] I use the total market cap of the stocks in CRSP at the end of December.

includes stocks in NYSE, AMEX, NASDAQ. Consumption expenditure is the sum of non durable consumption plus services from Table 2.3.5[13] of the National Income and Product Accounts (NIPAs), available on the Bureau of Economic Analysis (BEA) website, as it is standard in the literature. Gross National Product data used in Section 1.4.4 are from FRED. For the industry section, I construct my dividend and return dataset using the 10 industry definitions of Fama and French using M&A cash dividends from CRSP pre-1980 and SDC post-1980.

I compare my measure with the two standard dividend measures extracted from the CRSP value-weighted index. The first is extracted from the CRSP value-weighted index at the annual frequency, and it is affected by the reinvestment of the dividends in the market. The second is constructed by summing up the monthly dividends extracted from the monthly CRSP value-weighted index and it is not subject to the reinvestment problem. Table 1.1 reports descriptive statistics and correlations of dividend growth measures over the full sample, pre-war and post-war. Figure 1.2 shows the dividend growth series graphically.

Looking at at the table, we see that adding M&A cash dividends increases both the mean and volatility of the dividend growth rate over the various sample periods. This is expected, since M&A activity is volatile and tends to happen in waves (Harford (2005)). We also note that pre-war most of the statistics are very similar, while post-war they tend to be different. This is because M&A cash dividends become substantial, both in absolute amounts and relative to ordinary cash dividends, only post-war. The full-sample correlation between the CRSP dividend growth measures with and without the reinvestment issue is only 42%, as already discussed by Chen (2009). The contemporaneous correlation between the dividend growth measures and stock returns is positive but small for my dividend measure (between 5% and 7% across sub-samples) whereas it

[13] Personal Consumption Expenditures by Major Type of Product.

is above 60% for the measure with the reinvestment of dividends. These facts highlight the importance of reinvestments. Time variation in the correlation of dividend growth with stock returns is, however, limited for all three dividend measures.

The correlation between the dividend growth measure without reinvestment and the one that includes M&A cash dividends is 73% over the full sample and 55% post-war, confirming that M&A cash flows became relevant over the last 40 years. Most interestingly, the AR(1) coefficient of the dividend growth measure with reinvestment is negative, while slightly positive for the two other measures. This is due to the fact that the AR(1) of the returns series is negative post-war (-0.07) and therefore impacts the dividend growth dynamics when dividends are reinvested in the market.

Looking at the statistical properties of the dividend-price ratios (Table 1.2 and Figure 1.3), we see that adding M&A cash dividends to the dividend-price ratio slightly increases its mean (from 3.75% to 4.13%) and decreases its volatility (1.58% to 1.35%). The correlation with the standard dividend-price ratios is around 70% over the various sample periods. However, the most important result is the relatively small autocorrelation of my dividend-price ratio, 0.67 (0.71) over the full sample (post-war), far from the very high estimates found in the literature (e.g., Ang and Bekaert (2007), Lettau and Nieuwerburgh (2008), Cochrane (2008), Cochrane (2011)). As a consequence, its limited persistence should alleviate some of the problems with statistical inference of predictive regressions in the presence of near-integrated regressors (Stambaugh (1999), Campbell and Yogo (2006), Moon and Velasco (2014)).

Table 1.1: Dividend growth: descriptive statistics. **Panel A**: full sample statistics (1926-2012). **Panel B**: pre-war sample statistics (1926-1945). **Panel C**: post-war sample statistics (1945-2012). The table reports the mean (column 1), standard deviation (column 2), minimum (column 3), maximum (column 4), skewness (column 5), kurtosis (column 6), first-order autoregressive coefficient AR(1) (column 7), correlation between dividend growth and stock returns (column 8) and the correlations amongst the three measures. Annual nominal data.

Panel A: 1926-2012											
	mean	*std*	*min*	*max*	*skew*	*kurt*	*AR(1)*	$\rho(dg, ret)$	*correlations*		
1. Dividends with reinvestment	1.06	0.15	0.69	1.45	0.31	0.28	-0.09	0.62	1		
2. Dividends without reinvestment	1.05	0.12	0.60	1.53	-0.50	5.85	0.28	0.09	0.42	1	
3. Dividends inclusive of M&A cash	1.09	0.17	0.62	1.60	-0.14	1.09	0.13	0.07	0.26	0.73	1

Panel B: 1926-1945											
	mean	*std*	*min*	*max*	*skew*	*kurt*	*AR(1)*	$\rho(dg, ret)$	*correlations*		
1. Dividends with reinvestment	1.01	0.18	0.69	1.40	0.29	0.28	0.26	0.67	1		
2. Dividends without reinvestment	1.02	0.21	0.60	1.53	-0.06	1.18	0.25	0.11	0.50	1	
3. Dividends inclusive of M&A cash	1.05	0.22	0.62	1.55	-0.10	0.63	0.22	0.05	0.46	0.97	1

Panel C: 1946-2012											
	mean	*std*	*min*	*max*	*skew*	*kurt*	*AR(1)*	$\rho(dg, ret)$	*correlations*		
1. Dividends with reinvestment	1.07	0.14	0.75	1.45	0.47	0.29	-0.28	0.61	1		
2. Dividends without reinvestment	1.06	0.07	0.83	1.30	0.56	2.32	0.31	0.06	0.39	1	
3. Dividends inclusive of M&A cash	1.10	0.15	0.75	1.60	0.08	0.96	0.05	0.07	0.10	0.54	1

Table 1.2: Dividend-price ratio: descriptive statistics. **Panel A**: full sample statistics (1926-2012). **Panel B**: pre-war sample statistics (1926-1945). **Panel C**: post-war sample statistics (1945-2012). The table reports the mean (column 1), standard deviation (column 2), minimum (column 3), maximum (column 4), skewness (column 5), kurtosis (column 6), first-order autoregressive coefficient AR(1) (column 7) and the correlations amongst the three measures. Annual data.

Panel A: 1926-2012

	mean	*std*	*min*	*max*	*skew*	*kurt*	*AR(1)*	*correlations*		
1. Dividends with reinvestment	3.88%	1.49%	1.11%	7.23%	0.31	-0.41	0.90	1		
2. Dividends without reinvestment	3.75%	1.58%	1.14%	9.47%	0.79	1.06	0.79	0.94	1	
3. Dividends inclusive of M&A cash	4.15%	1.35%	2.16%	9.70%	1.16	1.98	0.67	0.73	0.84	1

Panel B: 1926-1945

	mean	*std*	*min*	*max*	*skew*	*kurt*	*AR(1)*	*correlations*		
1. Dividends with reinvestment	5.15%	1.04%	3.83%	7.23%	0.76	-0.46	0.48	1		
2. Dividends without reinvestment	5.15%	1.61%	3.32%	9.47%	1.01	0.90	0.35	0.79	1	
3. Dividends inclusive of M&A cash	4.89%	1.64%	3.01%	9.70%	1.25	2.00	0.38	0.79	0.99	1

Panel C: 1946-2012

	mean	*std*	*min*	*max*	*skew*	*kurt*	*AR(1)*	*correlations*		
1. Dividends with reinvestment	3.50%	1.40%	1.11%	7.16%	0.56	-0.02	0.94	1		
2. Dividends without reinvestment	3.33%	1.31%	1.14%	6.51%	0.42	-0.34	0.89	0.98	1	
3. Dividends inclusive of M&A cash	3.93%	1.16%	2.16%	6.84%	0.66	-0.59	0.71	0.70	0.75	1

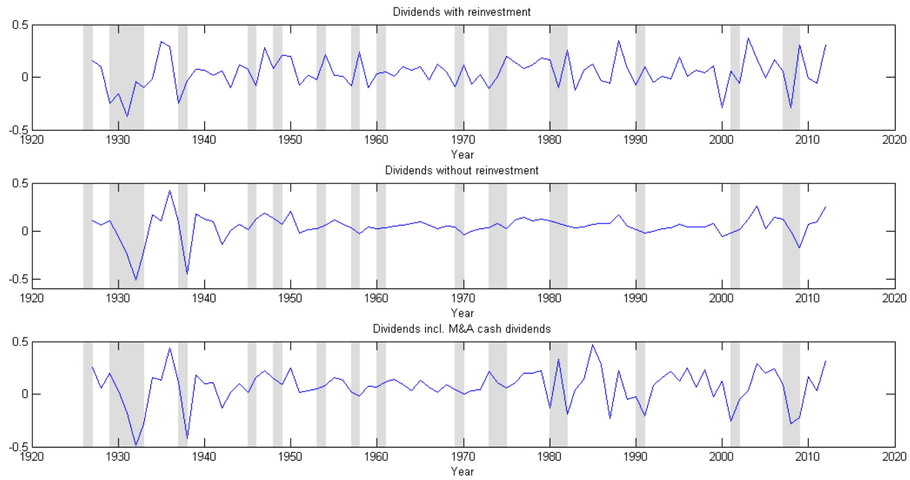

Figure 1.2: Dividend growth rates (logs)

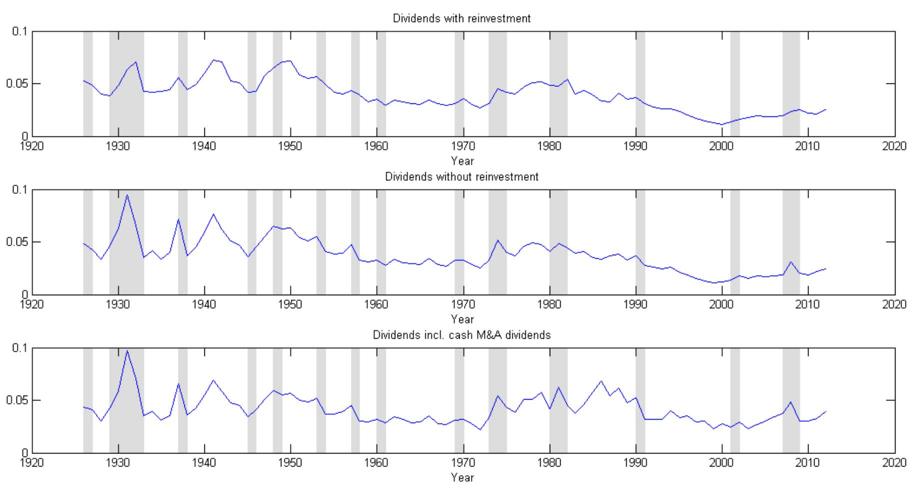

Figure 1.3: Dividend-price ratios

1.3 Other Cash Flow Measures

Distribution events occurring during a company's lifecycle can be divided into ordinary and non ordinary ones. Ordinary distributions are usually considered to be recurring cash dividends[14] and some liquidation dividends.[15] Non ordinary distributions include cash and stock dividends originating from an acquisition or reorganization of a company, subscription rights, stock dividends of any kind and buybacks and issuances. Which cash flows to include in the dividend measure has a substantial impact on dividend growth and return (when using the dividend-price ratio as predictor) predictability. Most authors (e.g., Chen (2009), van Binsbergen and Koijen (2010)) hold the traditional "shareholder view," which implies that the relevant cash flows are those effectively received by the individual shareholders, in contrast to the recent "macro" perspective where aggregate economy's cash flows, such as repurchases and issuances of equity or debt, are also included (Robertson and Wright (2006), Boudoukh et al. (2007), Larrain and Yogo (2008)).

1.3.1 The role of M&A cash dividends

The present value relationship of Campbell and Shiller (1988a) suggests that the price of a security today is the infinite sum of expected future dividends discounted at a rate that can be time-varying. Most studies use only ordinary dividends and find dividend growth to be unpredictable by the dividend-price ratio (Cochrane (2008)). As a consequence, some authors (Robertson and Wright (2006), Boudoukh et al. (2007), Larrain and Yogo (2008)) recently proposed to use aggregate cash flows.

Critics of the macro view argue that buybacks and issues are cash flows at the

[14] CRSP distribution codes 1xxx.

[15] CRSP distribution codes 2xxx

aggregate economy level, but not necessarily for the individual investors.[16] That is, fewer or more shares today only affects the future dividend per share amount in the present value equation.

Allen and Michaely (2003) note that both measures, either using only ordinary dividends or including repurchases and issues, are incomplete. In fact, shareholders also receive cash payouts and stock dividends from corporations through mergers and acquisitions that are accomplished through cash or stock transactions. That is, shareholders of the acquired firms receive either a cash or stock payment that can be viewed as a liquidating or final dividend. This amount is non trivial and it is time-varying. Allen and Michaely (2003) state that "these types of liquidating dividends seem to have a significant weight in the aggregate payout of US corporations. For example, in 1999, proceeds from cash M&As were more than the combined cash distributed to shareholders through dividends and repurchases combined...this component of payout has been largely ignored in the literature. Over the last decade such (M&A cash) payments have been around $240bn per year, or over 50% of aggregate payout if we also include dividends and repurchases. Measuring and understanding this component of payout policy is an important task for future research."

Figure 1.4 shows the dynamics of an M&A transaction financed only with cash. It is clear that cash financed mergers are like a liquidating dividend for the target shareholders. Figures 1.5 and 1.6 show how M&A transactions can be financed with buybacks and stock issues, respectively. Figure 1.5 shows that the compensation paid to target investors often comes directly from the existing assets of the buyer firm, directly in the form of stocks to the target shareholders or indirectly in the form of cash to the buyers' shareholders.[17] The choice of how to finance an M&A transaction is a function

[16] Remaining shareholders receive compensation in the form of capital appreciation and are not required to participate in a secondary equity offering.

[17] This holds unless the acquiring company issues debt or stock, receives the cash and pay target investors with that cash.

of other variables (e.g., tax regime, ownership control, etc.) and is beyond the scope of this chapter. In conclusion, M&A cash dividends are an important source of investors' cash flows.

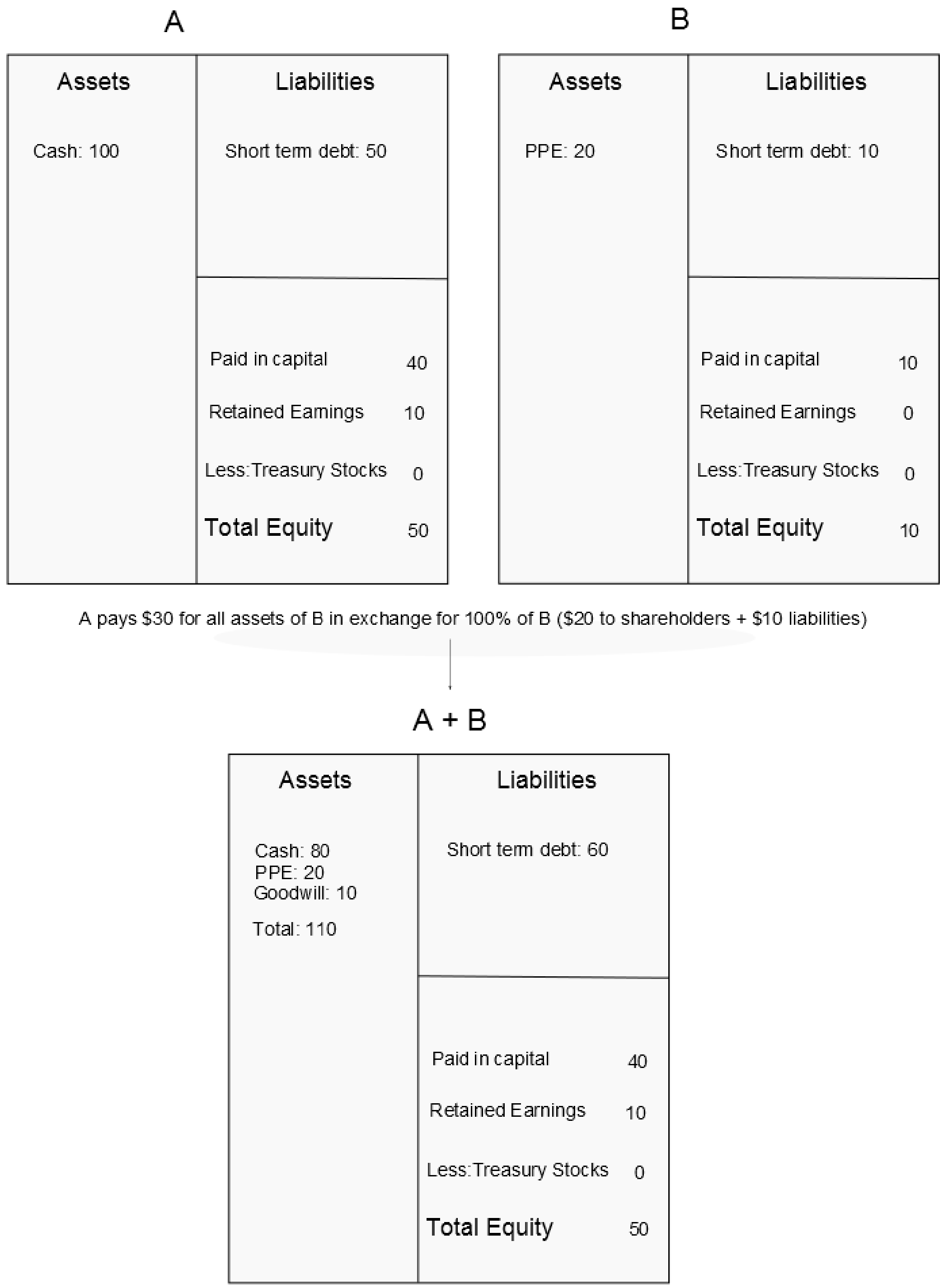

Figure 1.4: Dynamics of a M&A transaction financed only with cash.

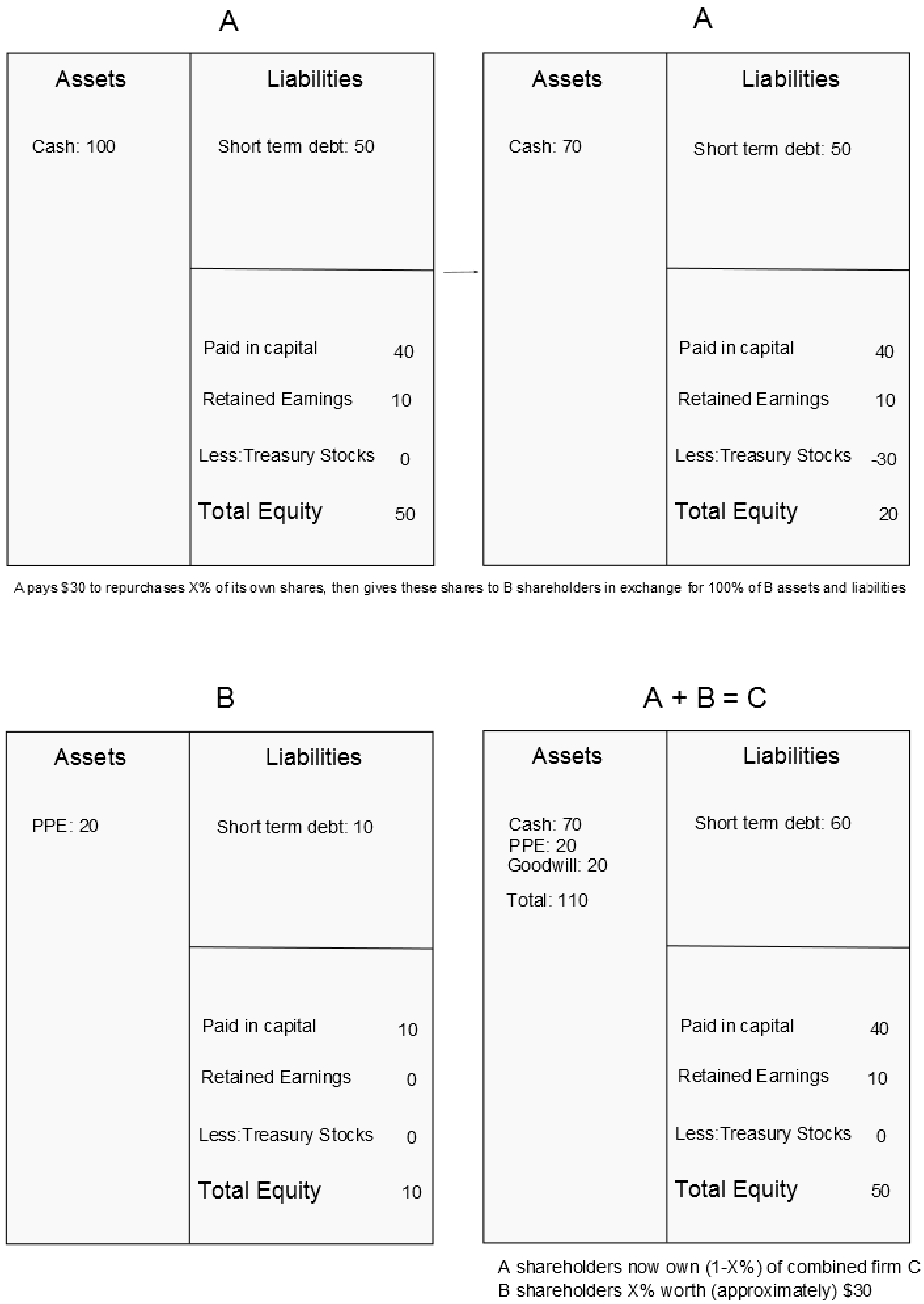

Figure 1.5: Dynamics of a stock M&A transaction financed with buybacks.

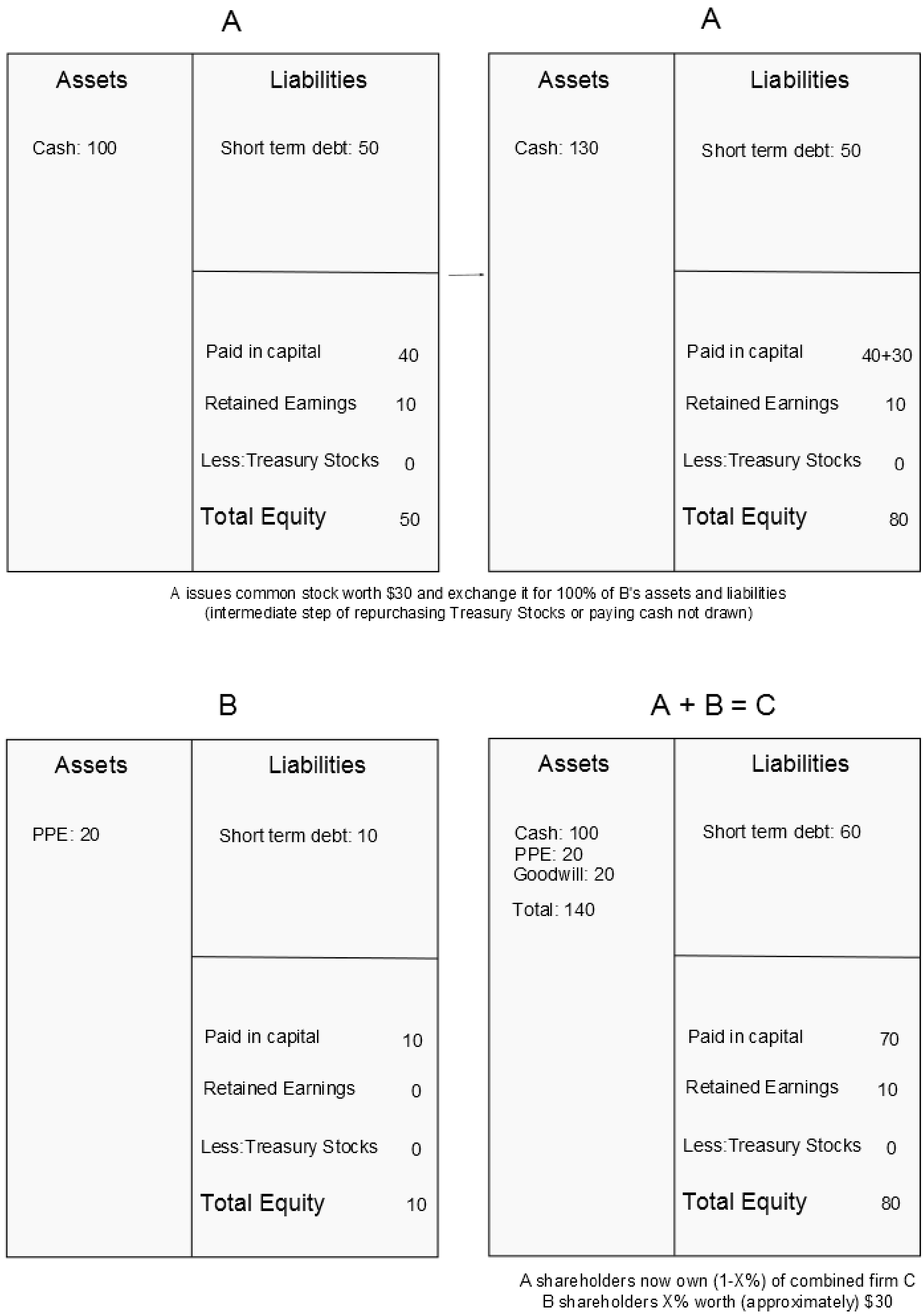

Figure 1.6: Dynamics of a M&A transaction financed by issuing stock.

1.3.2 Stock M&A cash flows versus repurchases and issues

In this section I discuss the empirical evidence on M&A stock dividends, repurchases and issues.

M&A stock dividends are also cash flows at the aggregate level if the acquiring company first repurchases its own shares and reissues them to target shareholders (see Figure 1.5). As discussed in Fama and French (2001), "share repurchases are often treated as non-cash dividends, but this is not the case when repurchased stock is reissued to the acquired firm in a merger," in the form of stock dividends to the target shareholders. Shareholders of the acquiring company do not put up any additional cash, the firm's cash is used to repurchase shares that are then given to the shareholders of the target company, who ultimately get a stock dividend that can be liquidated to a new stock market participant netting a positive cash flow.[18]

Fama and French (2001) report that cash dividends are disappearing and that repurchases are often used to finance stock M&A deals (pages 35-37). Boudoukh et al. (2007) propose an aggregate cash flow measure that includes repurchases and issues and suggest that repurchases might have been substituting cash dividends over the last 20 years following the institution of SEC rule 10b-18 in 1982 and for tax reasons.[19]

Stock issues and their cash flow implications are also worth discussing. If a stock issue happens today, there is a negative cash flow at the aggregate shareholder level. However, if these stocks' cash flows are later used to finance an M&A transaction, then the "net" aggregate shareholders' cash flows are zero. This net effect is not caught in the net payout measure of Boudoukh et al. (2007), which only considers the aggregate value of issues.[20] This implies that only the net issues "kept within the company" (that is, not

[18] The only exception is if the acquirer finances an acquisition of a target company by issuing stocks reissued as stock dividend to the target shareholders. In this case the net aggregate effect would be null (see Figure 1.6).

[19] The Tax Reform Act was enacted in 1986.

[20] Boudoukh et al. (2007) acknowledge that "the drawback of this measure is that it captures issuances

used to finance M&A activities later on) are cash flows at the aggregate shareholder level and should be taken into account.

Moreover, a measure of dividends that takes into account repurchases and issues, such that used by Boudoukh et al. (2007), but not M&A cash dividends, is sometimes negative (e.g. that is, in some years or months the net payout value is negative) and therefore cannot be used to calculate dividend growth or a dividend-price ratio.[21]

The implications of these facts is that asset pricing tests employing incomplete measures of cash distributions to shareholders are less likely to accurately capture economic effects if M&A cash dividends, a fundamental source of individual shareholders' and aggregate cash flows, are ignored in the dividend-price ratio and dividend growth measures.

I next review the standard methodology used to calculate dividend data and point out why M&A cash flows are excluded.

1.3.3 Methodology and dividends for CRSP index

Several authors use the CRSP NYSE-AMEX-NASDAQ value-weighted market index to extract aggregate dividends. Most researchers extract dividends "top-down" from the CRSP index as the difference between the cum dividend return (VWRETD) minus the ex dividend return (VWRETX) multiplied by the previous ex dividend index level. That is, $D_t = (R_t^{cum} - R_t^{ex}) \times P_{t-1}^{ex}$. Some studies (e.g., Campbell and Shiller (1988b), Goyal and Welch (2003), Ang and Bekaert (2007), Chen (2009)) first calculate CRSP monthly dividends and sum them up to construct an annual measure. Others (e.g., Cochrane (1992), Boudoukh et al. (2007), Cochrane (2008), Koijen and Nieuwerburgh

not generating cash flows (e.g., acquisitions and stock grants)."

[21] Boudoukh et al. (2007) adjust the npy by adding 0.1 to the yield.

(2011)) back out CRSP dividends using directly the annual index.

However, it is important to notice a few points. First, the CRSP value-weighted index is a value-weighted portfolio built using all issues listed on the NYSE-NASDAQ-AMEX exchanges, except ADRs.[22] It is therefore a proxy for the overall market index. This implies that the constituents of the index are not only US common stocks, but other securities such as certificates, SBIs, units, ETFs and closed-end funds listed on those exchanges are also included. In fact, around 12.55% (474,894) of the total monthly observations (3,782,752) from 1926 up to 2012 downloaded from the CRSP stock files database, excluding ADRs and issues listed on ARCA, refer to non-US common stocks (share codes 10 and 11). More precisely, 4% of the observations refer to common shares of non-US companies (share code 13), 3% refer to closed-end funds and unit investment trusts (share code 14), 3% refer to SBIs and Units (share codes 4x and 7x, except 73), around 1% to ETFs (share code 73), 0.2% to closed-end fund companies incorporated outside the US (share code 15), around 1% refer to REITs, 0.2% refer to certificates. These non common stock issues average 5.30% of the CRSP market capitalization over the full sample period, but over the last decade their weight increased to around 17%, a non trivial amount.

As a consequence, the dividend measure extracted from the index is a noisy proxy for common stock dividends, resembling more a measure of the aggregate dividends within the CRSP universe. Moreover, this is potentially biased because a security that contains a common stock (e.g., ETFs) and receives a dividend from it, does not necessarily distribute the full gross amount of that dividend back to the investors.[23]

[22] `http://www.crsp.com/documentation/product/stkind/index_methodologies/stock_file_indices.html`.

[23] As an example, the ProShares Ultra S&P500 2x leveraged ETF (ticker: SSO) is included in the CRSP database (permno: 91307). As stated on the company website, only investment income and capital gains, net of expenses, are distributed. Moreover, also short ETFs such as the ProShares Short SmallCap600 (ticker: SBB) are present in CRSP (permno: 91717). This will result in even more biased distribution amounts.

Second, and most importantly, dividends extracted from the CRSP value-weighted index only include ordinary and liquidation dividends.[24] This implies that a series of cash flows, such as cash dividends received during a takeover, are not taken into account in the dividend measure. More precisely, the returns of the cum dividend and ex dividend CRSP value-weighted indices are the same in those non-ordinary distribution cases. However, ignoring such sources of dividends, especially in recent years where cash dividends have been substituted by other forms of cash flows (Boudoukh et al. (2007)) is hardly justifiable.

Lastly, as highlighted by Chen (2009), there is the issue of reinvestment of dividends that appears when extracting dividends from the CRSP index at the annual frequency. CRSP computes quarterly or annual return series by reinvesting the dividends at the cum-dividend stock market return. As a consequence, the dividend amounts extracted by the CRSP annual index are contaminated by the return dynamics and the dividend growth measure inherits the behavior of stock returns. Given the fact that returns are more volatile than dividends, dividend growth predictability is severely impacted. As noted by Koijen and Nieuwerburgh (2011), the discrepancy in the dividend growth and dividend-price ratio series resulting from different reinvestment assumptions is substantial.

1.4 Predictive Regressions

Campbell and Shiller (1988a) derive the present value identity

$$dp_t = c + E_t\left[\sum_{j=1}^{\infty} \rho^{j-1}(r_{t+j})\right] - E_t\left[\sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j})\right] \tag{1.1}$$

[24] Specifically, only CRSP distribution codes 1xxx, 2xx2/2xx8 and, if available, 6xx2/6xx8 are included. CRSP also requires all these entries to have a factor to adjust price (facpr) equal to 0 or -1.

where $r_{t+j} = log\left[\frac{P_{t+j}+D_{t+j}}{P_{t+j-1}}\right]$ is the log return, $\Delta d_{t+j} = log\left[\frac{D_{t+j}}{D_{t+j-1}}\right]$ is log dividend growth rate, $dp_t = log\left[\frac{D_t}{P_t}\right]$ is the log dividend-price ratio and c is a constant. Equation (1.1) is the theoretical justification for why the dividend-price ratio should predict expected returns or dividend growth rates (or both), and it can be interpreted as a dynamic generalization of the constant dividend-price ratio in the Gordon model.

1.4.1 Dividend and consumption growth predictability

I estimate standard dividend and consumption growth predictability regressions

$$\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{dg,t+1} \tag{1.2}$$

$$\Delta c_{t+1} = a_c + b_c(d_t - p_t) + \varepsilon_{cg,t+1} \tag{1.3}$$

where $\Delta d_{t+1} = ln(D_{t+1}/D_t)$ is the real[25] log dividend growth, D_{t+1} is the dividend amount at time $t+1$, Δc_{t+1} is the real log consumption growth, $d_t - p_t = ln(D_t/P_t)$ is the log dividend-price ratio and P_t is the price at time t.

Panel A in Table 1.3 and Table 1.4 reports the results of regression (1.2) at annual and quarterly frequencies, respectively. Panel C reports the results of the consumption growth regression (1.3).[26]

In Panel A, the dividend growth and dividend-price ratio are always constructed using the same dividend specification (e.g., if the dividend growth includes M&A cash dividends, so will the dividend-price ratio). Over the full sample we see that there is no predictability using the standard dividend measures used in the literature (measures 1 and 2). However, adding M&A cash dividends to both dividend growth and the

[25] The regressions in this section based on nominal data yield very similar results.

[26] Panels A-B-C of Table 1.5 show the results of the same regressions using public-only M&A cash dividends. The results are qualitatively the same, except for return predictability, which is slightly weaker.

dividend-price ratio results in strong predictability. Contrary to Chen (2009)[27], I find predictability post-war. The adjusted R^2 of my measure is always higher than those of the other dividend measures and the coefficients always significant. It seems therefore that the absence of dividend growth predictability reported so far, both pre- and post-war, reflects the incompleteness of the standard measures of cash flow employed in the literature which do not take into account M&A cash dividends.[28] Lastly, the sign on the coefficient of the dividend-price ratio of my measure inclusive of M&A cash dividends is, as it should be, negative. This is not the case with the coefficients of the first measure where dividends are reinvested.

As far as consumption growth predictability is concerned, we see that it is substantially predictable by my dividend-price ratio. I find an adjusted R^2 of 11.77% over the full sample and 6.75% (10.03%) post-war at the annual (quarterly) frequency with statistically significant coefficients, in contrast with both the standard dividend-price ratios. This is consistent with the predictions of the long-run risk asset pricing model (Beeler and Campbell (2012), Bansal et al. (2012), Jagannathan and Marakani (2015)).

In conclusion, these results suggest that M&A cash dividends help solve the dividend growth predictability puzzle described by Cochrane (2008). Moreover, they have substantial explanatory power for aggregate consumption growth. Taking into account these dividends results in a more complete measure of actual cash flows received by investors. Not considering cash dividends from mergers and acquisitions reduces predictability. These dividends are important not only because their amount is cyclical and non trivial in recent years, when their relative proportion of total aggregate cash flows becomes relevant, but especially because the presence of dividend growth predictability implies that cash flow news help explain the volatility of stock prices. Stock price vari-

[27] His sample ends in 2005 and misses the market turmoil of 2007-2011 where cash flow amounts have been subject to substantial shocks due to the financial crisis.

[28] Table 1.5 Panel D reports the regression of the standard dividend growth on my dividend-price ratio inclusive of M&A cash dividends. Results are now significant post-war.

ation, often thought to be only justified by discount rate news (e.g., Cochrane (2011)), can thus be traced back to cash flow news. I explore this issue more in detail in Section 1.4.3, which discusses the fraction of the dividend-price ratio variance that can be attributed, respectively, to time-varying expected returns and dividend growth.

Out-of-sample analysis

Following Goyal and Welch (2008) and Ferreira and Santa-Clara (2011), we generate real-time, out-of-sample forecasts of dividend growth using a sequence of expanding estimation windows.[29] More precisely, we take a subsample of the first s observations $t = 1, 2, ..., s$ of the entire sample of T observations and estimate our dividend growth regression (1.2). We denote the expected dividend growth conditional on time s information by $g_{s+1} = E_{s+1|s}(\Delta d_{s+1})$. We then take the coefficients $\hat{a}^i_s$, $\hat{b}^i_s$ estimated using information available up to time s and predict the dividend growth at time $s+1$:

$$\hat{g}_{s+1} = \hat{a}^i_s + \hat{b}^i_s(d^i_s - p_s) \tag{1.4}$$

where $i = 1, 2, 3$ indicate the cash flow measures described in the previous sections.

We follow this process for $s = s_0, ..., T-1$, generating a sequence of out-of-sample dividend growth forecasts. In order to start the procedure, we set the initial sample size s_0 equal to half of the full sample (e.g., 45 years[30]). Using out-of-sample forecasts based on previously available information replicates what a forecaster could have done in real time. We evaluate the performance of the out-of-sample forecasts

[29] The literature focuses on out-of-sample excess returns predictability, but the methodology can be applied to dividend growth as well.

[30] We choose $s_0 = 45$ (e.g. 1926-1971) in order to have enough initial data to get a reliable estimate at the start of evaluation period (45 years, approximately half of the sample, from 1966-2012) and to be consistent with Goyal and Welch (2008) who also choose an evaluation period of half the sample in their out-of-sample analysis.

through the out-of-sample (OOS) R^2:

$$OOS\ R^2 = 1 - \frac{MSE_P}{MSE_M}, \tag{1.5}$$

where MSE_P is the mean square error of the out-of-sample predictions from the model:

$$MSE_P = \frac{1}{T - s_0} \sum_{s=s_0}^{T-1} (\Delta d_{s+1} - \hat{g}_{s+1})^2 \tag{1.6}$$

and MSE_M is the mean square error of the historical sample mean:

$$MSE_M = \frac{1}{T - s_0} \sum_{s=s_0}^{T-1} (\Delta d_{s+1} - \Delta \bar{d}_s)^2 \tag{1.7}$$

where $\Delta \bar{d}_s$ is the historical mean of dividend growth up to time s. This out-of-sample R^2 is positive (negative) when the model with the lagged dividend-price ratio predicts dividend growth better (worse) than the historical average of dividend growth. We evaluate the statistical significance of the results using the MSE-F statistic proposed by McCracken (2007),

$$MSE - F = (T - s_0)(\frac{MSE_M - MSE_P}{MSE_P}) \tag{1.8}$$

which tests for the equality of the two MSEs and takes into account nested forecast models. Table 1.6 (top panel) reports the results. We see that the only measure with strong out-of-sample dividend growth predictability is the one that takes into account M&A cash dividends. OOS R^2 values are positive and equal to 6.82% (1.73%) at annual (quarterly) frequency for that measure, while those of the other measures are all negative. Moreover, the MSE-F statistic is 3.00 (3.02), significant at the 5% level. Using other values for s_0 as a robustness check results in even stronger out-of-sample predictability. Figures 1.7 and 1.8 (top panel) show the cumulative squared prediction errors of the prevailing mean minus the cumulative squared prediction errors of the alternative model (Equation (1.4)) at the annual and quarterly frequency, respectively. We see that using my measure the out-of-sample performance is always strong except during the period 2000-2005 at quarterly frequency, where the historical mean performs better. The performance of the standard dividend measures are flat or negative, confirming that the alternative model does not improve on a forecast based on the prevailing mean for those measures. Overall, this confirms the in-sample results and suggests that M&A cash dividends are a fundamental component of aggregate cash flows both statistically and economically.

In conclusion, using my M&A cash dividends in the construction of both dividend growth and the dividend-price ratio, we have strong dividend growth predictability both in-sample and out-of-sample, in contrast to the results obtained using standard dividend measures.

1.4.2 Return predictability

We estimate the following standard return predictive regression:

$$r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{r,t+1} \tag{1.9}$$

where r_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R^{CRSP}) - ln(R_f)$, with R_{t+1}^{CRSP} is the CRSP cum dividend gross return (e.g., VWRETD) and R_f is the gross risk-free rate.) and $d_t - p_t$ is the lagged dividend-price ratio constructed using the various dividend measures. Annual (quarterly) results are reported in Panel B of Table 1.3 (Table 1.4).

Over the full sample my predictor has an annual (quarterly) adjusted R^2 of 4.40% (1.59%), the highest, but it is post-war, when M&A cash dividends become substantial, that we have the largest predictability with an adjusted (annual) R^2 of 10.62% (2.83%), almost 5% (1.5%) more than those of the standard dividend measures. All coefficients are statistically significant at the 5% level. As already discussed by Chen (2009), there is no return predictability pre-war. I also calculate the Stambaugh (1999) coefficient bias for my cash M&A measure to avoid possible criticisms related to statistical inference in the presence of near-integrated regressors (e.g., Ferson et al. (2003), Campbell and Yogo (2006), Moon and Velasco (2014)). The bias is extremely small (0.02) because both the autocorrelation of my cash M&A dividend-price ratio and the error covariance (correlation) are small, 0.67 (Table 1.2) and -0.033 (-0.75), respectively, over the full sample.

Out-of-sample analysis

We follow the same procedure adopted in the out-of-sample analysis of dividend growth[31] to forecast the excess return equation (1.9).[32] Results are reported in Table 1.6 (bottom panel). We see that out-of-sample my dividend measure is the only one that is strongly significant with a MSE-F statistic of 3.44 (3.79) and with an out-of-sample R^2 of 7.73% (2.16%) at annual (quarterly) frequency. The standard dividend-price ratios perform badly out-of-sample, as already reported by Goyal and Welch (2008). It is important to note that my dividend measure beats both measures out-of-sample using any initial estimation sample s_0.

Figures 1.7 and 1.8 (bottom panel) show the cumulative squared prediction errors of the prevailing mean minus the cumulative squared prediction errors of the alter-

[31] A vast literature discusses out-of-sample excess returns predictability. See, for example, Goyal and Welch (2008), Campbell and Thompson (2008), Rapach et al. (2010), Cenesizoglu and Timmermann (2012), Johannes et al. (2014).

[32] Campbell and Thompson (2008) state that all forecasts tend to underpredict returns when log returns are used. This suggests that our forecasts are conservative.

native model (1.9). We see that my M&A cash dividend measure does not undergo any dry spells with poor out-of-sample performance, in constrast with the standard dividend measures which have an extremely weak out-of-sample performance over the period 1994-1999. All measures perform well from 1999 to 2003, but only my cash M&A measure performs well after that. Goyal and Welch (2008) (p. 1456) report that out-of-sample predictability is dependent on years up to and especially during the years of the Oil Shock 1973-1975. We see that this is only partially true for my predictor. In fact, both at annual and quarterly frequency, the alternative model performs very well during the period 1998-2005. Goyal and Welch (2008) and Campbell and Thompson (2008) note that the ability of valuation ratios to forecast stock returns is weakest in the period post-1980, which includes the great equity bull market at the end of the twentieth century. My measure performs best over that period, suggesting that today merger activity has a strong impact on future stock returns.

In conclusion, using my M&A cash dividends in the construction of the dividend-price ratio results in strong evidence of return predictability, both in-sample and out-of-sample.

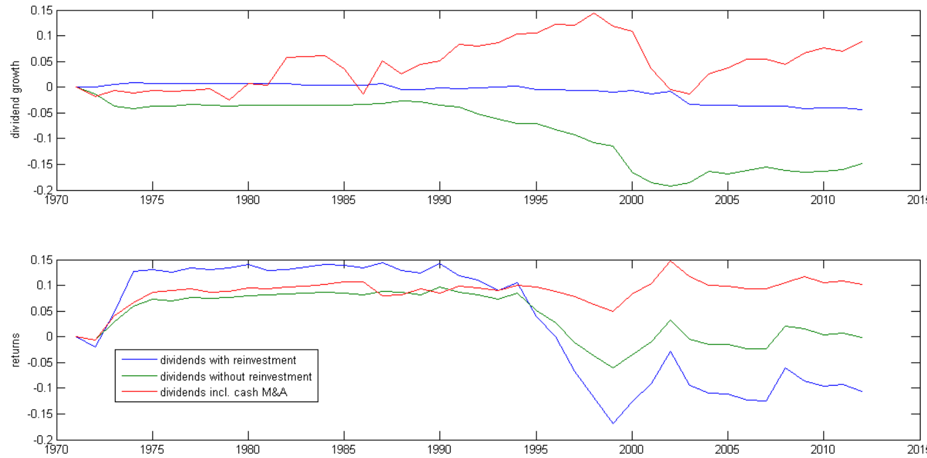

Figure 1.7: Out-of-sample dividend growth and excess return predictability (annual). This figure plots the cumulative squared prediction errors of the null minus the cumulative squared prediction error of the alternative. The alternative are models (1.4) and (1.9). The null is the prevailing mean.

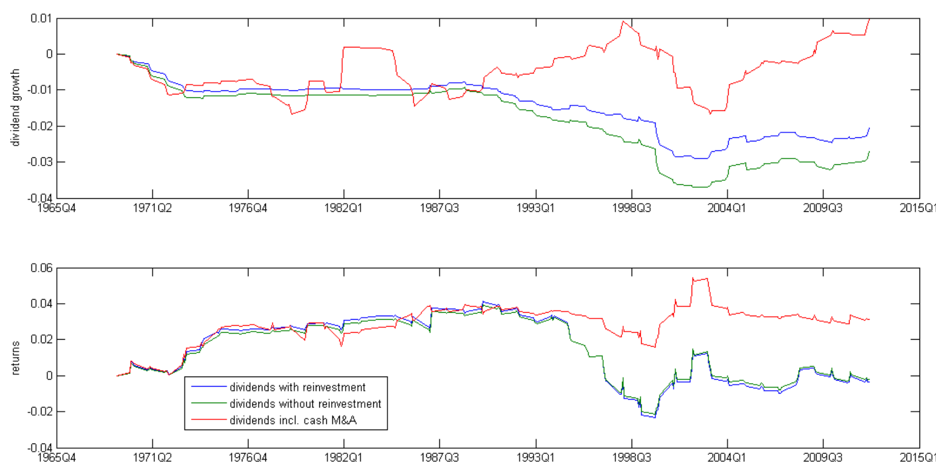

Figure 1.8: Out-of-sample dividend growth and excess return predictability (quarterly). This figure plots the cumulative squared prediction errors of the null minus the cumulative squared prediction error of the alternative. The alternative are models (1.4) and (1.9). The null is the prevailing mean.

Table 1.3: Predictive regressions of excess returns, dividend and consumption growth at annual frequency. **Panel A**: Δd_{t+1} is the real dividend log growth between t and $t+1$, $d_t - p_t$ is the log dividend-price ratio. **Panel B**: r_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R_{t+1}^{CRSP}) - ln(R_f)$), $d_t - p_t$ is the log dividend-price. R_{t+1}^{CRSP} is the CRSP cum dividend gross return (VWRETD), R_f is the gross risk-free rate. **Panel C**: Δc_{t+1} is the real (log) consumption growth. Annual real data, sample 1926-2012 (1929-2012 for consumption growth). Newey West (5 lags) standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Annual	**1926-2012**		**1926-1945**		**1946-2012**	
Panel A: $\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{dg,t+1}$						
	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$
1. Dividends with reinvestment	.00	-1.18%	-.05	-5.47%	.02	-0.99%
(t-stat)	(0.12)		(-0.24)		(0.83)	
2. Dividends without reinvestment	-.07	8.03%	-.61***	77.54%	-.01	-1.07%
(t-stat)	(-1.67)		(-7.78)		(-0.48)	
3. Dividends inclusive of M&A cash	-.23***	17.54%	-.60***	77.08%	-.12*	4.34%
(t-stat)	(-3.09)		(-9.82)		(-1.97)	
Panel B: $r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{r,t+1}$						
	b_r	*adj.* $R^2(\%)$	b_r	*adj.* $R^2(\%)$	b_r	*adj.* $R^2(\%)$
1. Dividends with reinvestment	.10***	3.58%	.53**	7.69%	.11***	5.99%
(t-stat)	(2.82)		(2.44)		(3.19)	
2. Dividends without reinvestment	.08*	1.86%	.02	-5.82%	.11**	6.67%
(t-stat)	(1.73)		(0.17)		(3.12)	
3. Dividends inclusive of M&A cash	.15**	4.40%	0.07	-5.20%	.20***	10.62%
(t-stat)	(2.40)		(0.57)		(3.59)	
Panel C: $\Delta c_{t+1} = a_c + b_c(d_t - p_t) + \varepsilon_{cg,t+1}$						
	b_c	*adj.* $R^2(\%)$	b_c	*adj.* $R^2(\%)$	b_c	*adj.* $R^2(\%)$
1. Dividends with reinvestment	-.00	-1.21%	-.02	-6.12%	-.00	-1.56%
(t-stat)	(-0.11)		(-0.33)		(-0.00)	
2. Dividends without reinvestment	-.01	2.10%	-.10**	30.73%	-.00	-0.81%
(t-stat)	(-0.88)		(-2.45)		(-0.65)	
3. Dividends inclusive of M&A cash	-.03*	11.77%	-.10***	32.19%	-.02**	6.75%
(t-stat)	(-1.74)		(-2.92)		(-2.30)	

Table 1.4: Predictive regressions of excess returns, dividend and consumption growth at quarterly frequency. **Panel A**: Δd_{t+1} is the 4-quarters rolling real dividend growth, $d_t - p_t$ is the log dividend-price ratio. **Panel B**: r_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R^{CRSP}_{t+1}) - ln(R_f)$), $d_t - p_t$ is the log dividend-price. R^{CRSP}_{t+1} is the CRSP cum dividend gross return, R_f is the gross risk-free rate. **Panel C**: Δc_{t+1} is the 4-quarters rolling real consumption growth. Quarterly real data, sample 1926-2012 (1947-2012 for consumption growth). Newey West (5 lags) standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Quarterly	**1926-2012**		**1926-1945**		**1946-2012**	
Panel A: $\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{dg,t+1}$						
	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$
1. Dividends with reinvestment	-.01	2.90%	-.12***	33.85%	-.00	-0.36%
(t-stat)	(-1.57)		(-5.48)		(-0.12)	
2. Dividends without reinvestment	-.02*	4.53%	-.12***	38.64%	-.00	-0.06%
(t-stat)	(-1.88)		(-6.49)		(-0.49)	
3. Dividends inclusive of M&A cash	-.05***	8.22%	-.13***	42.63%	-.03*	1.84%
(t-stat)	(-3.50)		(-8.30)		(-1.66)	
Panel B: $r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{r,t+1}$						
	b_r	*adj.* $R^2(\%)$	b_r	*adj.* $R^2(\%)$	b_r	*adj.* $R^2(\%)$
1. Dividends with reinvestment	.02	0.69%	.05	-0.52%	.03**	1.60%
(t-stat)	(1.51)		(0.66)		(2.27)	
2. Dividends without reinvestment	.02	0.63%	.05	-0.68%	.03**	1.49%
(t-stat)	(1.48)		(0.62)		(2.20)	
3. Dividends inclusive of M&A cash	.05**	1.59%	.06	-0.18%	.05***	2.83%
(t-stat)	(2.00)		(0.77)		(3.25)	
Panel C: $\Delta c_{t+1} = a_c + b_c(d_t - p_t) + \varepsilon_{cg,t+1}$						
	b_c	*adj.* $R^2(\%)$	b_c	*adj.* $R^2(\%)$	b_c	*adj.* $R^2(\%)$
1. Dividends with reinvestment	-.00	0.05%	-	-	-.00	0.05%
(t-stat)	(-0.61)		-		(-0.61)	
2. Dividends without reinvestment	-.00	0.44%	-	-	-.00	0.44%
(t-stat)	(-0.84)		-		(-0.84)	
3. Dividends inclusive of M&A cash	-.01***	10.03%	-	-	-.01***	10.03%
(t-stat)	(-2.89)		-		(-2.89)	

Table 1.5: Predictability: Robustness checks. **Panel A**: Predictive regressions of dividend growth using a dividend measure that includes only CRSP M&A cash dividends. Δd_{t+1} is the real dividend log growth between t and $t+1$, $d_t - p_t$ is the log dividend-price ratio. **Panel B**: Predictive regressions of excess returns using a dividend measure that includes only CRSP M&A cash dividends. r_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R^{CRSP}_{t+1}) - ln(R_f)$), $d_t - p_t$ is the log dividend-price. R^{CRSP}_{t+1} is the CRSP cum dividend gross return (VWRETD), R_f is the gross risk-free rate. **Panel C**: Predictive regressions of consumption growth using a dividend measure that includes only CRSP M&A cash dividends. Δc_{t+1} is the real (log) consumption growth. **Panel D**: Predictive regressions of standard dividend growth (CRSP with reinvestment) using a dividend-price ratio that includes M&A cash dividends.

	1926-2012		**1926-1945**		**1946-2012**	
Panel A: $\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{dg,t+1}$						
	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$
Dividends inclusive of public-only M&A cash (annual)	-.21***	17.86%	-.60***	77.08%	-.14**	7.79%
(t-stat)	(-3.38)		(-9.82)		(-2.37)	
Dividends inclusive of public-only M&A cash (quarterly)	-.04***	7.18%	-.13***	42.63%	-.03*	2.46%
(t-stat)	(-3.29)		(-8.30)		(-1.87)	
Panel B: $r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{r,t+1}$						
	b_r	*adj.* $R^2(\%)$	b_r	*adj.* $R^2(\%)$	b_r	*adj.* $R^2(\%)$
Dividends inclusive of public-only M&A cash (annual)	.10*	1.94%	.08	-5.20%	.11**	4.33%
(t-stat)	(1.85)		(0.57)		(2.26)	
Dividends inclusive of public-only M&A cash (quarterly)	.03*	0.86%	.06	-0.18%	.03**	1.27%
(t-stat)	(1.68)		(0.77)		(2.13)	
Panel C: $\Delta c_{t+1} = a_c + b_c(d_t - p_t) + \varepsilon_{cg,t+1}$						
	b_c	*adj.* $R^2(\%)$	b_c	*adj.* $R^2(\%)$	b_c	*adj.* $R^2(\%)$
Dividends inclusive of public-only M&A cash (annual)	-.02	9.16%	-.10***	32.19%	-.01*	5.17%
(t-stat)	(-1.63)		(-2.92)		(-1.79)	
Dividends inclusive of public-only M&A cash (quarterly)	-.00**	8.29%	-	-	-.00**	8.29%
(t-stat)	(-2.51)		-		(-2.51)	
Panel D: $\Delta d^{CRSP}_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon_{dg,t+1}$						
	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$	b_d	*adj.* $R^2(\%)$
Dividends inclusive of M&A cash (annual)	.06	0.55%	-.06	-4.53%	0.12**	5.20%
t-stat	(1.37)		(-0.64)		(2.51)	
Dividends inclusive of M&A cash (quarterly)	-.03**	4.99%	-.10***	33.23%	0.00	-0.37%
t-stat	(-2.13)		(-5.43)		(0.10)	

Table 1.6: Out-of-sample dividend growth and excess return predictability. OOS $R^2 = 1 - \frac{MSE_P}{MSE_M}$. The MSE-F statistic of McCracken (2007) is defined as $MSE - F = (T - s_0)(\frac{MSE_M - MSE_P}{MSE_P})$. The cash flows in (1) are the dividends extracted from the annual CRSP index return (e.g., VWRETD, subject to the reinvestment issue), in (2) are the sum of monthly dividends extracted from CRSP index returns (no reinvestment issue), in (3) are the sum of all the dividends present in CRSP database with distribution codes 1xxx, 2xxx with M&A cash dividends being CRSP dividends with distribution codes from 3xxx up to 3400 before 1980 and SDC M&A cash dividends with US public buyer and US target post 1980. Annual and quarterly real data, estimation sample $s_0 = 45$ ($s_0 = 172$) years (quarters) from 1926-1971, forecasting sample 1972-2012. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Panel A: Dividend Growth

	Annual		*Quarterly*	
	OOS R^2	*MSE-F*	*OOS R^2*	*MSE-F*
1. Dividends with reinvestment	-4.76%	-1.86	-16.50%	-24.36
2. Dividends without reinvestment	-63.81%	-15.97	-25.23%	-34.65
3. Dividends inclusive of M&A cash	6.82%	3.00**	1.73%	3.02**

Panel B: Returns

	Annual		*Quarterly*	
	OOS R^2	*MSE-F*	*OOS R^2*	*MSE-F*
1. Dividends with reinvestment	-7.97%	-3.02	-0.23%	-0.40
2. Dividends without reinvestment	-0.06%	-0.02	-0.15%	-0.27
3. Dividends inclusive of M&A cash	7.73%	3.44***	2.16%	3.79***

1.4.3 Long run predictability

Following Cochrane (2008), we look at long-horizon coefficients implied from the one-year regression coefficients in order to determine how much of the variation in the dividend-price ratio is explained by changes in dividends and discount rates. Consider a first-order VAR represention of log returns, log dividend-price ratios and log dividend growth

$$r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon^r_{t+1} \tag{1.10}$$

$$\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon^d_{t+1} \tag{1.11}$$

$$d_{t+1} - p_{t+1} = a_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1} \tag{1.12}$$

The Campbell and Shiller (1988a) present value identity (1.1) implies that the regression coefficients in (1.10)-(1.12) are related by

$$b_r \approx 1 - \rho\phi + b_d \tag{1.13}$$

where ρ is defined as $\rho = \frac{exp(-\overline{dp})}{1+exp(-\overline{dp})}$, where $\overline{dp}$ is the mean log dividend-price ratio. If we divide (1.13) by $1-\rho\phi$ we obtain

$$\frac{b_r}{1-\rho\phi} - \frac{b_d}{1-\rho\phi} \approx 1 \tag{1.14}$$

$$b^{lr}_r - b^{lr}_d \approx 1 \tag{1.15}$$

It is clear that one of (1.10)-(1.12) is redundant, as one can infer the data, coefficients and error of any one equation from those of the other two. The long run coefficients b^{lr}_r and b^{lr}_d in (1.15) are, respectively, the regression coefficients of long-run returns $\sum_{j=1}^{\infty}\rho^{j-1}r_{t+j}$ and dividend growths $\sum_{j=1}^{\infty}\rho^{j-1}\Delta d_{t+j}$ on the dividend-price ratio at time t. More precisely, by combining the VAR described above with the present value relation

(1.1), we can calculate the VAR implied long-horizon coefficients for each horizon K as

$$b_r^K = \frac{b_r(1-\rho^K\phi^K)}{1-\rho\phi}$$

$$b_d^K = \frac{b_d(1-\rho^K\phi^K)}{1-\rho\phi}$$

$$b_{dp}^K = \rho^K\phi^K$$

Note that when $K = \infty$ we obtain again equation (1.15). b_r^{lr} and b_d^{lr} represent the fraction of the variance of the dividend-price ratio that can be attributed, respectively, to time-varying expected returns and dividend growth (see Cochrane (2008), van Binsbergen and Koijen (2010), Golez (2014)). Table 1.7 reports the estimates of the VAR system and long run coefficients, along with the t-statistics of two null (joint) hypotheses. The first tests $b_r^{lr} = 0, b_d^{lr} = -1$, while the second $b_r^{lr} = 1, b_d^{lr} = 0$ as done in Cochrane (2008) and Maio and Santa-Clara (2014). The standard errors of the long-run coefficients b_r^{lr} and b_d^{lr} are calculated by the delta method.

Looking at b_r^{lr} and b_d^{lr} for the measure with reinvestment of dividends, we find that approximately all the variation in the dividend-price ratio is accounted for by changes in the discount rate and none by news about future cash flows, consistent with Cochrane (2008) and Golez (2014), and the null of no return predictability is rejected with a t-stat value of 2.65. Using the measure without reinvestment of dividends, we see that changes in the dividend-price ratio are equally explained by variation in discount rates and dividend growth, but both tests reject the statistical significance of the coefficients. However, when we use my dividend measure that includes M&A cash dividends we see a different result. Both expected returns and expected dividend growth help explain the variation in the dividend-price ratio. Expected dividend growth explains approximately 60% of the variation in the dividend-price ratio, while time-varying discount

Table 1.7: OLS forecasting regressions of returns, dividend growth and dividend-price ratio on the lagged dividend-price ratio using the various dividend measures. The first row presents the regression $r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon^r_{t+1}$, the second row $\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \varepsilon^d_{t+1}$ and the third row $d_{t+1} - p_{t+1} = a_{dp} + \phi(d_t - p_t) + \varepsilon^{dp}_{t+1}$. r_{t+1} is the CRSP cum dividend log return (e.g., $\ln(R_{t+1})$), $d_t - p_t$ is the log dividend-price. The lagged dividend-price ratio used as regressor includes in columns (1)-(3) the dividends extracted from the annual CRSP index return (e.g., VWRETD, subject to the reinvestment issue), in columns (4)-(6) the sum of monthly dividends extracted from CRSP index returns (no reinvestment issue), in columns (7)-(9) the sum of all the dividends present in CRSP database with distribution codes 1xxx, 2xxx with M&A cash dividends being CRSP dividends with distribution codes from 3xxx up to 3400 before 1980 and SDC M&A cash dividends with US public buyer and US target post 1980. The "implied" column calculates each coefficient based on the other two coefficients and the identity $b_r = 1 - \rho\phi + b_d$. ρ is calculated as $\rho = \frac{exp(-\overline{dp})}{1+exp(-\overline{dp})}$, with the mean log dividend-price ratio being constructed under the different dividend measures. The long-run returns forecast coefficient b^{lr}_r is computed as $b^{lr}_r = \frac{b_r}{1-\rho\phi}$ and the long-run dividend-growth forecast coefficient b^{lr}_d as $b^{lr}_d = \frac{b_d}{1-\rho\phi}$. Annual nominal data, sample 1926-2012. Newey West (5 lags) standard errors.

	Divs with reinv.			**Divs w/o reinv.**			**Divs incl. M&A cash**		
	$\hat{b}, \hat{\phi}$	$\sigma(\hat{b})$	*implied*	$\hat{b}, \hat{\phi}$	$\sigma(\hat{b})$	*implied*	$\hat{b}, \hat{\phi}$	$\sigma(\hat{b})$	*implied*
r	0.096	0.036	0.099	0.071	0.048	0.071	0.154	0.066	0.121
Δ d	0.003	0.030	0.001	-0.082	0.059	-0.081	-0.234	0.089	-0.201
dp	0.936	0.040	0.938	0.876	0.054	0.876	0.670	0.069	0.636
ρ	0.966			0.967			0.962		
b^{lr}_r	**1.004**	0.38		0.467	0.31		**0.432**	0.19	
b^{lr}_d	**0.027**	0.32		-0.534	0.38		**-0.66**	0.25	
$t(H_0 : b^{lr}_r = 0, b^{lr}_d = -1)$	**2.65*****			1.50			**2.33****		
$t(H_0 : b^{lr}_r = 1, b^{lr}_d = 0)$	0.08			-1.39			**-2.61*****		

rates account for the remaining 40%. Moreover, we see that both (joint) hypotheses are rejected, implying that both expected return and dividend growth help explain variation in the dividend-price ratio. These results show that expectations about future cash flows affect today's stock prices, even more than news about future discount rates. This is not completely unexpected, as news on mergers and acquisitions is amongst the most price-sensitive information available on the market and it is reassuring to know that using a more comprehensive measure of dividends we can link cash flows news and stock market volatility.

1.4.4 Predictability during M&A waves

It is a well known fact that M&A activity seems to happen in waves (e.g., Mitchell and Mulherin (1996), Mulherin and Boone (2000), Andrade et al. (2001), Brealey and Myers (2003)). We also know that evidence of return predictability is stronger during recessions, as recently suggested by Rapach et al. (2010), Henkel et al. (2011), Dangl and Halling (2012) and Gargano et al. (2014). Therefore, a natural question is whether there exists a link between these M&A waves and business cycles. In other words, is high (low) M&A activity related to expansions (recessions)? Is predictability higher during periods of low M&A activity?

To answer the first question, I define periods of high M&A activity to be those years above the 80% percentile of my nominal M&A cash dividends timeseries standardized by the US nominal GNP to account for the non-stationarity of the series.[33] Since we deal with annual data starting from 1926, this results in 18 years of high M&A activity, all of them post-1972. The remaining years are labeled periods of low M&A activity. I use data from the NBER US Business Cycle Expansions and Contractions table to define the recession periods. I establish the link between M&A activity and expansions/recessions by looking at the degree of association, ϕ[34], and at Pearson's χ^2 test between the (binary) variables NBER recessions and periods of low M&A activity post-1972. I find $\phi = 0.23$ over the full sample with a statistically significant χ^2 statistic of 4.80, rejecting the null hypothesis of independence between the two variables. This positive relationship between periods of low M&A activity and recessions can also be noted by looking at years 2002 and 2009 when M&A activity was extremely low. This result is unsurprising, as companies prefer to use cash for acquisitions when capital mar-

[33] Using other top percentiles (e.g., 90% or 70%) or real M&A cash dividends the results are qualitatively the same.

[34] $\phi = \sqrt{\frac{\chi^2}{N}}$ where N is the grand total of observations (e.g., 87). In a 2x2 case the ϕ coefficient coincides with the Pearson correlation coefficient.

kets are liquid, debt issuance is easier, and equity valuations are high (Harford (2005)). In order to analyze the relationship between recessions and cash M&A activity, we can also look at the mean and median dividend-price ratio under the two regimes. The median log dividend-price ratios are -3.45 and -3.19 during the high M&A and low M&A periods, respectively. The ratio is therefore countercyclical (Campbell et al. (1996), Ch.7), higher during periods of low M&A activity, which are positively correlated with recessions. During recessions market prices are low and dividends are sticky (Fama and French (2002)), resulting in higher dividend-price ratios.

I now look at both in-sample and out-of-sample excess return predictability. I run the following predictive regressions over the full sample:

$$r_{t+1} = \alpha + \beta(d_t - p_t)I_{\{t=high\_M\&A\}} + \gamma(d_t - p_t)(1 - I_{\{t=high\_M\&A\}}) + \varepsilon_{t+1} \qquad (1.16)$$

$$r_{t+1} = a + \eta(d_t - p_t) + \phi(d_t - p_t)I_{\{t=high\_M\&A\}} + \varepsilon_{t+1} \qquad (1.17)$$

where r_{t+1} is the CRSP cum dividend log excess return[35] and $d_t - p_t$ is the lagged dividend-price ratio constructed using my cash M&A measure.

Equation (1.17) allows us to test the $H_0 : \phi = \beta - \gamma = 0$ (e.g., whether there is different in-sample predictability between the two periods), while equation (1.16) shows the individual high versus low M&A period coefficients. We notice that in-sample both coefficients are statistically significant, with the coefficient of the high M&A period (0.15) approximately equal to the low M&A period one (0.15). Since dividend-price ratios are slightly higher during the low M&A periods, the total effects (e.g., $\beta(d_t - p_t)I_{\{t=high\_M\&A\}}$ and $\gamma(d_t - p_t)(1 - I_{\{t=high\_M\&A\}})$) are similar. We cannot reject the null hypothesis that the ϕ coefficient is different from 0. In other words, these results confirm our previous finding that adding M&A cash dividends improves (excess)

[35] e.g., $\ln(R^{CRSP}) - ln(R_f)$. R^{CRSP}_{t+1} is the CRSP cum dividend gross return (e.g., VWRETD) and R_f is the gross risk-free rate.

returns predictability in general, but there is not much difference between the two periods for (in-sample) predictability. However, the most interesting results appear out-of-sample. I compute R^2 values separately for the recession and expansion samples. I find an out-of-sample R^2 of 12.47% during low cash M&A periods (e.g., recessions), while the R^2 during expansions is 3.12%.[36] This suggests that most of the out-of-sample return predictability comes from periods of low M&A cash flows. This is consistent with Rapach et al. (2010), Henkel et al. (2011), who find a similar result using a regime-switching VAR methodology (p.576, "in recessions, ... the VAR or the RSVAR would be preferable (out-of-sample) to the historical average") and with Dangl and Halling (2012). Summing up, our results suggest that we have strong out-of-sample return predictability during periods of low M&A activity, related to recessions, but relatively little during periods of high M&A activity, related to expansion periods.

1.4.5 Predictability from individual components

In the previous section we looked at the predictability of "complete" dividend-price or payout ratios. We now look at the impact of the individual cash flow components on return predictability. We run the following regression:

$$r_{t+1} = \alpha + \beta(d_t^{CRSP} - p_t) + \gamma(d_t^{cash\_M\&A} - p_t) + \delta(ney_t) + \varepsilon_{t+1} \tag{1.18}$$

where r_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R_{t+1}^{CRSP}) - ln(R_f)$), R_{t+1}^{CRSP} is the CRSP cum dividend gross return (VWRETD), R_f is the gross risk-free rate, $d_t^{CRSP} - p_t$ is the log CRSP without reinvestment dividend-price ratio, $d_t^{cash\_M\&A} - p_t$ is the cash M&A only log dividend-price ratio and $ney_t = ln(1+NP)$ is the net issues ratio

[36] Our out-of-sample analysis runs from 1972 up to 2012 (40 years), with 17 years of high M&A and 23 years of low M&A. Results using different initial estimation samples are qualitatively similar.

of Boudoukh et al. (2007) that includes only repurchases and issues.[37] Our full sample starts in 1961 since M&A cash dividends are zero during some years before that date and end in 2010, the last available date for *ney*. Table 1.8 reports the results of the regression of the excess return on the various cash flow components. We notice that when taken individually, none of the components is statistically significant, which suggests an omitted variable bias. However, when we include all the predictors in the regression we see that the only significant component is the cash M&A dividend-price ratio. In other words, M&A cash dividends have substantial predictive power for excess returns compared to the ordinary dividends or repurchases and issues. It is also interesting that the net equity yield is never significant, neither by itself nor when pooled with the other predictors, consistent with the findings of Eaton and Paye (2013), who argue that this is mainly due to the issuance component. This suggests that repurchases and issues tend to be less important than M&A cash dividends for return predictability. Overall, these results confirm the importance of the M&A cash dividends: M&A activity has a big impact on stock prices, not only ex-ante (e.g., stock prices incorporate the probability of a takeover) but also ex-post (e.g., higher M&A activity today predicts higher returns tomorrow).

[37] They add 1 because the ratio can be negative.

Table 1.8: Excess return predictability regressions on the individual cash flow components. r^{CRSP}_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R^{CRSP}_{t+1}) - ln(R_f)$), $d^i_t - p_t$ is the log dividend-price. R^{CRSP}_{t+1} is the CRSP cum dividend gross return (VWRETD), R_f is the gross risk-free rate. $d^{CRSP}_t - p_t$ is the log CRSP without reinvestment dividend-price ratio, $d^{cash\_M\&A} - p_t$ is the cash M&A only log dividend-price ratio and $ney_t = ln(1+NP)$ is the net issues ratio of Boudoukh et al. (2007) that includes only repurchases and issues. Annual data, sample 1961-2010. Newey West (3 lags) standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

$d^{CRSP}_t - p_t$	.09			.12**	.08		.09
	(1.67)			(2.25)	(1.11)		(1.27)
$d^{cash\_M\&A}_t - p_t$		.01		.02**		.02*	.02**
		(1.49)		(2.17)		(1.99)	(2.30)
ney_t			-1.63		-.44	-2.60	-1.27
			(-1.03)		(-0.22)	(-1.56)	(-0.59)
adj. R^2	1.94%	0.86%	-0.24%	5.68%	-0.14%	3.13%	4.34%

1.4.6 Comparison with additional cash flow measures

I now look at alternative total cash flow predictors that have been recently proposed to address the conflicting evidence on return predictability. I focus on the net payout yield of Boudoukh et al. (2007)[38], which includes repurchases and issues and the payout price ratio of Robertson and Wright (2006), which uses the total payout (e.g., repurchases, issues, M&A, private equity transactions) of the economy based on FED data.[39] None of these predictors addresses the dividend growth predictability puzzle, because repurchases and net issues are not dividends.

Net Payout Yield

Table 1.9 reports the in-sample predictability results.[40] We see that the net payout yield performs better over the full sample. However, the performance is sensitive to the inclusion of data from the years 1929-1930, when the net payout measure performs extremely well, consistent with the findings of Cochrane (2008) and Eaton and Paye (2013). Post-war the net payout yield, as suggested by the adjusted R^2, has substantially less explanatory power than my dividend-price ratio. As reported in Figure 1 of Boudoukh et al. (2007), the share repurchases and issues start to be relevant only post-1970. As a consequence, one would expect the net payout yield to perform better when repurchases and issues start to be a substantial amount. This does not seem to be the case. In other words, as reported in Figure 1.1, both M&A and repurchases/issues cash flows boom post-war, but the M&A cash dividends tend to explain excess return predictability the most. We now analyze out-of-sample predictability (Table 1.10 and Figure 1.9 (top panel)), using half of the sample (e.g., 45 years) for initial estimation. We see that the net payout yield performs relatively poorly out-of-sample, with an OOS

[38] We use their “lcrspnpy” variable, their best measure, which is available monthly up to 2010.
[39] The FED does not provide a clear split of the components nor details on how they compute them.
[40] Their data are available only for the period 1927-2010.

R^2 of -17.03%. My measure instead has an OOS R^2 of 8.35% and a large and significant MSE-F statistic of 3.65.

In conclusion, my cash M&A measure seems to perform better than the net payout yield both in-sample (post-war) and out-of-sample.

Payout Price Ratio

The CRSP database does not take into account intercorporate cross-holdings. As a consequence, using it to extract dividends might result in a positive biased amount of ordinary dividends that could affect predictability results. Robertson and Wright (2006) therefore propose to use a total payout ratio using the total nonfinancial U.S. corporate sector data from the Federal Reserve Board Financial Accounts Tables.[41] We follow their approach and use Table B.102 line 36 (Market Value of equities outstanding) for the market value of equities, Table F.102 line 39 (net new equity issues)[42] for the net equity issues (it includes cash M&A plus repurchases minus public and private equity issues) and Table F.102 line 3 (net dividends) for the ordinary cash dividends. We construct annual total payout summing the net dividends and the net equity issues every year. Table 1.9 reports the in-sample predictability results. Data are available only post-war starting from 1946, so our full sample has 66 observations. The total payout price ratio has a barely statistically significant coefficient at the 5% level and an adjusted R^2 of 5.34% compared to the 10.62% of our measure. Out-of-sample (initial estimation sample $s_0 = 45$, thus forecasts start in 1991) we see that my measure has more explanatory power than the payout price measure, which has an OOS R^2 of -8.56% (Table 1.10). Figure 1.9 (bottom panel) reports the cumulative squared prediction errors of the null minus the alternative model. We see that the performance of the total payout price ratio is good for the first few years and then collapses in 2007 at the beginning of the recent

[41] also known as Flow of Funds tables.
[42] equivalently, Table R.102 line 10.

Table 1.9: Comparison of excess return predictability between my dividend-price ratio inclusive of M&A cash dividends versus (1) the net payout yield measure of Boudoukh et al. (2007), (2) the payout price ratio proposed by Robertson and Wright (2006). r_{t+1}^{CRSP} is the CRSP cum dividend log excess return (e.g., $\ln(R_{t+1}^{CRSP}) - ln(R_f)$), $d_t^i - p_t$ is the log dividend-price. R_{t+1}^{CRSP} is the CRSP cum dividend gross return (VWRETD), R_f is the gross risk-free rate. (1) is the net payout yield measure ("lcrspnpy" variable) of Boudoukh et al. (2007). The lagged dividend-price ratio used in (3) includes the sum of all the dividends present in CRSP database with distribution codes 1xxx, 2xxx with M&A cash dividends being CRSP dividends with distribution codes from 3xxx up to 3400 before 1980 and SDC M&A cash dividends with US public buyer and US target post 1980. Annual data, sample 1927-2010. Newey West (5 lags) standard errors. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

Annual	**1927-2010**		**1927-1945**		**1946-2010**		**1946-2012**	
Panel A: $r_{t+1} = a_r + b_r(d_t - p_t) + \varepsilon_{r,t+1}$								
	b_d	adj. $R^2(\%)$	b_d	adj. $R^2(\%)$	b_d	adj. $R^2(\%)$	b_d	adj. $R^2(\%)$
1. Net Payout ratio	.29***	9.39	.73***	27.07	.22***	5.71	-	-
t-stat	(2.84)		(7.79)		(3.07)		-	
2. Payout Price ratio	-	-	-	-	-	-	.09**	5.34
t-stat	-		-		-		(2.01)	
3. Dividends incl. M&A cash	.15**	4.47	.08	-5.35	.20***	**10.98**	.20***	**10.62**
t-stat	(2.40)		(0.65)		(3.66)		(3.59)	

financial crisis. Overall, my dividend-price ratio inclusive of M&A cash dividends has stronger performance than the payout price ratio both in-sample and out-of-sample.

Table 1.10: Out-of-sample excess return predictability. OOS $R^2 = 1 - \frac{MSE_P}{MSE_M}$. The MSE-F statistic of McCracken (2007) is defined as $MSE - F = (T - s_0)(\frac{MSE_M - MSE_P}{MSE_P})$. (1) is the net payout yield measure ("lcrspnpy" variable) of Boudoukh et al. (2007), (2) is the payout price ratio of Robertson and Wright (2006), (3) the dividend-price ratio inclusive of M&A cash dividends. Estimation sample: (1) $s_0 = 45$ years from 1926-1971, forecasting sample 1972-2011; (2) $s_0 = 45$ years from 1946-1991, forecasting sample 1992-2012; (3) $s_0 = 45$ years from 1926-1971, forecasting sample 1972-2011. Annual data. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

	OOS $R^2(\%)$	*MSE-F*
1. Net Payout ratio	-17.03	-5.82
2. Payout Price ratio	-8.56	-1.66
3. Dividends inclusive of M&A cash	8.35	3.65***

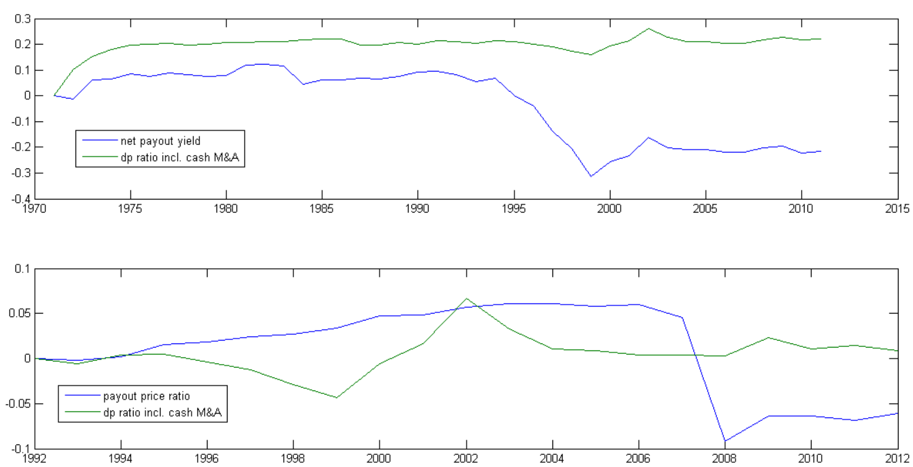

Figure 1.9: Comparison of out-of-sample predictability between the net payout yield, the payout price ratio and my dp ratio. This figure plots the cumulative squared prediction errors of the null minus the cumulative squared prediction error of the alternative for the net payout yield, the payout price ratio and my dividend-price ratio inclusive of M&A cash dividends. The alternative is model (1.9). The null is the prevailing mean.

1.5 Industries

In this section I discuss evidence of return and dividend predictability at the industry level. I showed in the previous sections that M&A cash flows matter for predictability at the aggregate level. Therefore, given the fact that some industries are subject to more M&A activity than others, a natural question is whether there is more predictability in those industries.

Figure 1.10 shows the relative weight of the M&A cash dividends for each industry. We notice that the relative industry proportions change over time. Some industries M&A dividends are concentrated during specific years (e.g., Hitech industry during the early '90s and over the last 10 years). Over the post-1970 period, the consumer non-durables, manufacturing, Hitech and Wholesale/Retail industries account for the lion's share of these cash flows, respectively, with 12.44%, 18.09%, 17.28% and 20.07% median weight.[43]

1.5.1 Predictive regressions

We run the following timeseries predictive regressions

$$\Delta d^i_{t+1} = a + b^i(d^i_t - p^i_t) + \varepsilon^i_{t+1} \tag{1.19}$$

$$r^i_{t+1} = \alpha + \beta^i(d^i_t - p^i_t) + \varepsilon^i_{t+1} \tag{1.20}$$

and panel regressions

$$\Delta d_{i,t+1} = a_i + b(d_{i,t} - p_{i,t}) + \varepsilon_{i,t+1} \tag{1.21}$$

[43] Using the mean gives approximately the same ranking, but given the number of outliers for some industries, the median is a better descriptive measure.

$$r_{i,t+1} = \alpha_i + \beta(d_{i,t} - p_{i,t}) + \varepsilon_{i,t+1} \tag{1.22}$$

where $i = 1, 2, ..9$ represent the industries.[44] Tables 1.11 and 1.12 report the results.[45] Dividend growth predictability seems robust across industries and subperiods, with the only exception of the healthcare industry pre-war. Similarly to the aggregate level analysis, dividend growth predictability is stronger pre-war. Moreover, the coefficients on the dividend-price ratio are all statistically significant and negative, as it should be, for every industry. The panel regression, using the standard errors suggested by Petersen (2009) in order to account for both industry and time effects[46], confirms the statistical significance of the results.

Looking at return predictability, we see that over the full sample there is some predictability for the industries with the largest concentration of M&A cash flows. However, post-war we observe substantial in-sample return predictability for those industries that have higher M&A cash dividends. Consumer non-durables, manufacturing and wholesale/retail healthcare industries have large adjusted R^2 and statistically significant coefficients. Those are the industries that, together with the Hitech industry, have the largest share of M&A cash dividends.[47]

A possible explanation for this presence of dividend growth and return predictability at the industry level is the fact that my dividend-price ratios are less persistent than the standard ones traditionally considered in the literature (see Koijen and Nieuwerburgh (2011)). In fact, the maximum autoregressive coefficient AR(1) over the full sample is 0.746 for industry 8 and post-war is 0.73, still for industry 8, substantially

[44] I exclude the "Others" industry because it is not a proper industry.

[45] Running equations (1.19) and (1.20) with de-meaned variables, as suggested by Koijen and Nieuwerburgh (2011), gives similar results.

[46] We also run the panel with fixed effects, which implies clustering by industry effect only, as a robustness check. The results are even more statistically significant than the ones reported here.

[47] We also tried to predict the aggregate stock market return using the dividend-price ratios of the individual industries, but we find that no specific industry has predictive power for the aggregate stock market.

Table 1.11: **Panel A**: Predictive regressions of dividend growth inclusive of M&A cash dividends at the industry level. **Panel B**: Predictive regressions of dividend growth (standard measure, without M&A cash dividends) at the industry level. Time-series regressions are based on (1.19). Panel indicates the panel regression (1.21) with all the industries. Standard errors are clustered on both industry and time dimensions to account for both firm and time effects in the panel dataset (Petersen (2009)). Newey West standard errors (3 lags) are used for individual timeseries regressions. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Annual data, 1926-2012.

	Panel A						**Panel B**					
	1926-2012		**1926-1945**		**1946-2012**		**1926-2012**		**1926-1945**		**1946-2012**	
	b	$R^2(\%)$	b	$R^2(\%)$	b	$R^2(\%)$	b	$R^2(\%)$	b	$R^2(\%)$	b	$R^2(\%)$
1. Consumer Non-Durables	-.31***	16.85%	-.38***	46.22%	-.30***	14.35%	-.09***	8.07%	-.38***	46.22%	-.06**	3.40%
t-stat	(-3.37)		(-3.68)		(-2.87)		(-3.30)		(-3.68)		(-2.55)	
2. Consumer Durables	-.53***	34.53%	-.83***	70.79%	-.47***	28.08%	-.23***	14.33%	-.83***	71.01%	-.14**	5.30%
t-stat	(-4.84)		(-4.94)		(-3.63)		(-3.19)		(-4.96)		(-2.26)	
3. Manufacturing	-.42***	27.30%	-.71***	75.68%	-.33***	16.56%	-.19**	15.44%	-.71***	75.71%	-.06	1.19%
t-stat	(-4.39)		(-8.50)		(-2.99)		(-2.25)		(-8.44)		(-1.15)	
4. Energy	-.46***	25.72%	-.59***	44.38%	-.40***	18.97%	-.16**	10.69%	-.59***	44.38%	-.05	1.04%
t-stat	(-4.19)		(-4.41)		(-3.05)		(-2.12)		(-4.41)		(-1.35)	
5. HiTech	-.28***	16.48%	-.61***	59.19%	-.24**	11.54%	-.08*	5.37%	-.60***	57.99%	-.06	2.55%
t-stat	(-2.91)		(-4.38)		(-2.23)		(-1.73)		(-4.13)		(-1.18)	
6. Telecom	-.27***	12.74%	-.17**	15.76%	-.30**	12.93%	-.02	0.45%	-.17**	15.76%	-.02	-0.40%
t-stat	(-2.79)		(-2.09)		(-2.61)		(0.92)		(-2.09)		(-0.69)	
7. Wholesale and Retail	-.27***	14.45%	-.40***	53.49%	-.25***	11.04%	-.06*	4.73%	-.40***	53.49%	-.05	3.48%
t-stat	(-3.66)		(-5.31)		(-2.95)		(-1.92)		(-5.31)		(-1.49)	
8. Healthcare	-.15***	6.22%	-.13	-0.71%	-.18**	6.76%	-.04	1.48%	-.13	-0.71%	-.03	1.05%
t-stat	(-2.79)		(-1.62)		(-2.62)		(-1.55)		(-1.62)		(-1.31)	
9. Utilities	-.22***	17.10%	-.37***	70.31%	-.16**	6.85%	-.14***	19.42%	-.37***	70.31%	-.03	-0.31%
t-stat	(-3.87)		(-8.35)		(-2.19)		(-2.82)		(-8.35)		(-1.13)	
Panel	-.27***	16.81%	-.45***	44.69%	-.23***	12.87%	-.08***	6.92%	-.45***	44.57%	-.05***	3.24%
t-stat	(-4.31)		(-6.19)		(-3.55)		(-3.25)		(-6.14)		(-3.05)	

lower than the near unity value previously found at the aggregate level in the literature. Structurally adjusting the dividend-price ratios (e.g., Lettau and Nieuwerburgh (2008)) results in similar values.

In conclusion, it seems that M&A cash dividends not only matter at the aggregate level, but also have a notable impact at the industry level, especially post-war, when these cash flows start to be a non-trivial amount.

Table 1.12: **Panel A**: Predictive regressions of excess returns at the industry level (dp ratio inclusive of M&A cash dividends). **Panel B**: Predictive regressions of excess returns at the industry level (standard dp ratio). Time-series regressions are based on (1.20). The "% weight" column represent the median of the M&A cash dividends proportion over the sample 1971-2012. Panel indicates the panel regression (1.22) with all the industries. Standard errors are clustered on both industry and time dimensions to account for both firm and time effects in the panel dataset (Petersen (2009)). Newey West standard errors (3 lags) are used for individual timeseries regressions. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Annual data, 1926-2012.

	Panel A							**Panel B**					
	1926-2012		**1926-1945**		**1946-2012**		**1971-2012**	**1926-2012**		**1926-1945**		**1946-2012**	
	β	$R^2(\%)$	β	$R^2(\%)$	β	$R^2(\%)$	*% weight*	β	$R^2(\%)$	β	$R^2(\%)$	β	$R^2(\%)$
1. Consumer Non-Durables	.14***	9.40%	.13*	-2.88%	.14***	12.87%	**12.44%**	.08*	2.47%	.12*	-2.88%	.11**	4.98%
t-stat	(4.08)		(1.88)		(3.63)			(1.83)		(1.88)		(2.20)	
2. Consumer Durables	.08	0.52%	-.03	-5.74%	.10*	2.11%	4.31%	.05	-0.45%	-.03	-5.74%	.06	0.09%
t-stat	(1.58)		(-0.31)		(1.69)			(0.92)		(-0.31)		(1.13)	
3. Manufacturing	.09*	1.53%	-.06	-5.36%	.13**	8.31%	**18.09%**	.07	0.58%	-.06	-5.38%	.13**	6.56%
t-stat	(1.73)		(-0.42)		(2.62)			(1.03)		(-0.41)		(2.36)	
4. Energy	.14***	4.29%	.21**	2.74%	.11*	2.55%	6.66%	.09*	2.09%	.21**	2.74%	.08	1.61%
t-stat	(2.65)		(2.62)		(1.66)			(1.81)		(2.62)		(1.55)	
5. HiTech	.06	0.49%	-.08	-4.86%	.08	2.39%	**17.28%**	.03	0.03%	-.07	-5.01%	.05	1.79%
t-stat	(1.22)		(-0.50)		(1.51)			(0.76)		(-0.46)		(1.13)	
6. Telecom	.16***	7.21%	.43***	18.10%	.13***	4.93%	4.09%	.11**	8.05%	.43***	18.10%	.10*	7.99%
t-stat	(3.34)		(3.06)		(2.76)			(2.25)		(3.06)		(1.91)	
7. Wholesale and Retail	.09**	3.80%	.17	-0.80%	.08**	4.43%	**20.07%**	.06*	1.79%	.17	-0.80%	.06**	2.72%
t-stat	(2.34)		(1.50)		(2.01)			(1.94)		(1.50)		(2.12)	
8. Healthcare	.08*	3.13%	.13	-1.34%	.12***	6.73%	5.58%	.06	1.65%	.13	-1.34%	.13**	6.42%
t-stat	(1.72)		(0.74)		(2.75)			(1.26)		(0.74)		(2.50)	
9. Utilities	.11*	3.05%	.18	0.16%	.08*	1.83%	3.15%	.13**	3.78%	.18	0.16%	.10**	3.53%
t-stat	(1.97)		(1.22)		(1.76)			(2.27)		(1.22)		(2.10)	
Panel	.08***	2.76%	.07	1.01%	.08***	3.98%		.05*	1.59%	.07	1.02%	.05**	2.59%
t-stat	(3.05)		(0.84)		(3.13)			(1.88)		(0.85)		(2.29)	

1.6 Trading Strategy

I now assess the economic importance of my dividend-price ratio inclusive of M&A cash dividends by executing an out-of-sample trading strategy that combines the stock market with the risk-free asset, following Ferreira and Santa-Clara (2011). This strategy could have been exploited in real-time by investors. We assume investors have the standard mean-variance utility function and can invest either in the stock market or in the risk-free asset. They maximize expected utility, each period, using out-of-sample estimates of expected returns constructed using the standard return predictive regression described in Section 1.4.2 and calculate the optimal portfolio weights. More precisely,

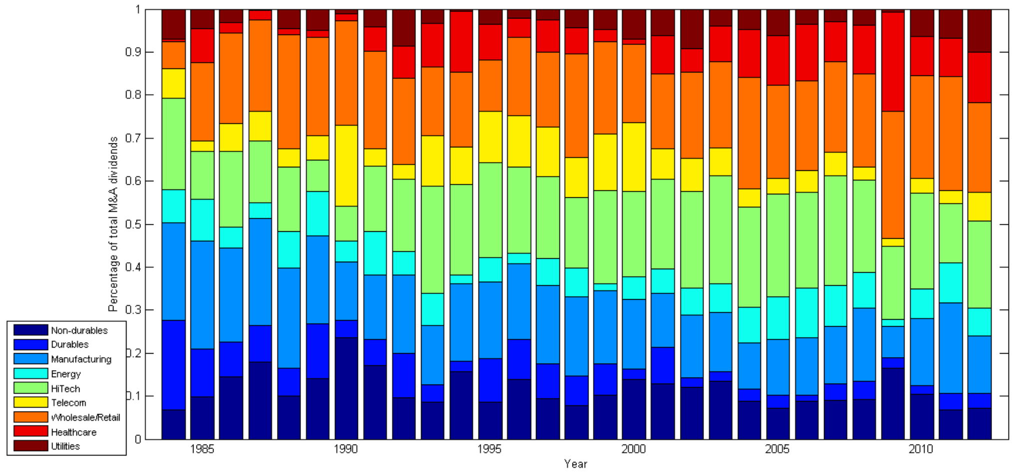

Figure 1.10: Relative weight of M&A cash dividends for each of the 9 industries. The sample consists of the SDC M&A cash dividends for each industry. All figures are in percentages.

the investor's problem is

$$\max_{w} EU(r_p) = \max_{w} E(r_p) - \frac{\gamma}{2}\sigma_p^2 = \max_{w} w\mu_E + (1-w)rf - w^2\frac{\gamma}{2}\sigma_E^2$$

where w are the optimal weight on the stock market, μ_E is the expected return on the stock market, rf is the risk free return, γ is the risk-aversion coefficient and σ_E^2 is the variance of the stock market. Every period we calculate the optimal weights

$$w_s = \frac{\hat{\mu}_s - rf_{s+1}}{\gamma \hat{\sigma^2}_s}$$

where $\hat{\mu}_s$ is the forecast of the market return between s and s+1 based on the various predictors, rf_{s+1} denotes the risk-free return from time s to s+1 (known at time s), γ is assumed to be 2[48], and $\hat{\sigma^2}_s$ is the variance of the stock market returns estimated using

[48] as in Ferreira and Santa-Clara (2011). Results are similar using other values of γ.

data up to time s. We then calculate the portfolio return at the end of each period as

$$rp_{s+1} = w_s r_{s+1} + (1 - w_s) rf_{s+1}$$

where r_{s+1} is the realized stock market return. We start our out-of-sample analysis as always at half the sample $s_0 = 45$ and iterate the process until the end of the sample, resulting in time series of returns for portfolios constructed exploiting information in the predictors. We evaluate the performance of each predictor calculating the Sharpe ratio and the certainty equivalent returns

$$ce = \overline{rp} - \frac{\gamma}{2}\sigma^2(rp)$$

where $\overline{rp}$ is the sample mean portfolio return and $\sigma^2(rp)$ is the sample variance portfolio return. We report the change in certainty equivalents relative to investing with the prevailing mean forecast of the equity risk premium, as done in Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), Cenesizoglu and Timmermann (2012).[49] The certainty equivalent gain can be interpreted as the fee an investor would be willing to pay to exploit the information in the predictor variable. We also report the Sharpe ratio values and gains. Table 1.13 summarizes the results.

We notice that the portfolios constructed using my cash M&A dividend-price ratio as predictor outperform both those formed using the standard dividend-price ratios and those formed using the prevailing mean by a substantial amount. The certainty equivalent gain is 1.88% for our trading strategy while negative for the other two. Looking at the Sharpe ratio gains we confirm the great performance of my measure. The Sharpe ratio of our trading strategy (26.55%) is almost twice that of a strategy based

[49] Marquering and Verbeek (2004) report the average realized utility in absolute terms, while Shanken and Tamayo (2012) and Johannes et al. (2014) report the statistics within a Bayesian framework.

Table 1.13: Certainty equivalent gains, Sharpe ratios and Sharpe ratio Gains of the various trading strategies. This table presents out-of-sample portfolio choice results at annual frequencies from predictive regressions of the standard CRSP dividend-price ratios (measures 1 and 2) and of my cash M&A measure (measure 3). The certainty equivalent (Sharpe ratio) gains are the difference between the CE (Sharpe Ratios) using the equity risk premium forecast of the various predictors relative to using the historical mean as forecast and setting $\gamma = 2$. Annual data, out-of-sample 1972-2012.

Predictor	**CE Gains**	**Sharpe Ratio**	**Sharpe ratio gains**
1. Dividends with reinvestment	-0.79%	1.99%	-0.1182
2. Dividends without reinvestment	0.10%	8.94%	-0.0487
3. Dividends inclusive of M&A cash	**1.88%**	**26.55%**	**0.1274**

on the historical mean (13.81%). Both trading strategies that exploit the information in the standard dividend-price ratios underperform the simple trading strategy based on the historical mean, as already noted by Ferreira and Santa-Clara (2011).

In conclusion, a real-time trading strategy that uses excess returns predicted with my dividend-price ratio inclusive of M&A cash dividends outperforms trading strategies that use the prevailing mean or excess return forecasts implied by standard dividend-price ratios to construct portfolios.

1.7 Conclusions

M&A cash dividends have become an important component of total cash flows received by shareholders over the last 45 years. When these cash flows are taken into account, I find strong dividend growth, consumption growth and return predictability at the aggregate level, both in-sample and out-of-sample. In this chapter I show that variation in the dividend-price ratio is substantially explained by changes in expected dividend growth, in contrast to previous findings. This implies that we can now trace price movements back to news about future cash flows. Related to this point, I find that M&A cash dividends are the only component with explanatory power for future returns, while net equity issues are irrelevant when M&A cash dividends are taken into account. I also find that return predictability is stronger during periods of low M&A activity, related to recessions, and that a dividend-price ratio that includes M&A cash dividends performs better, both in-sample and out-of-sample, than the payout ratios of Boudoukh et al. (2007) and Robertson and Wright (2006).

At the industry level I find return predictability only for those industries with the largest M&A cash dividends, while dividend growth predictability is pervasive across all industries. Evaluating predictability jointly as suggested by Cochrane (2008), it is clear that cash dividends from mergers and acquisitions cannot be neglected because they play a fundamental role in explaining variation in stock market prices and cash flows, both at the aggregate and industry levels.

1.8 Bibliography

Allen, Franklin, and Roni Michaely, 2003, Payout policy, *Handbook of the Economics of Finance* 1A.

Andrade, Gregor, Mark Mitchell, and Erik Stafford, 2001, New evidence and perspectives on mergers, *The Journal of Economic Perspectives* 15(2), 103–120.

Ang, Andrew, and Geert Bekaert, 2007, Stock return predictability: Is it there?, *The Review of Financial Studies* 20(3), 651–707.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risk model for asset prices, *Critical Finance Review* 1, 183–221.

Beeler, Jason, and John Y. Campbell, 2012, The long-run risk model and aggregate asset prices: An empirical assessment, *Critical Finance Review* 1, 141:182.

Bollerslev, Tim, Lai Xu, and Hao Zhou, 2013, Stock return and cash flow predictability: The role of volatility risk, *Working Paper* .

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the importance of measuring payout yield: implications for empirical asset pricing, *The Journal of Finance* 62(2), 877–915.

Brealey, Richard A., and Stewart C. Myers, 2003, *Principles of Corporate Finance* (McGraw-Hill).

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1996, *The Econometrics of Financial Markets* (Princeton University Press).

Campbell, John Y., and Robert J. Shiller, 1988a, The dividend-price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies* 1(3), 195–228.

Campbell, John Y., and Robert J. Shiller, 1988b, Stock prices, earnings and expected dividends, *The Journal of Finance* 43(3), 661–676.

Campbell, John Y., and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *The Review of Financial Studies* 21(4), 1509–1531.

Campbell, John Y., and Motohiro Yogo, 2006, Efficient tests of stock return predictability, *Journal of Financial Economics* 81(1), 27–60.

Cenesizoglu, Tolga, and Allan Timmermann, 2012, Do return prediction model add economic value?, *Journal of Banking and Finance* 36, 2974–2987.

Chen, Long, 2009, On the reversal of return and dividend growth predictability: A tale of two periods, *Journal of Financial Economics* 92(1), 128–151.

Chen, Long, Zhi Da, and Xinlei Zhao, 2013, What drives stock price movements, *The Review of Financial Studies* 26(4), 841–876.

Cochrane, John H., 1992, Explaining the variance of price-dividend ratios, *The Review of Financial Studies* 5(2), 243–280.

Cochrane, John H., 2008, The dog that did not bark: a defense of return predictability, *The Review of Financial Studies* 21(4), 1533–1575.

Cochrane, John H., 2011, Presidential address: Discount rates, *The Journal of Finance* 66(4), 1047–1108.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *The Journal of Finance* 63(4), 1977–2011.

Dangl, Thomas, and Michael Halling, 2012, Predictive regressions with time-varying coefficients, *Journal of Financial Economics* 106(1), 157–181.

Eaton, Gregory W., and Bradley S. Paye, 2013, Net cash disbursements to equity holders: Implications for stock return predictability and long-horizon consumption and portfolio decisions, *Working Paper* .

Engsted, Tom, and Thomas Q. Pedersen, 2010, The dividend-price ratio does predict dividend growth: international evidence, *Journal of Empirical Finance* 17(4), 585–605.

Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: changing firm characteristics or lower propensity to pay?, *Journal of Financial Economics* 60(1), 3–43.

Fama, Eugene F., and Kenneth R. French, 2002, Testing trade-off and pecking order predictions about dividends and debt, *The Review of Financial Studies* 15(1), 1–33.

Ferreira, Miguel A., and Pedro Santa-Clara, 2011, Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514–537.

Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin, 2003, Spurious regressions in financial economics?, *The Journal Of Finance* 58(4), 1393–1414.

Gargano, Antonio, Davide Pettenuzzo, and Allan Timmermann, 2014, Bond return predictability: Economic value and links to the macroeconomy, *Working Paper* .

Golez, Benjamin, 2014, Expected returns and dividend growth rates implied by derivatives markets, *The Review of Financial Studies* 27(3), 790–822.

Goyal, Amit, and Ivo Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science* 49(5), 639–654.

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 21(4), 1455–1508.

Harford, Jarrad, 2005, What drives merger waves?, *Journal of Financial Economics* 77(3), 529–560.

Henkel, Sam James, J. Spencer Martin, and Federico Nardari, 2011, Time-varying short-horizon predictability, *Journal of Financial Economics* 99(3), 560–580.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2007, Do industries lead stock markets?, *Journal of Financial Economics* 83(2), 367–396.

Jagannathan, Ravi, and Srikant Marakani, 2015, Price dividend ratio factors: Proxies for long run risk, *Review of Asset Pricing Studies* 5(1), 1–47.

Johannes, Michael, Arthur Korteweg, and Nicholas Polson, 2014, Sequential learning, predictability and optimal portfolio returns, *The Journal of Finance* 69(2), 611–644.

Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of present values, *The Journal of Finance* 68(5), 1721–1756.

Koijen, Ralph S.J., and Stijn Van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467–491.

Kothari, S. P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time-series analysis, *Journal of Financial Economics* 44(2), 169–203.

Larrain, Borja, and Motohiro Yogo, 2008, Does firm value move too much to be justified by subsequent changes in cash flow?, *Journal of Financial Economics* 87(1), 200–226.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth and expected stock returns, *The Journal of Finance* 56(3), 815–849.

Lettau, Martin, and Sydney Ludvigson, 2005, Expected returns and expected dividend growth, *Journal of Financial Economics* 76(3), 583–626.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *The Review of Financial Studies* 21(4), 1607–1652.

Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74(2), 209–235.

Li, Yan, David T. Ng, and Bhaskaran Swaminathan, 2013, Predicting market returns using aggregate implied cost of capital, *Journal of Financial Economics* 110(2), 419–436.

Lintner, John, 1956, Distribution of incomes of corporations among dividends, retained earnings, and taxes, *The American Economic Review* 46(2), pp. 97–113.

Maio, Paulo, and Pedro Santa-Clara, 2014, Dividend yields, dividend growth, and return predictability in the cross-section of stocks, *Journal of Financial and Quantitative Analysis* Forthcoming.

Marquering, Wessel, and Marno Verbeek, 2004, The economic value of predicting stock index returns and volatility, *Journal of Financial and Quantitative Analysis* 39(2), 407–429.

Marsh, Terry A., and Robert C. Merton, 1987, Dividend behaviour for the aggregate stock market, *The Journal of Business* 60(1), 1–40.

McCracken, Michael W., 2007, Asymptotics for out of sample tests of granger causality, *Journal of Econometrics* 140(2), 719–752.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *The Journal of Political Economy* 112(1), 1–47.

Mitchell, Mark L., and J.Harold Mulherin, 1996, The impact of industry shocks on takeover and restructuring activity, *Journal of Financial Economics* 41(2), 193–229.

Moon, Seongman, and Carlos Velasco, 2014, On the properties of regression tests of stock return predictability using dividend-price ratios, *Journal of Financial Econometrics* 12(1), 151–173.

Mulherin, J. Harold, and Audra L Boone, 2000, Comparing acquisitions and divestitures, *Journal of Corporate Finance* 6(2), 117–139.

Pesaran, M. Hashem, and Allan Timmermann, 1995, Predictability of stock returns: Robustness and economic significance, *The Journal of Finance* 50(4), 1201–1228.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *The Review of Financial Studies* 22(1), 435–480.

Piatti, Ilaria, and Fabio Trojani, 2013, Dividend growth predictability and the price-dividend ratio, *Working Paper* .

Rangvid, Jesper, Maik Schmeling, and Andreas Schrimpf, 2014, Dividend predictability around the world, *Journal of Financial and Quantitative Analysis* Forthcoming.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *The Review of Financial Studies* 23(2), 821–862.

Robertson, Donald, and Stephen Wright, 2006, Dividends, total cashflows to shareholders, and predictive return regressions, *The Review of Economics and Statistics* 88(1), 91–99.

Shanken, Jay, and Ane Tamayo, 2012, Payout yield, risk, and mispricing: A bayesian analysis, *Journal of Financial Economics* 105, 131–152.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54(3), 375–421.

van Binsbergen, Jules H., Wouter H. Hueskes, Ralph S.J. Koijen, and Evert B. Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110(3), 503–519.

van Binsbergen, Jules H., and Ralph S.J. Koijen, 2010, Predictive regressions: A present-value approach, *Journal of Finance* 65, 1439–1471.

Chapter 2

Do Private Firms' Dividends Predict Stock Returns?

Abstract

In this chapter I show that information about fundamentals of the aggregate economy derived from closely held firms help predict stock returns. I construct a new economy-wide dividend-price ratio that takes into account dividends and market capitalization of both listed (public) and non-listed (private) U.S. companies and show that it strongly predicts stock returns with in-sample and out-of-sample annual adjusted R^2 of 15.35% and 16.28%, compared to the standard dividend-price ratio values of 5.32% and -1.14%, respectively. I also find that changes in dividends of private firms lead those of public firms and that the economy-wide dividend-price ratio subsumes the standard dividend-price ratio.

2.1 Introduction

A large literature in asset pricing addresses whether stock returns are predictable by a set of economic variables.[1] Following the influential work of Campbell and Shiller (1988), the dividend-price ratio has been the most widely used predictor of stock returns.[2] However, the evidence on the ability of the standard dividend-price ratio to predict returns, whether in-sample or out-of-sample, is mixed.[3] A common feature of most of the stock market predictors proposed in the literature is that they are constructed using financial data of public firms. Yet private firms are a substantial component of the U.S. economy. Asker et al. (2015) estimate that in 2010 private U.S. firms accounted

[1] Cochrane (1992), Lettau and Ludvigson (2001), Goyal and Welch (2003), Lewellen (2004), Menzly et al. (2004), Lettau and Ludvigson (2005), Robertson and Wright (2006), Campbell and Yogo (2006), Ang and Bekaert (2007), Boudoukh et al. (2007), Cochrane (2008a), Goyal and Welch (2008), Campbell and Thompson (2008), Lettau and Nieuwerburgh (2008), Chen (2009) van Binsbergen and Koijen (2010), Ferreira and Santa-Clara (2011), Shanken and Tamayo (2012), Li et al. (2013), Golez (2014), Johannes et al. (2014), Huang et al. (2015) amongst others.

[2] Several dividend-price ratios have been proposed. See, for example, Lettau and Nieuwerburgh (2008), Chen (2009), Golez (2014).

[3] See Rapach and Zhou (2013) for an overview.

for 52.8% of aggregate non-residential fixed investment, 68.7% of private-sector employment, 58.7% of sales, and 48.9% of aggregate pre-tax profits. They also find that nearly all of the 5.7 million firms in the U.S. are private (only 0.06% are listed) and that private firms predominate even amongst larger ones: 86.4% of firms with 500 or more employees were privately held in 2010. Private firms employ workers, who consume goods, invest in machinery and human capital. Perhaps unsurprisingly, it is often small privately-held startups that develop disruptive technologies leading to aggregate productivity shocks (e.g., Uber, Dropbox).

If macroeconomic factors such as consumption, investment or productivity drive risk premia in financial markets (Cochrane (2008b)), then stock market predictors that use aggregate economic information should perform well. Moreover, if stock markets are indeed intertwined with the real economy and reflect its fundamentals, it is important to incorporate information about the status of the aggregate economy by taking into account data on private firms such as dividends or investments in the construction of stock market predictors.

Financial and accounting information on private firms are, however, either unavailable or incomplete. Only recently datasets containing such information (e.g., Capital IQ, Sageworks) have been released. Unfortunately their data coverage is limited in scope and time[4] and therefore information about individual private firms cannot yet be used in empirical studies. Private firms, by law, are not required to make their financial statements public, unless they are SEC-filers.[5] However, private SEC-filers firms are only a few hundreds and are amongst the largest private firms, inducing a strong sample bias. A solution to this problem is to use financial information of private firms, both S- and C-corporations, available from the Federal Reserve Flow of Funds Tables. Using the

[4] For example, Sageworks has data starting from 2001.

[5] These are private firms that meet one of two requirements: (a) they have $10 millions or more in assets and 500 or more shareholders (2000 after passage of the JOBS Act in 2012); (b) they have public debt outstanding.

Flow of Funds data, I extract private firms' dividends and "implicit" market capitalization and construct an economy-wide dividend-price ratio, $dp^{economy}$. Data construction is described in detail in Section 2.2.

This chapter is the first to estimate dividends and valuations of private firms and to use them, jointly with public firms' data, in the construction of an economy-wide dividend-price ratio. This economy-wide dividend-price ratio differs from the standard dividend-price ratio in two ways: it represents the total U.S. economy, rather than the subset of listed companies, and excludes intercorporate holdings in the calculation of aggregate net dividends and non-financial sector market capitalization.

The main contribution of this chapter is to highlight the fundamental role of private firms' financial information in predicting (public) stock market returns. Most of the predictors proposed in the literature are based on public firms' information (e.g., earnings-price ratio, svar, book-to-market, yield spreads), while only a few of them (e.g., ik, cay) look at aggregate economy dynamics, without explicitly disentangling the effects of the public and private components. I show that the predictive performance of the economy-wide dividend-price ratio is strong both in-sample and out-of-sample, at annual and quarterly frequency, compared to the set of cash flow predictors proposed in the literature[6] (Goyal and Welch (2008), Boudoukh et al. (2007)). I find an in-sample adjusted R^2 of 15.35% (2.86%) at the annual (quarterly) frequency. Out-of-sample R^2 is 16.28% (2.43%). I also construct a dividend-price ratio for private firms and find that most of the predictive power of the economy-wide one is attributable to private firms and that the former subsumes the public dividend-price ratio in predictive regressions. These empirical results suggest that there is valuable information in the real private economy that substantially affects stock markets' risk premia. In other words, there is a news

[6] In non tabulated results I show that $dp^{economy}$ forecasts the equity risk premium better than a set of 22 predictors proposed in the literature. Results are available upon request.

component in the private sector of the economy that is not immediately incorporated in the stock market and therefore helps predicting it. Moreover, I find that lagged values of the dividend growth of private firms Granger-cause the dividend growth of public firms, consistent with private firms being more flexible than public firms in setting their dividend policy in response to cash flow shocks. Lastly, I show that a trading strategy, rebalanced annually (quarterly), that exploits the predictive power of the economy-wide dividend-price ratio is extremely profitable with a Sharpe ratio of 36.02% (41.86%) and large certainty equivalent gains.

This chapter is related to the vast return predictability literature and, specifically, to those papers discussing the dividend-price ratio as predictor of stock returns. Lettau and Nieuwerburgh (2008) show that return predictability by financial ratios is weak due to parameter instability and propose to adjust the dividend-price ratio for shifts in the steady-state of the economy. Chen (2009) discusses the importance of dividend reinvestment in the construction of the dividend-price ratio and show that returns are predictable only post-war. van Binsbergen and Koijen (2010) highlight the importance of the dividend reinvestment issue for return predictability within a present value where both expected returns and dividends growth are modeled as latent processes. Golez (2014) adjusts the standard dividend-price ratio for changes in expected dividend growth using estimates implied by the derivatives market and show that it predicts future stock returns better both in-sample and out-of-sample. Sabbatucci (2015) shows that including dividends from M&A activity in the dividend-price ratio results in stronger predictability both in-sample and out-of-sample. Boudoukh et al. (2007) adjust the dividend-price ratio for buybacks and issues, while Robertson and Wright (2006) use data from the Flow of Funds in the construction of a total payout ratio.

This chapter also fits into the recent and growing literature on private firms, which account for a substantial share of the economy (Michaely and Roberts (2012), Asker et al. (2015), Ferre-Mensa (2015)). Michaely and Roberts (2012) compare the dividend policy of U.K. private and public firms and find that private firms smooth dividends significantly less than their public counterparts. Gao et al. (2013) compare the cash policies of public and private U.S. firms and show that despite higher financing frictions, private firms hold, on average, about half as much cash as public firms do. Asker et al. (2015) investigate whether short-termism distorts the investment decision of public firms by comparing them to private firms and find that public firms invest substantially less and are less responsive to changes in investment opportunities than private firms. Ferre-Mensa (2015) investigate the private firms' ability to disclose confidential information to selected investors and find that private firms hold lower levels of precautionary cash than similar-sized public firms.

The chapter proceeds as follows. Section 2.2 describes the data. Section 2.3 shows the predictability results, discusses the lead-lag relationship between private and public firms' dividends and compare the economy-wide dividend-price ratio with other aggregate predictors. Section 2.4 presents a trading strategy that exploits the predictive performance of the economy-wide dividend-price ratio. Section 2.5 concludes.

2.2 Data Description

I use data on public and private firms from the Federal Reserve Flow of Funds Tables over the period 1945-2013. The Federal Reserve Flow of Funds Tables report, separately, data for the financial and non-financial corporate sectors. It is thus possible to construct a "total" dividend-price ratio defined as

$$dp^{total} = \frac{corporate\ dividends}{MV\ Equity\ nonfin + MV\ Equity\ fin}$$

The numerator, the corporate business dividends (series: FA096121073), includes U.S. dividends paid by the financial business (series: FA796121073), non-financial corporate sector (series: FA106121075) and dividends paid by U.S. financial and non-financial firms to the rest of the world (series: FA266121073). These dividends include, as I will discuss later, dividends paid by private firms.

As far as the denominator is concerned, the Flow of Funds Tables decompose the market value of equity of the non-financial corporate sector (series: FL103164103) as the market value of publicly listed non-financial firms (series: FL103164113) less intercorporate holdings (series: FL103164193) plus the market value of equity of closely held non-financial firms (series: FL103164123).[7] Similarly, the market value of equity of the financial sector (series: FL793164105) can be decomposed as the sum of four components: the market value of equity of public financial firms (series: FL793164113), the market value of equity of closely held financial firms (series: FL793164123), the Federal Government equity ownership (series: FL313064105) and the Monetary Authority equity ownership (series: FL713064103). As a consequence, The dp^{total} can be rewritten as

$$\frac{corporate\, dividends}{MV\ publ\, nonfin - intercorp\, holdings + MV\ priv\, nonfin + MV\ publ\ fin + MV\ priv\ fin + MV\ govt + MV\ mon}$$

The market value of private firms is directly calculated by the Federal Reserve by multiplying the net worth (revenues) data of S-corps (C-corps) by the correspondent financial

[7] Beginning in 1996, it is estimated as the sum of the market value of non-financial C-corporations and S-corporations. The market value of S-corporations is estimated by multiplying the net worth data of S-corporations in nonfinancial industries (identified by 2-digit NAICS codes) from the IRS, SOI Table S-Corporation Returns: Balance Sheet and Income Statement Items, by Major Industry, by the average ratio of market value to net worth from Standard and Poor's Compustat for public companies in the same nonfinancial industries. The market value of C-corporations is estimated by multiplying the revenue data of companies that appear on Forbes' annual list of America's Largest Private Companies by the ratio of total market value to total revenue of public companies from Standard and Poor's Compustat with similar industry, employment, and revenue profiles. The total market value of C-corporations is split between financial and nonfinancial corporations using the same splits available from the S-corporations calculations. The total market value of C-corporations and S-corporations is adjusted downward by 25 percent to reflect the lack of liquidity of closely held shares. Prior to 1996, nonfinancial business' closely held equity is included with nonfinancial business' public corporate equities.

multiples of public firms and adjusting for the lack of liquidity. This estimate of the market value of private firms is arguably the most accurate one, given the private information available to the Federal Reserve, and it is unbiased in the sense that no discretionary assumptions have been made in order to calculate it.

I further rearrange the terms in the denominator noting that the market value of equity of public financial firms plus the market value of equity of public non-financial (net of intercorporate holdings) is the total market value of equity of public firms. Similarly, the market value of equity of closely held non-financial firms plus the market value of equity of closely held financial firms is equal to the total market value of equity of closely held firms. Pooling the Government and Monetary authority holdings together (e.g., "others"), it is possible to rewrite dp^{total} as

$$dp^{total} = \frac{corporate\ dividends}{MV\ public + MV\ private + MV\ others}$$

The corporate dividends (series: FA096121073) are the sum of public and private firms dividends, since the Flow of Funds dividends data are taken from the NIPA Tables[8], which cover the whole U.S. corporate universe. We note that the market value of government and monetary authority is negligible for most of the sample period (it is non zero only for the last 24 quarters, accounting for 1-2% over the 2009-2011 period, with a peak of 10.54% in Q1 2009 when the U.S. Government bailed out the financial sector, and now less than 0.5% after the Government disinvestment). As a consequence, I define an economy-wide dividend-price ratio as

$$dp^{economy} = \frac{Divs\ public + Divs\ private}{MV\ public + MV\ private}$$

[8] `http://www.bea.gov/faq/index.cfm?faq_id=200`. Net dividends is measured as gross dividends paid by U.S. corporations in cash or other assets, plus U.S. receipts of dividends from the rest of the world, net of dividend payments to the rest of the world, less dividends received by U.S. corporations. Quarterly dividends are seasonally adjusted at annual rates. For additional details, see `http://www.bea.gov/scb/pdf/2011/03%20March/0311_profits.pdf`

In other words, the economy-wide dividend-price ratio is the sum of public and private firms' dividends divided by the total market capitalization of public and private firms.

The aggregate stock market return is the excess market return, computed as the log return on the value-weighted CRSP index with dividends (e.g., VWRETD) minus the log risk-free rate. I compare the $dp^{economy}$ with other cash flow ratios, namely the log dividend-price ratio (dp^{public}), log dividend-yield (DY) and the log net payout yield (NPY). I also compare the $dp^{economy}$ with other "economy-wide" predictors, such as the log investment-to-capital ratio (IK) and the log consumption-wealth ratio (CAY).[9]

Table 2.1 reports the summary statistics of the data. The $dp^{economy}$ has a higher mean and a lower volatility than the dividend-price ratio of the public firms. This implies that the relative (to market capitalization) amount of dividends paid out by private firms is greater than that paid out by public firms. This is confirmed by looking at Figures 2.1 and 2.2 (top panels), which show the nominal dividend amounts of both public and private firms, extracted as the difference from the total economy dividends and the public[10] dividends at annual and quarterly frequency, respectively. It is evident that starting from the mid-90s the dividends of private firms start to be larger than those distributed by U.S. public firms, consistent with Fama and French (2001).

2.3 Empirical Results

2.3.1 Lead-lag relationship

Dividends of public firms are known to be sticky (Fama and French (2002)). A common explanation is public scrutiny related to the principal-agent conflict between managers and shareholders. Managers are reluctant to change dividend payout policy

[9] The data on these measures are available from Amit Goyal and Michael R. Roberts's websites.
[10] from CRSP.

Table 2.1: This table provides summary statistics for the dividend-price ratio of the economy ($dp^{economy}$), dividend-price ratio of public firms (dp^{public}), dividend yield (dy), net payout yield (npy), investment-to-capital ratio (ik) and consumption-wealth ratio (cay). For each variable, the time-series average (Mean), standard deviation (Std), skewness (Skew), kurtosis (Kurt), minimum (Min) and maximum (Max) are reported. The sample period is 1946-2013.

	Mean	Std	Skew	Kurt	Min	Max
dp$^{\text{economy}}$	3.93%	1.10%	0.550	-0.381	1.99%	6.66%
dp$^{\text{public}}$	3.42%	1.41%	0.530	-0.191	1.14%	7.20%
dy	3.68%	1.57%	0.756	0.666	1.11%	8.77%
npy	11.59%	2.16%	-0.439	0.976	4.84%	16.73%
ik	3.50%	0.36%	0.284	-0.475	2.80%	4.32%
cay	0.06%	2.63%	0.383	0.006	-4.85%	7.41%

when profitability declines because they fear being punished by the market. An implication of this is that managers also tend to avoid raising dividend payments when profitability rises in order to limit future dividend cuts. An interesting yet unexplored question is whether private firms, which are not subject to the public scrutiny of investors and therefore should be more flexible in setting dividend policy, change their dividend policy more often than public firms. Unfortunately the data needed to answer this question, a long-enough time-series of dividends paid by individual private firms, is not available yet. However, if private firms have more flexibility in setting dividend policy, when the aggregate economy is hit by a shock they should be the first to change their dividend policy. In other words, we can test whether aggregate changes in private firms' dividend policy lead those of public firms.

Let $z_t = \begin{bmatrix} \Delta d_t^{public} \\ \Delta d_t^{private} \end{bmatrix}$, where $d_t = log(d_t/d_{t-1})$ for annual data and $d_t = log(d_t/d_{t-4})$ for quarterly data. We can write down the following VAR model

$$z_t = A_0 + \sum_i A_i z_{t-i} + \varepsilon_t \tag{2.1}$$

Table 2.2: This table reports the coefficients of the VAR system $z_t = A_0 + \sum_i A_i z_{t-i} + \varepsilon_t$. Annual data, sample 1947-2013. Quarterly data, sample 1952Q1-2013Q4. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

	Annual			**Quarterly**		
	i=1	*i=2*	*i=3*	*i=5*	*i=9*	*i=13*
$A_{11,i}$	.134	-.018	-.056	.166**	-.011	-.174***
	(1.06)	(-0.14)	(-0.44)	(2.31)	(-0.15)	(-2.63)
$A_{12,i}$	-.011	-.073	-.121**	.045**	-.073***	-.084***
	(-0.19)	(-1.38)	(-2.19)	(2.03)	(-3.10)	(-3.54)
$A_{21,i}$	.507*	-.386	-.028	.011	.009	-.056
	(1.68)	(-1.21)	(-0.09)	(0.05)	(0.04)	(-0.28)
$A_{22,i}$	-.055	.164	-.057	-.223***	-.081	-.082
	(-0.42)	(1.30)	(-0.43)	(-3.28)	(-1.14)	(-1.13)
Obs.		64			232	

where $A_i = \begin{bmatrix} A_{11,i} & A_{12,i} \\ A_{21,i} & A_{22,i} \end{bmatrix}$, and i=1,2,3 for yearly data (i=5,9,13 for quarterly data[11]). I then test the null hypothesis that lagged values of $\Delta d_t^{private}$ do not Granger-cause Δd_t^{public}, which is equivalent to testing that the coefficients $A_{12,i}$ are jointly equal to zero. Table 2.2 reports the VAR coefficients. The null hypothesis of no Granger-causality is rejected with a p-value of 0 at the quarterly frequency (χ^2 statistic of 26.9), and at the 10% level at the annual frequency (χ^2 statistic of 6.8). It is interesting to note the sign of the quarterly coefficients: the 1-year lagged private dividend growth loads positively on future public dividend growth while the 2- and 3-year coefficients' sign is flipped, which suggests public firms initially respond by increasing their dividends' payout but later reduce it. Figures 2.1 and 2.2 (bottom panels) compare the public and private firms' dividend growths. The dividend growth of private firms is highly volatile, indicating that private firms might have greater flexibility than public firms in adjusting their dividend policy.

[11] This is to ensure non-overlapping quarterly data.

Figure 2.1: Aggregate dividends and dividend growth (annual data). Split between private and public firms' dividends. Sample: 1946-2013.

2.3.2 Forecasting the market

I consider the standard predictive regression model,

$$r_{t+1} = \alpha + \beta X_t + \varepsilon_{t+1} \tag{2.2}$$

where r_{t+1} is the excess market return (e.g., CRSP value-weighted log return minus the log risk free rate) and X_t are four cash flow and two "economy-wide" predictors used

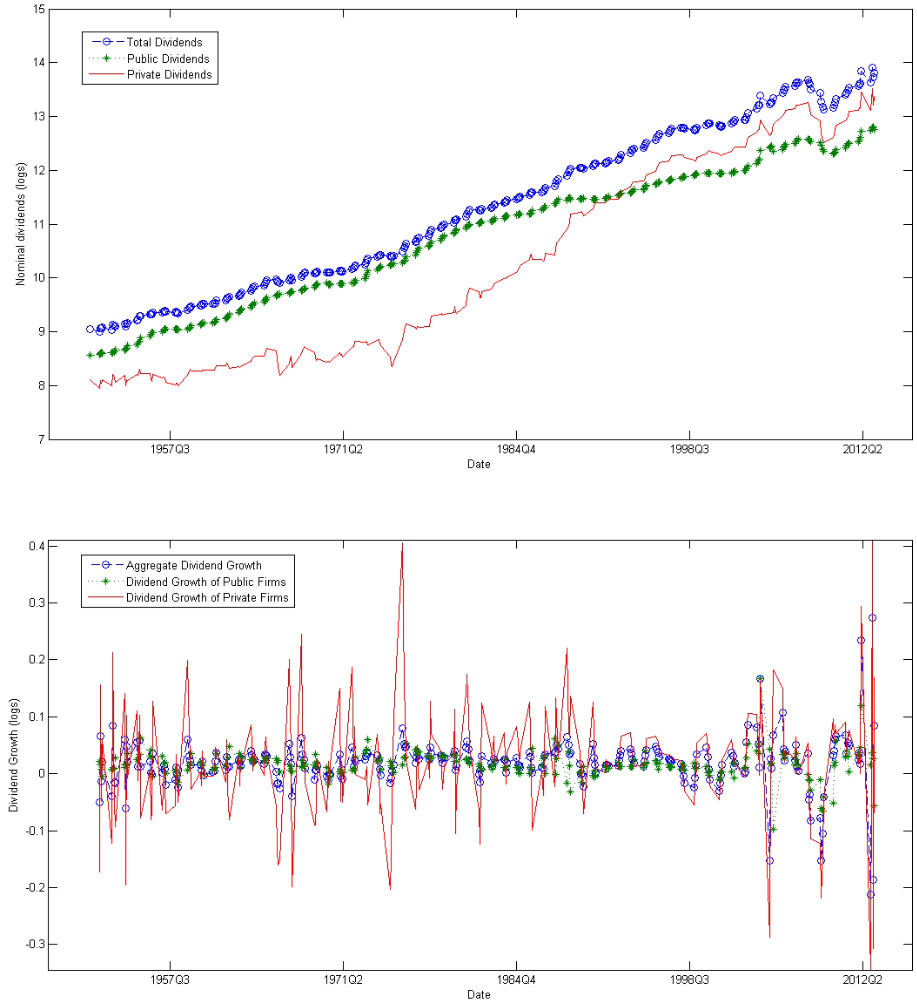

Figure 2.2: Aggregate dividends and dividend growth (quarterly data). Split between private and public firms' dividends. Sample: 1951Q4-2013Q4.

in the literature. Table 2.3 reports the results. The economy-wide dividend-price ratio substantially outperforms any other predictor at annual frequency with an adjusted R^2 of 15.35% and a statistically significant β of 0.25 (Newey-West t-statistic of 4.72). The dp^{public} and dy have approximately one-third of the explanatory power and economic magnitude of $dp^{economy}$, but the coefficients are still statistically significant (2.50 and 2.26, respectively). The investment-to-capital, cay and npy also perform quite well, with adjusted R^2 of 8.09%, 4.10% and 5.71%, respectively. Correcting the coefficients for the Stambaugh's bias, the dp^{public} becomes insignificant, while the other predictors are unaffected. At quarterly frequency, the $dp^{economy}$ performance is still strong with a coefficient equal to 6% (t-stat of 2.80) and an adjusted R^2 of 2.86%. The performances of the investment-to-capital ratio (adj. R^2 of 2.74%) and cay (adj. R^2 of 2.74%) are also good at quarterly frequency, although these predictors are only available on a shorter sample period. Moreover, the weight parameters of cay are estimated using full sample information, generating a look-ahead bias. The dp^{public} and dy performs poorly, with statistically insignificant coefficients (at the 5% confidence level) and adjusted R^2 around 1%.

It is important to note that $dp^{economy}$, ik and cay ratios are the only predictors to use information from the aggregate economy. More precisely, the investment-to-capital ratio is the ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy (Cochrane (1991)). The cay is constructed using aggregate consumption (e.g., "c"), aggregate wealth (e.g., "a"), and aggregate income (e.g., "y") (Lettau and Ludvigson (2001)). Overall, this seems to suggest that there might be a news component in the private sector of the economy that is not incorporated in the stock market and therefore helps predicting it. In Section 2.3.5 I compare these three economy-wide predictors and show that the $dp^{economy}$ is the most important one for predictability.

2.3.3 Out-of-sample forecasts

The in-sample analysis provides more efficient and stable parameter estimates since it utilizes all available information. However, in-sample predictability cannot be exploited in real-time, since it induces a look-ahead bias, and is subject to over-fitting. In addition, out-of-sample tests are less affected by small-sample size distortions such

Table 2.3: Comparison of in-sample excess return predictive regressions. The excess return is the CRSP value-weighted return over the risk free rate. Predictors data is from Goyal and Welch (2008), except the net payout yield (Boudoukh et al. (2007)). †: only 66 annual data; † † †: only 247 quarterly data; ⋆: only 65 annual and 237 quarterly data; •: 66 annual data and 247 quarterly data. Annual data: 1946-2013. Quarterly data: 1951Q4-2013Q4. Newey West 3 lags (5 lags) standard errors for annual (quarterly) regressions. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

	Annual		**Quarterly**	
	β	R^2	β	R^2
1.dp$^{\text{economy}}$	**.25*** **	**15.35%**	**.06*** **	**2.86%**
t-stat	**(4.72)**		(2.80)	
2.dp$^{\text{public}}$	.10**	5.32%	.02	0.77%
t-stat	(2.50)		(1.65)	
3. dy•	.08**	3.16%	.03*	1.10%
t-stat	(2.26)		(1.81)	
4. npy⋆	0.22***	5.71%	.04	0.48%
t-stat	(2.85)		(1.34)	
5. ik†	-.52**	8.09%	-.15***	2.74%
t-stat	(-2.62)		(-2.87)	
6. cay† † †	1.52**	4.10%	.70***	2.74%
t-stat	(2.18)		(2.92)	

as the Stambaugh bias. Hence, I investigate the out-of-sample performance of the same set of predictors discussed in the previous section.

Following Goyal and Welch (2008) and Ferreira and Santa-Clara (2011), I generate real-time, out-of-sample forecasts of excess returns using a sequence of expanding estimation windows. More precisely, I take a subsample of the first s observations $t = 1, 2, ..., s$ of the entire sample of T observations and forecast the equity risk premium using Equation 2.2. I denote the expected equity risk premium conditional on time s information by $g_{s+1} = E_{s+1|s}(r_{s+1})$. I then take the coefficients $\hat{a}^i_s$, $\hat{b}^i_s$ estimated using information available up to time s and predict the equity risk premium at time $s+1$:

$$\hat{g}_{s+1} = \hat{a}^i_s + \hat{b}^i_s X^i_s \tag{2.3}$$

where i indicate each of the twenty-one predictors. I follow this process for $s = s_0, ..., T-1$, generating a sequence of out-of-sample forecasts. In order to start the procedure, I set the initial sample size s_0 equal to one third (one fifth) of the full sample using annual (quarterly) data (e.g., 22 years and 50 quarters). Using out-of-sample forecasts based on previously available information replicates what a forecaster could have done in real time. I evaluate the performance of the out-of-sample forecasts based on the widely used Campbell and Thompson (2008) out-of-sample (OOS) R^2:

$$OOS\ R^2 = 1 - \frac{MSE_P}{MSE_M}, \tag{2.4}$$

where MSE_P is the mean square error of the out-of-sample predictions from the model:

$$MSE_P = \frac{1}{T - s_0} \sum_{s=s_0}^{T-1} (r_{s+1} - \hat{g}_{s+1})^2 \tag{2.5}$$

and MSE_M is the mean square error of the historical sample mean:

$$MSE_M = \frac{1}{T - s_0} \sum_{s=s_0}^{T-1} (r_{s+1} - \bar{r}_s)^2 \tag{2.6}$$

where $\bar{r}_s$ is the historical mean of the equity risk premium up to time s. The OOS R^2 is positive (negative) when the model estimated with predictor X_i predicts the equity risk premium better (worse) than the historical average risk premium. In other words, it measures the proportional reduction in mean squared forecast error for the predictive regression forecast relative to the historical average benchmark, where the latter corresponds to the constant expected return model ($r_s = \alpha + \varepsilon_s$) and $\bar{r}_s = \frac{1}{s}\sum_{t=1}^{s} r_t$. The statistical significance of the forecasts is evaluated using the MSE-F statistic proposed by McCracken (2007),

$$MSE - F = (T - s_0)(\frac{MSE_M - MSE_P}{MSE_P}) \tag{2.7}$$

which tests for the equality of the two MSEs and takes into account nested forecast models. As discussed in Goyal and Welch (2008), Campbell and Thompson (2008) and Rapach et al. (2010) amongst others, most of the standard predictors perform poorly out-of-sample. This is not the case with $dp^{economy}$. Table 2.4 and Figures 2.3 and 2.4 show that $dp^{economy}$ performs well both at the annual and quarterly frequency. It generates a positive OOS R^2 of 16.28% (2.43%) at the annual (quarterly) frequency with a statistically significant MSE-F statistic of 8.75 (4.91). Amongst the other predictors, at the annual frequency the *ik* and cay are the only predictors to have positive OOS R^2. The former has a positive R^2 of 7.87%, while the latter of 3.96%. At quarterly frequency, the ik still has a positive and statistically significant OOS R^2 of 1.78%, while the cay OOS R^2 is not statistically meaningful. The out-of-sample performance of all the other predictors is weak, consistent with the results of Goyal and Welch (2008). In other words, the great majority of the predictors does not improve on a forecast based on the prevailing mean of the equity risk premium. Overall, the out-of-sample results confirm the ones in-sample and suggest that private firms contain fundamental information affecting the stock market.

2.3.4 Predictability attribution: public or private dividend-price ratios?

The economy-wide dividend-price ratio $dp^{economy}$ strongly predicts the equity risk premium both in-sample and out-of-sample. It is therefore important to understand

Table 2.4: Out-of-sample excess return predictability. OOS $R^2 = 1 - \frac{MSE_P}{MSE_M}$. The MSE-F statistic of McCracken (2007) is defined as $MSE - F = (T - s_0)(\frac{MSE_M - MSE_P}{MSE_P})$. r_{t+1} is the CRSP cum dividend log excess return (e.g., $\ln(R_{t+1}^{CRSP}) - ln(R_f)$), dp_t is the log dividend-price. R_{t+1}^{CRSP} is the CRSP cum dividend gross return (VWRETD), R_f is the gross risk-free rate. Estimation sample $s_0 = 22$ years (one third of the sample) at annual frequency ($s_0 = 21$ for ik) and $s_0 = 50$ quarters (one fifth of the sample) at quarterly frequency. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

	Annual		**Quarterly**	
	OOS R^2	MSE-F	OOS R^2	MSE-F
1. $dp^{economy}$	**16.28%**	**8.75*****	**2.43%**	**4.91*****
2. dp^{public}	-1.14%	-0.51	-0.58%	-1.14
3. **dy**	-0.89%	-0.40	-0.62%	-1.21
4. **npy**	-0.45%	-0.19	-2.14%	-3.90
5. **ik**	7.87%	3.85**	1.78%	3.57**
6. **cay**	3.96%	1.86*	0.44%	0.86

whether this result is driven by the additional information content of private firms. I decompose the $dp^{economy}$ into two components: the standard public dividend-price ratio dp^{public} and a private dividend-price ratio, $dp^{private}$, which summarizes cash flow and valuation information of non-listed firms. The private economy dividend-price ratio is obtained as the orthogonal residual ε_t from the contemporaneous regression

$$dp_t^{economy} = \alpha + \beta dp_t^{public} + \varepsilon_t \tag{2.8}$$

Unfortunately the dividends of private firms are not publicly available and the market values of private firms are available only starting from 1996. In principle, it could be possible to estimate both the dividends and market capitalization of private firms as the difference between the total economy quantities and those of the public firms. However, there is an inconsistency that arises using this approach to calculate $dp^{private}$ since CRSP does not adjust for intercorporate holdings, neither in dividends nor in market capitalization. As a consequence, the resulting "difference-based" $dp^{private}$ ratio is estimated with significant noise and measurement error. The regression-based approach avoids this issue and generates a $dp^{private}$ that is less volatile (0.19 vs 0.65) and orthogonal to dp^{public}. Moreover, the regression-based $dp^{private}$ has a 83% correlation with the "difference-based" one.[12] Figure 2.5 shows the time-series of the three dividend-price ratios at annual and quarterly frequency, respectively. The time-series pattern is similar for all three dividend-price ratios until the 90s, when the dp^{public} and $dp^{economy}$ start to diverge with the former shifting down. The $dp^{private}$ remains stable, causing the $dp^{economy}$ to decline modestly up to 2000 and then bouncing back up. The dp^{public} bounces back towards the most recent period, but its mean value is still below the one of the 80s.

In order to gauge the importance of the three dividend-price ratios, I run univariate and bivariate predictive regressions of the equity risk premium using dp^{public} and $dp^{private}$. I also run the following encompassing test

$$r_{t+1} = \alpha + \beta_1 dp_t^{economy} + \beta_2 dp_t^{public} + \varepsilon_{t+1} \tag{2.9}$$

to determine whether the $dp^{economy}$ ratio subsumes the standard dp^{public}, which is equivalent to testing the null hypothesis $H_0 : \beta_1 = 0$ and $\beta_2 \neq 0$. Table 2.5 reports the regression results.

The results show that the predictive power of the dp^{public}, while being important at the annual frequency, disappears once the $dp^{economy}$ is accounted for, suggesting that the information content of the public firm is incorporated in the $dp^{economy}$. The dp^{public} has no predictive power using quarterly data, while the results for the $dp^{private}$ and $dp^{economy}$ are similar.

Overall, the results suggest that private firms have important information content

[12] The difference-based $dp^{private}$ is still strongly significant in a bivariate predictive regression together with the dp^{public} at the annual frequency, but not at quarterly frequency.

Table 2.5: Decomposition of in-sample return predictability: dp^{public} vs. $dp^{private}$. Encompassing test: $r_{t+1} = \alpha + \beta_1 dp_t^{economy} + \beta_2 dp_t^{public} + \varepsilon_{t+1}$. Excess return r_{t+1} is the CRSP value-weighted return minus the risk-free rate. Annual data, sample 1947-2013. Quarterly data, sample 1952Q1-2013Q4. Newey West standard errors (3 lags annual, 5 lags quarterly). ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

	Annual				**Quarterly**			
dp^{public}	.10**		0.10**	-.03	.02		.02	-.00
	(2.50)		(2.40)	(-0.42)	(1.65)		(1.64)	(-0.14)
$dp^{private}$		0.28**	0.28***			0.06**	0.06**	
		(2.65)	(2.81)			(1.99)	(2.02)	
$dp^{economy}$				.28***				.06**
				(2.81)				(2.02)
adj. R^2	5.32%	9.03%	14.32%	14.32%	0.77%	1.71%	2.47%	2.47%
Obs.	67	67	67	67	248	248	248	248

that is slowly incorporated into market prices and therefore generates predictability.

2.3.5 Comparison with "economy-wide" predictors

Section 2.3.2 provides evidence suggesting that the predictors with the largest economic and statistical predictive power, namely $dp^{economy}$, ik and cay ratios, are constructed with information related to the aggregate economy. In order to determine which aggregate economy information is truly important for predicting the stock market, I investigate the joint predictability of $dp^{economy}$, ik and cay. Table 2.6 reports the regression results. The $dp^{economy}$ ratio still performs well even after controlling for the other predictors, both at annual and quarterly level. The adjusted R^2 ranges from 14.50% to 18.22% when including the $dp^{economy}$, but it is only 11.64% without it. The cay and ik performances are weak at annual frequency when controlling for the $dp^{economy}$, with non statistically significant coefficients. However, at quarterly frequency both the cay and ik perform similarly to the $dp^{economy}$, with R^2 slightly above 4%.

Overall, these results show that fundamentals about the aggregate economy have strong predictive power for the equity risk premium, whether these fundamentals reflect

Table 2.6: Comparison of in-sample return predictability performance of the "economy-wide" predictors: $dp^{economy}$, cay and ik. Excess return r_{t+1} is the CRSP value-weighted return minus the risk-free rate. Annual data, sample 1947-2013. Quarterly data, sample 1952Q1-2013Q4. Newey West standard errors (3 lags for annual data, 5 lags for quarterly data). ***, ** and * denote significance at the 1%, 5% and 10% level, respectively.

	Annual				Quarterly			
$dp^{economy}$	.23***	0.21***		0.19***	.05**	.05**		.04
	(3.69)	(5.05)		(3.46)	(2.02)	(2.23)		(1.64)
cay	0.48		1.48**	0.69	.53**		.57**	.45
	(0.54)		(2.25)	(0.78)	(1.98)		(2.48)	(1.74)
ik		-0.33*	-0.48**	-0.34*		-0.12**	-0.11**	-0.10**
		(-1.84)	(-2.59)	(-1.95)		(-2.18)	(-2.45)	(-2.01)
adj. R^2	14.50%	18.22%	11.64%	17.83%	4.08%	4.27%	4.30%	5.01%
Obs.	67	66	66	66	247	248	247	247

aggregate consumption, aggregate investments or financial information of private firms.

2.4 Trading Strategy

I now assess the economic value of stock market forecasts based on $dp^{economy}$ by executing an out-of-sample trading strategy that combines the stock market with the risk-free asset, following Ferreira and Santa-Clara (2011). This strategy could have been exploited in real-time by investors. I assume investors have the standard mean-variance utility function and can invest either in the stock market or in the risk-free asset. They maximize expected utility, each period, using out-of-sample estimates of expected returns constructed using the standard return predictive regression described in Section 2.3.2 and calculate the optimal portfolio weights. More precisely, the investor's problem is

$$\max_w EU(r_p) = \max_w E(r_p) - \frac{\gamma}{2}\sigma_p^2 = \max_w w\mu_E + (1-w)rf - w^2\frac{\gamma}{2}\sigma_E^2$$

where w are the optimal weight on the stock market, μ_E is the expected return on the stock market, rf is the risk free return, γ is the risk-aversion coefficient and σ_E^2 is the variance of the stock market. Every period I calculate the optimal weights

$$w_s = \frac{\hat{\mu}_s - rf_{s+1}}{\gamma \hat{\sigma}^2{}_s}$$

where $\hat{\mu}_s$ is the forecast of the market return between s and s+1 based on the various predictors, rf_{s+1} denotes the risk-free return from time s to s+1 (known at time s), γ can assume different values (e.g., 1,2,3) and $\hat{\sigma}^2{}_s$ is the variance of the stock market returns estimated using data up to time s. I then calculate the portfolio return at the end of each period as

$$rp_{s+1} = w_s r_{s+1} + (1 - w_s) rf_{s+1}$$

where r_{s+1} is the realized stock market return. I start our out-of-sample analysis at $s_0 = 20$ ($s_0 = 50$) for annual (quarterly) data and iterate the process until the end of the sample, resulting in time series of returns for portfolios constructed exploiting the information content of the $dp^{economy}$. I evaluate the performance of the trading strategy by calculating the Sharpe ratio and certainty equivalent returns

$$ce = \overline{rp} - \frac{\gamma}{2}\sigma^2(rp)$$

where $\overline{rp}$ is the sample mean portfolio return and $\sigma^2(rp)$ is the sample variance portfolio return. I report the average portfolio return and the gain in certainty equivalent relative to investing with the prevailing mean forecast of the equity risk premium, as done in Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), Cenesizoglu and Timmermann (2012).[13] The certainty equivalent gain can be interpreted as the fee

[13] Marquering and Verbeek (2004) report the average realized utility in absolute terms, while Shanken and Tamayo (2012) and Johannes et al. (2014) report the statistics within a Bayesian framework.

Table 2.7: Average portfolio returns, certainty equivalent gains, Sharpe ratios and Sharpe ratios gains. This table presents out-of-sample portfolio choice results at annual and quarterly frequencies from predictive regressions of the equity returns on the lagged $dp^{economy}$. The certainty equivalent (Sharpe ratio) gains are the difference between the certainty equivalent (Sharpe ratios) using the equity risk premium forecast predicted by $dp^{economy}$ relative to using the historical mean. Estimation period: 20 years (one third of the sample) or 50 quarters (one fifth of the sample). Quarterly data are not annualized.

	Annual				**Quarterly**			
	$\overline{rp}$	CE Gains	SR	SR Gains	$\overline{rp}$	CE Gains	SR	SR Gains
$\gamma = 1$	27.12%	8.29%	36.02%	17.21%	7.70%	1.37%	20.93%	8.13%
$\gamma = 2$	16.04%	4.10%	36.03%	17.15%	4.49%	0.69%	20.93%	8.13%
$\gamma = 3$	12.34%	2.70%	35.95%	17.07%	3.41%	0.46%	20.93%	8.14%

an investor would be willing to pay to exploit the information in the predictor variable. I also report the Sharpe ratio values and gains. Table 2.7 summarizes the results.

The trading strategy based on the predictive power of the $dp^{economy}$ performs incredibly well. The average portfolio return ranges from 12.34% to 27.12% (3.41% to 7.70%) per year (per quarter). The certainty equivalent gain is always positive, ranging from 2.70% to 8.29% (0.46% to 1.37%) per year (per quarter). The Sharpe ratio of our trading strategy (36.02%) is more than twice that of a strategy based on the historical mean (18.95%).

In conclusion, a real-time trading strategy that uses forecasts of returns based on $dp^{economy}$ is highly profitable and outperforms trading strategies based on the prevailing mean of returns for standard values of risk aversion γ.

2.5 Conclusions

Is information about the aggregate economy important for stock market predictability? In this chapter I show that this is indeed the case. I construct an economy-wide dividend-price ratio, $dp^{economy}$, and show that it strongly predicts the equity risk premium, both in-sample and out-of-sample. Further, I present evidence of changes in private firms' dividends leading those of public firms, consistent with the hypothesis of private firms being more flexible in setting their dividend policy in response to cash flow shocks. I also show that $dp^{economy}$ subsumes the standard dp^{public} in predictive regressions and that the predictive power of $dp^{economy}$ depends on the information content of private firms.

Overall, the results presented in this chapter suggest that the aggregate economy contains valuable information for stock market predictability that is neither reflected in market prices nor incorporated in market-based predictors. This predictability, stemming from information about aggregate economy fundamentals, might have several explanations. For example, risk premia might be time-varying. Alternatively, it might be that news about these private firms' fundamentals diffuses with delay generating predictable patterns, consistent with a theory of slow diffusion of information. In other words, market participants might not understand the relevance or, simply, pay attention to aggregate economy fundamentals such as private firms' dividends. When such information is finally fully understood by the market, prices adjust, inducing predictability. Understanding why aggregate economy and, specifically, information about private firms' fundamentals generate stock market predictability is left for future research.

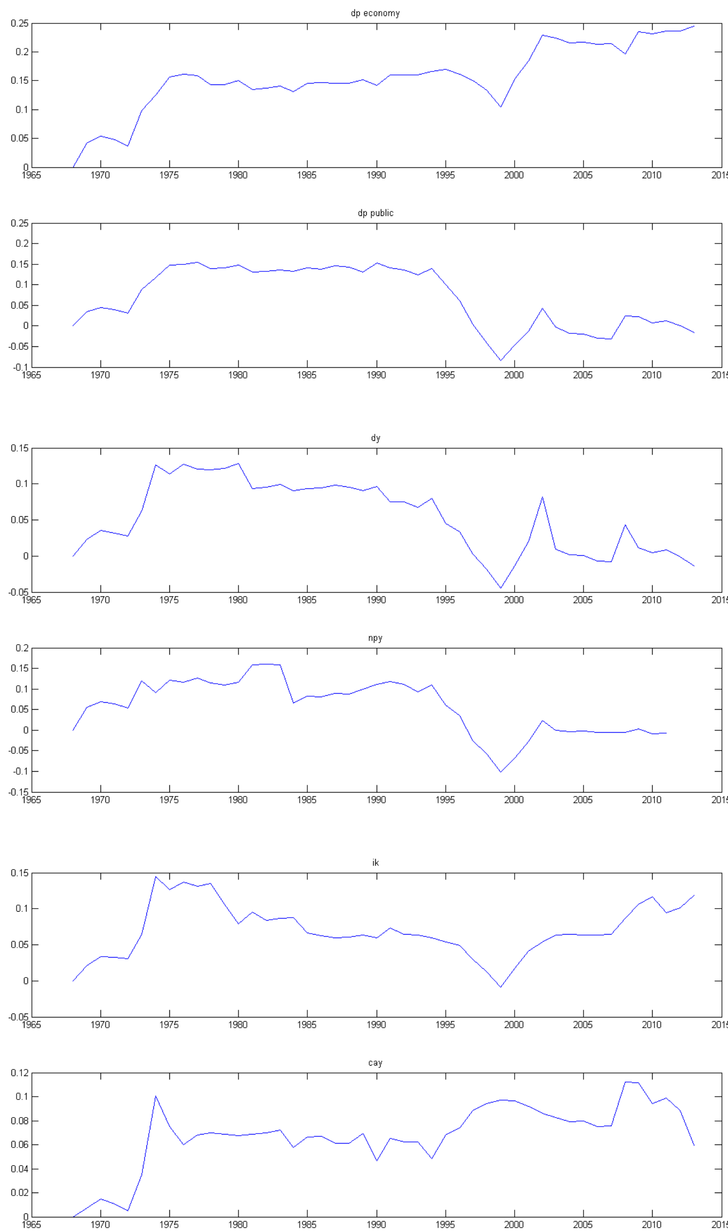

Figure 2.3: Out-of-sample predictability of excess return by the various cash flow and economy-wide predictors (annual data). Initial estimation sample for each predictor is one third of the total sample size.

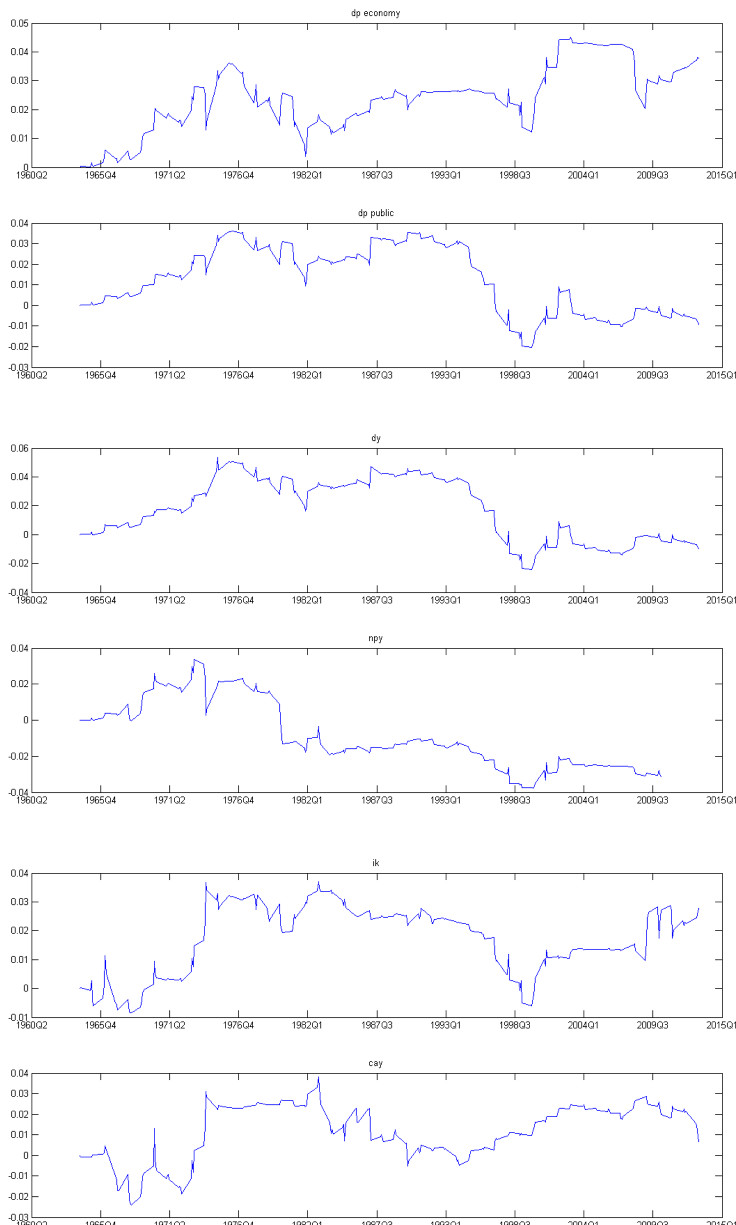

Figure 2.4: Out-of-sample predictability of excess return by the various cash flow and economy-wide predictors (quarterly data). Initial estimation sample for each predictor is one fifth of the total sample size.

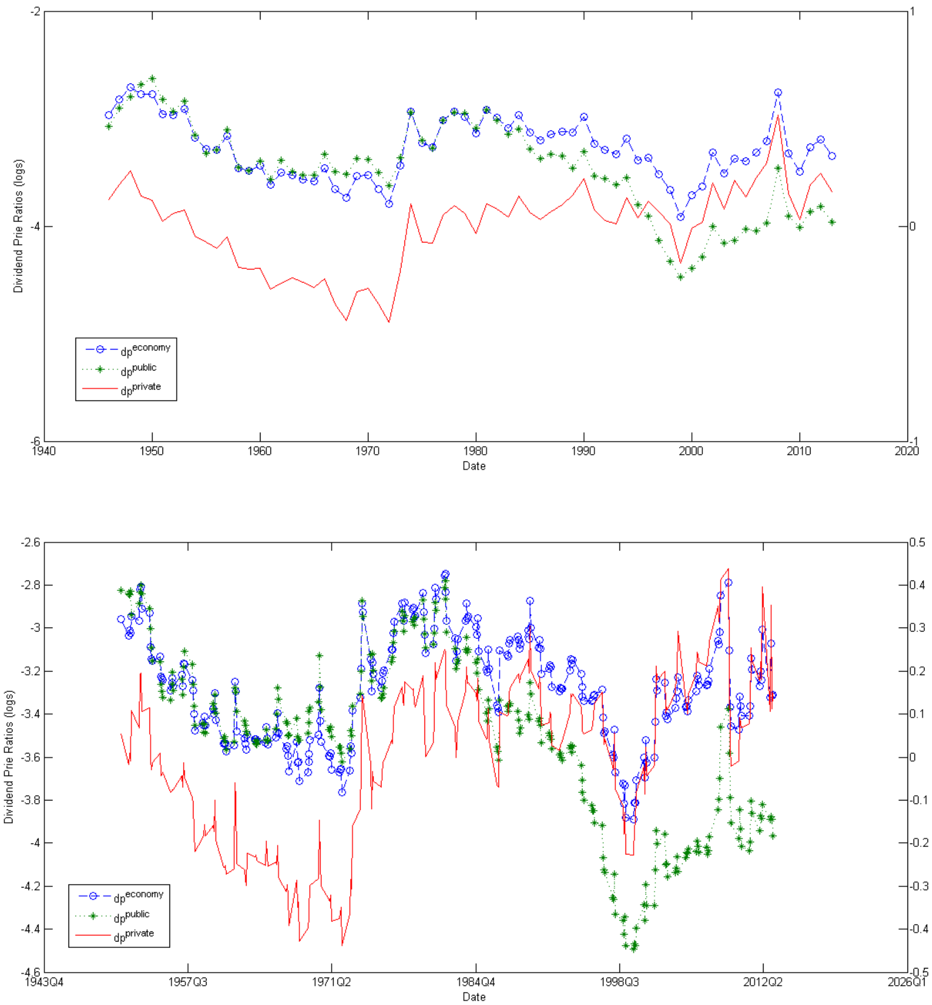

Figure 2.5: Dividend-price ratios. $dp^{economy}$, dp^{public} and $dp^{private}$. Top (bottom) panel: annual (quarterly) data. $dp^{private}$ is scaled on the right y-axis. Annual data: 1946-2013. Quarterly data: 1951Q4-2013Q4

2.6 Bibliography

Ang, Andrew, and Geert Bekaert, 2007, Stock return predictability: Is it there?, *The Review of Financial Studies* 20(3), 651–707.

Asker, John, Joan Ferre-Mensa, and Alexander Ljungqvist, 2015, Corporate investment and stock market listing: a puzzle?, *The Review of Financial Studies* 28(2), 342–390.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the importance of measuring payout yield: implications for empirical asset pricing, *The Journal of Finance* 62(2), 877–915.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies* 1(3), 195–228.

Campbell, John Y., and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *The Review of Financial Studies* 21(4), 1509–1531.

Campbell, John Y., and Motohiro Yogo, 2006, Efficient tests of stock return predictability, *Journal of Financial Economics* 81(1), 27–60.

Cenesizoglu, Tolga, and Allan Timmermann, 2012, Do return prediction model add economic value?, *Journal of Banking and Finance* 36, 2974–2987.

Chen, Long, 2009, On the reversal of return and dividend growth predictability: A tale of two periods, *Journal of Financial Economics* 92(1), 128–151.

Cochrane, John H., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *The Journal of Finance* 46(1), 209–237.

Cochrane, John H., 1992, Explaining the variance of price-dividend ratios, *The Review of Financial Studies* 5(2), 243–280.

Cochrane, John H., 2008a, The dog that did not bark: a defense of return predictability, *The Review of Financial Studies* 21(4), 1533–1575.

Cochrane, John H., 2008b, Financial markets and the real economy, *Handbook of the Equity Risk Premium* .

Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: changing firm characteristics or lower propensity to pay?, *Journal of Financial Economics* 60(1), 3–43.

Fama, Eugene F., and Kenneth R. French, 2002, Testing trade-off and pecking order predictions about dividends and debt, *The Review of Financial Studies* 15(1), 1–33.

Ferre-Mensa, Joan, 2015, The benefits of selective disclosure: evidence from private firms, *Working Paper* .

Ferreira, Miguel A., and Pedro Santa-Clara, 2011, Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514–537.

Gao, Huasheng, Jarrad Harford, and Kai Li, 2013, Determinants of corporate cash policy: Insights from private firms, *Journal of Financial Economics* 109(3), 623–639.

Golez, Benjamin, 2014, Expected returns and dividend growth rates implied by derivatives markets, *The Review of Financial Studies* 27(3), 790–822.

Goyal, Amit, and Ivo Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science* 49(5), 639–654.

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 21(4), 1455–1508.

Huang, Dashan, Fuwei Jiang, Jun Tu, and Goufu Zhou, 2015, Investor sentiment aligned: A powerful predictor of stock returns, *The Review of Financial Studies* 28(3), 791–837.

Johannes, Michael, Arthur Korteweg, and Nicholas Polson, 2014, Sequential learning, predictability and optimal portfolio returns, *The Journal of Finance* 69(2), 611–644.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth and expected stock returns, *The Journal of Finance* 56(3), 815–849.

Lettau, Martin, and Sydney Ludvigson, 2005, Expected returns and expected dividend growth, *Journal of Financial Economics* 76(3), 583–626.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the return predictability evidence, *The Review of Financial Studies* 21(4), 1607–1652.

Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74(2), 209–235.

Li, Yan, David T. Ng, and Bhaskaran Swaminathan, 2013, Predicting market returns using aggregate implied cost of capital, *Journal of Financial Economics* 110(2), 419–436.

Marquering, Wessel, and Marno Verbeek, 2004, The economic value of predicting stock index returns and volatility, *Journal of Financial and Quantitative Analysis* 39(2), 407–429.

McCracken, Michael W., 2007, Asymptotics for out of sample tests of granger causality, *Journal of Econometrics* 140(2), 719–752.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *The Journal of Political Economy* 112(1), 1–47.

Michaely, Roni, and Michael R. Roberts, 2012, Corporate dividend policies: lessons from private firms, *The Review of Financial Studies* 25(3), 711–746.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *The Review of Financial Studies* 23(2), 821–862.

Rapach, David E., and Goufu Zhou, 2013, Forecasting stock returns, *Handbook of Economic Forecasting* 2A, 328–383.

Robertson, Donald, and Stephen Wright, 2006, Dividends, total cashflows to shareholders, and predictive return regressions, *The Review of Economics and Statistics* 88(1), 91–99.

Sabbatucci, Riccardo, 2015, Are dividends and stock returns predictable? new evidence using m&a cash flows, *Working Paper* .

Shanken, Jay, and Ane Tamayo, 2012, Payout yield, risk, and mispricing: A bayesian analysis, *Journal of Financial Economics* 105, 131–152.

van Binsbergen, Jules H., and Ralph S.J. Koijen, 2010, Predictive regressions: A present-value approach, *Journal of Finance* 65, 1439–1471.

Chapter 3

Geographic Momentum

Abstract

We document geographic momentum: a positive lead-lag stock return relation between neighboring firms operating in different sectors. Geographic momentum yields risk-adjusted returns of 5-6% per year, about half that observed for industry momentum. However, while industry momentum is strongest among thinly traded, small firms, and/or those with scant analyst following, geographic momentum is *unrelated* to these proxies for information processing. We propose an explanation linking this to the structure of the investment analyst business, which is organized by sector, rather than by geographic region.

3.1 Introduction

Stock prices of firms with similar characteristics tend to move together. However, empirical studies document significant lead-lag relationships, indicating that, even among related firms, some stocks react to information before others. Indeed, the popular pairs trading strategy used by many quantitative hedge funds exploits, but apparently does not completely eliminate, this mispricing.

An important ingredient of any lead-lag strategy is categorization – i.e., sorting firms into groups based on one or more similar attributes. These attributes generally proxy for common sources of cash flow variation, and because investor clienteles tend to arise based on these attributes, they may also proxy for common sources of discount rate variation. In this chapter, we explore the possibility that common clienteles create a type of 'echo chamber,' whereby information is rapidly disseminated within groups, but less so outside them. Our main hypothesis is that scrutiny by a *common set of investors and analysts* facilitates the incorporation of information into prices, but that (even intense) scrutiny by non-overlapping sets of individuals lead to incomplete price reactions, and generates lead-lag effects.

We illustrate our intuition with a simple, descriptive model where one firm announces its earnings early (at date 1), with two others announcing later (date 2). Earnings are generated by an industry factor, a location factor, and a firm-specific factor. The early announcer shares (only) an industry factor with one of the late announcers, and (only) a location factor with the other late announcer. If markets are completely efficient, both late announcers' stock prices will respond to the earnings release of the early announcer. Generalizing this case to multiple firms, portfolios sorted by either location or industry will not exhibit momentum.

Consider now the polar opposite case, where late announcers' stock prices react only when they disclose earnings, i.e., late announcers fail to react when early announcers' earnings are released. This will generate both industry and geographic momentum, since the returns of the early announcers will lead the returns of the late announcers. Our focus is on an intermediate case, where only some firms underreact. The key source of heterogeneity is the extent to which two firms are scrutinized by a common set of investors: more overlap hastens the incorporation of common information into prices, thereby weakening any lead-lag relation in stock returns.

Our empirical tests are guided by the structure of the institutional investment business. In particular, both sell-side and buy-side analysts are typically organized along industrial, rather than geographic, lines. As a result, our model predicts lead-lag within industry groups only for the smallest, most thinly traded companies. In these cases, the potential for overlap is (trivially) limited by the small number of analysts covering any one stock. However, this is not the case for geographic peers in different sectors, for which analyst overlap is minimal, even between firms scrutinized by a large number of industry-focused analysts. Among such firms, the model assumes limited awareness of and/or communication between their respective analysts, leading to lead-lag between geographic peers, irrespective of how many analysts follow each firm *individually*.

To test these hypotheses, we run predictive regressions that include the lagged returns of a stock's industry (non-local) peers, as well as those of its local (non-industry) neighbors. Both regressors are significant. Within industry groups, we find that a 1% change in the prior month's returns forecasts continuation of about 20 basis points the following month, consistent with Moskowitz and Grinblatt (1999)'s original documentation of industry momentum. Among geographic peers we find similar patterns, with magnitudes roughly one-half to one-third the magnitude generated in our industry-sorted regressions.

Importantly however, industry momentum decreases in firm size, with the largest firms exhibiting less than half the predictability of small firms. Likewise, the magnitude of the abnormal returns displays a strong, decreasing relation with both trading volume and analyst coverage. These findings are consistent with our model: as the number of analysts following a firm increases, so too does the chance of overlap between one or more industry peers. Thus, industry momentum is expected (and is found) to be strongest for small firms with few analysts, and weaker or even absent among firms with high analyst following.

A starkly different picture emerges when examining lead-lags between geographic peers. Rather, geographic momentum appears to be completely unrelated to proxies for the richness of the information environment, a key distinction not only with industry momentum, but also with most other asset pricing anomalies. For example, when attempting to predict a firm's one-month returns from the one-month lagged area returns of its geographic neighbors, we estimate a sensitivity of 0.072 ($t = 3.78$) for the quartile with lowest trading volume, and 0.100 ($t = 4.94$) for the quartile with the most. With analyst coverage, the quartiles with strongest area-level predictability are those with the third (0.071, $t = 3.75$) and fourth (0.066, $t = 2.51$) highest levels. Cuts on firm size produce similar patterns, with geographic momentum being significant and stable in every quartile.

These non-results are likewise consistent with our stylized model, combined with our observations about the structure of the analyst business. For, even among industry leaders headquartered nearby – e.g., Google and Genetech (Bay Area), Target and General Mills (Twin Cities), or Home Depot and Coca Cola (Atlanta) – the lack of industry overlap means that few, if any analysts, will simultaneously cover both firms in the respective local pair. Thus, the recognition and/or transmission of area-level information is retarded just as much for large, heavily scrutinized large firms as it is for small, less scrutinized ones. The result is a trading strategy that, unlike anomalies involving stocks with high transactions cost and/or price impact (most of them), would appear profitable even were a large amount of capital deployed.

As an illustration, we calculate the returns of investment strategies that are implemented on a set of $20 \times 12 = 240$ city-industry portfolios, e.g., Denver-Manufacturing, Seattle-Health Care, Philadelphia-Telecommunications, and Boston-Utilities. Within every month, we rank each portfolio based on the average, one-month lagged returns of the eleven portfolios constructed in the same area, but outside the industry. The long-short hedge portfolio that buys (shorts) city/industry portfolios ranked in the top (bottom) 20% yields monthly profits of about 42 basis points per month, with a Fama-French (FF-3) alpha of 5.4% per year. Repeating this exercise, but focusing only on the top 20% of firms when ranked by market capitalization, we find profits of 43 basis points per month and a nearly identical FF-3 alpha of 5.6%.

Although our focus on geographically sorted lead-lags is novel, there is a large literature that explores non-synchronous return patterns. Atchison et al. (1987) was among the earliest to consider how these patterns generate serial correlation in portfolio returns. Lo and MacKinlay (1990) showed that size is a determinant of lead-lag effects across securities, with large firms leading small firms. Brennan et al. (1993), Badrinath et al. (1995) and Chordia and Swaminathan (2000) linked lead-lag return patterns to analyst coverage, institutional ownership and trading volume, respectively.[1] Relative to these earlier papers, our contribution is to more explicitly understand the channel linking the level of scrutiny to observed lead-lag relationships.

We are also not the first to lead-lag effects are generated by slow information diffusion. Hong et al. (2000), for example, finds that momentum – particularly when firms with negative returns are involved – weakens sharply with firm size and analyst coverage. This suggests that delayed awareness of, or reaction to, information is responsible for the sluggish price reaction observed in momentum. Other prominent examples include Cohen and Frazzini (2008), which examines the lead-lag relation between the stock returns of firms in a supply chain, and Cohen and Lou (2012), which documents underreaction between focused firms and conglomerates. We, however, are the first to explicitly tie the nature of the lead-lag relation to the *organization* of the analyst community, to examine how the lead-lag relation depends on investor scrutiny in alternative settings, and to document momentum within geographically-sorted portfolios.

Our focus on regional patterns in stock returns builds on Pirinsky and Wang (2006), which documents comovement (but not lead-lags) among firms headquartered in the same location, and on Korniotis and Kumar (2013), which uses fluctuations in state-level economic variables (e.g., house prices) as forecasting variables.[2] Both papers emphasize that discount rates may be influenced by local factors, particularly when a

[1] Numerous prior studies have examined lead-lag relationships in stock returns. Jegadeesh and Titman (1995) found that delayed reactions to common factors give rise to a size-related lead-lag effect in stock returns, while Mech (1993) and McQueen et al. (1996) showed that lead-lag effects can also be the result of non-synchronous trading or time-varying expected returns. Hou (2007) found that the lead-lag relationship between large and small firms found in the literature is predominantly an intra-industry phenomenon. Within the same industry, big firms lead small firms, and this effect is more important than the effect across industries.

[2] Other papers that examine the impact of location on asset prices and firm policies include Hong et al. (2008), Becker et al. (2011), John et al. (2011), Garcia and Norli (2012), Kumar et al. (2013), Tuzel and Zhang (2015) and Bernile et al. (2015).

firm's investors are geographically concentrated and undiversified. Our work suggests that common variation in cash flows may also be important for neighboring firms, and that the market's awareness of these regional linkages may be incomplete.

The chapter is organized as follows. Section 3.2 begins with a simple model of underreaction. The key assumption is that when two firms are covered by a common analyst, common sources of information are incorporates quickly into prices, compared to two firms with non-overlapping analysts. The remainder of the chapter is empirical, comparing lead-lag effects between industry peers (who are more likely to have common analysts) and firms headquartered nearby, but in different sectors (where analyst overlap is unlikely). In Section 3.3 we describe our sample, and in Section 3.4, we present our main results: significant lead-lags both at the industry and geographical level. How these sources of return predictability differ, one of our main interests, is described in Section 3.4.3. Specifically, while industry-level under-reaction is limited to small, thinly traded firms, area-level under-reaction persists even among large, highly scrutinized firms. Section 3.5 provides some alternative specifications and robustness checks. We conclude in Section 3.6.

3.2 Background and Theoretical Motivation

This section provides institutional details and a simple model that motivates our empirical analysis. Our characterization of the investment analyst business, discussed in subsection 3.2.1, describes prior research as well as stylized facts that demonstrate the high degree of sector concentration/focus observed among equity analysts. We then draw on these insights in subsection 3.2.2 to develop a simple model that generates cross-serial correlation at both the regional and industry level.

3.2.1 Analyst specialization by industry

Sell-side equity analysts tend to specialize by industry. This is not particularly surprising given the evidence of industry factors generating both investment rates and profitability.[3] Academic research confirms the importance of industry affiliation in the day-to-day operations, evaluation, and career paths of analysts. Kadan et al. (2012) document, for example, that industry expertise is a key dimension that defines an analyst's skill. Being recognized by institutional clients as an "all star" depends on a ranking between analysts covering firms in a given sector (Stickel (1992), Clement (1999)).

Figure 3.2 provides an indication of the importance of industry affiliation in the analyst community. The figure plots the percentage of firms (from 1993-2013) that are in an analyst's modal industry sector. For example, in 1995, the graph indicates that for the median analyst, about 83% of covered stocks were in the same industry. The interquartile range is also informative, indicating that on average, 75% of analysts spend two-thirds of their time on a single sector, with more than 25% being fully concentrated in one industry. A direct implication is that firms within the same industry tend to be covered largely by a common set of analysts. As we illustrate in the model below, this overlap in analyst coverage has an important effect on the lead-lag relation between the returns of individual stocks.

3.2.2 Industry and geographic momentum: a stylized model

We begin with a stylized model that explicitly links overlapping analyst coverage for a pair of stocks and lead-lag effects in their respective returns. The model

[3] Schmalensee (1985)'s seminal study used cross-sectional data from the year 1975 to decompose the rates of return on assets into industry, firm, and market-specific factors. Industry factors were identified as the most important in generating differences in performance between firms. Though the findings and interpretation have been challenged – most prominently by Rumelt (1991) – subsequent work, e.g., McGahan and Porter (1997) continues to identify industry affiliation as a key source of variation between business units.

assumes that in addition to firm specific shocks, there are two sources of potentially common variation – industry and regional factors. The workhorse assumption is a type of *industry focused limited attention* resulting from stock analysts' disproportionate focus on a small number of industry sectors. As our model illustrates, this focus implies that prices tend to more efficiently reflect common industry factors, especially for large, highly scrutinized firms. In contrast, the model predicts that prices may not immediately reflect regional shocks, but this mispricing will not depend on firm size or analyst coverage. The latter effect is exhibited as a lead-lag relation between neighboring firms operating in different industries.

Timing and Payoffs

The model has three dates, $t = 0,1,2$, and involves three firms $i \in \{1,2,3\}$. The interest rate is zero, and all investors are risk neutral. Each firm i realizes a liquidating dividend π_i at $t = 2$. The realization of the liquidating dividend depends on three factors: 1) an industry factor I, 2) a local factor L, and 3) a firm-specific factor ε.

There are two industries A and B, and two locations, X and Y. Firms 1 and 2 are in the same industry, and thus share industry shocks (I_A), but realize different values of the local shock, denoted L_X and L_Y, respectively. Firms 2 and 3, on the other hand, are both headquartered in location Y, but because they operate in different industries, are exposed, respectively, to I_A and I_B. Combining these assumptions, the realization of firm i's liquidating dividend at $t = 2$ is:

$$\begin{aligned}
\pi_1 &= I_A + L_X + \varepsilon_1 \\
\pi_2 &= I_A + L_Y + \varepsilon_2 \qquad (3.1)\\
\pi_3 &= I_B + L_Y + \varepsilon_3.
\end{aligned}$$

Industry, area, and firm-specific shocks are all normally distributed, i.e., $(I_A, I_B) \sim$

$N(0, \sigma_I^2 = \frac{1}{\tau_I})$, $(L_X, L_Y) \sim N(0, \sigma_L^2 = \frac{1}{\tau_L})$, and $(\varepsilon_1, \varepsilon_2, \varepsilon_3) \sim N(0, \sigma_\varepsilon^2 = \frac{1}{\tau_\varepsilon})$. The covariance between all signals, both within and across groups, is zero.

The relevant timing is shown in the timeline below. Initially, at $t = 0$, the expected liquidating dividends for each firm, and thus prices, are zero, i.e., $P^1_{t=0} = P^2_{t=0} = P^3_{t=0} = 0$. At $t = 1$, firms 1 and 3 both announce earnings, information which can be used to update the stock price of firm 2. The model ends at $t = 2$, where the realization of π_2 is observed.

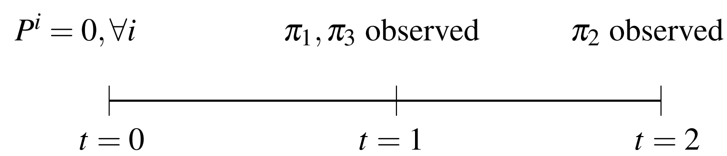

Figure 3.1: Timeline

Analyst reports

The model focuses on the stock price of firm 2, which shares an industry linkage (and only an industry linkage) with firm 1, and (only) a location linkage with firm 3. Analysts play an important role in the way stock prices are determined. Specifically, there exist a set of analysts indexed by $n \in \{1, 2, 3, ...N\}$ that cover stock 2, each of which may or may not also cover its industry peer (firm 1) or geographic neighbor (firm 3). Investors read analysts reports, and set the price of firm 2 as the expectation of π_2, conditional on the information produced by analysts that cover firm 2. Denoting the report produced by analyst n as r_n,

$$P^2_{t=1} = E[\pi_2 | (r_1, r_2, r_3, ...r_N)]. \tag{3.2}$$

The model takes a stylized view of analyst reports. In reality, analysts and investors collect and analyze information from a wide variety of sources, many of which are specific to firm being covered (e.g., talking with management, surveying customers, etc.). However, because we are interested in cross-serial correlation *between* firms, rather than *within* them, we focus on information about other companies that analysts may view as relevant. In particular, analyst n may choose to report π_1, the profit of firm 2's industry peer, and/or π_3, the profit of firm 2's geographic neighbor. There are thus four possible reports each analyst can produce: $\{\pi_1, \pi_3\}, \{\pi_1\}, \{\pi_3\}, \{\}$.

The first would correspond to an analyst that followed both firms 1 and 3, in addition to firm 2 (the subject of his report). The second and third, respectively, correspond to analysts that cover only firm 2's industry peer (firm 1) and geographic neighbor (firm 3). In the last case, the analyst covers neither firm 1 nor 3, and therefore reports neither's profits in his report.

Because investors of firm 2 read all available reports, they form expectations using the union of all information produced by the analyst community. Thus, from a pricing standpoint, it would make no difference whether all the information came from one analyst (e.g., $r_1 = \{\pi_1, \pi_3\}, r_2 = \{\}$), or whether the information is spread across analysts (e.g., $r_1 = \{\pi_1\}, r_2 = \{\pi_3\}$). The same intuition applies for more than two analysts, and for different values for the union of all reports. For example, the price formed with the set of reports ($r_1 = \{\pi_1\}, r_2 = \{\pi_1\}, r_3 = \{\pi_1\}, r_4 = \{\pi_1\}$) would be the same as that formed as with reports ($r_1 = \{\pi_1\}, r_2 = \{\}, r_3 = \{\}, r_4 = \{\}$).

Price and returns

Using the factor structure in Equation 3.1 as given, the stock price of firm 2 at $t = 1$ can take on four possible values:

$$P^2_{t=1} = \begin{cases} 0 & \text{if neither } \pi_1 \text{ nor } \pi_3 \text{ reported,} \\ \pi_1 \left(\frac{\sigma_I^2}{\sigma_I^2+\sigma_L^2+\sigma_\varepsilon^2} \right) & \text{if only } \pi_1 \text{ reported,} \\ \pi_3 \left(\frac{\sigma_L^2}{\sigma_I^2+\sigma_L^2+\sigma_\varepsilon^2} \right) & \text{if only } \pi_3 \text{ reported,} \\ \pi_1 \left(\frac{\sigma_I^2}{\sigma_I^2+\sigma_L^2+\sigma_\varepsilon^2} \right) + \pi_3 \left(\frac{\sigma_L^2}{\sigma_I^2+\sigma_L^2+\sigma_\varepsilon^2} \right) & \text{if both } \pi_1 \text{ and } \pi_3 \text{ reported.} \end{cases}$$

In the first case when neither π_1 nor π_3 is reported by the analyst community, no updating occurs, and $P^2_{t=1} = 0$. In the second case, only the industry signal is reported; here, $P^2_{t=1}$ is efficient with respect to industry information (π_1), but inefficient with respect to the geographical shock reflected by π_3. The third case is the converse, with $P^2_{t=1}$ capturing the impact of geographic, but not industry, information. The final case corresponds to the fully efficient case, where both industry and geographic shocks are appropriately incorporated into the stock price of firm 2.

We wish to characterize the conditional expected return of firm 2 from $t = 1$ to $t = 2$, using either the $t = 0$ to $t = 1$ return of firm 1 as the conditioning variable, $E[P^2_{t=2} - P^2_{t=1} | P^1_{t=1}]$, or the return of firm 3 over the same horizon, $E[P^2_{t=2} - P^2_{t=1} | P^3_{t=1}]$. The first corresponds to cross-serial correlation between industry peers (industry momentum), and the second to cross-serial correlation between local neighbors that are in different industries (geographic momentum).

Calculating these quantities requires the probabilities for the prices given above. However, the fact that π_1 and π_3 are statistically independent allows us to take a notational shortcut. Rather than having to specify prices for each price realization (four probabilities), all that is needed is the probability of π_1 being reported, irrespective of whether π_3 is reported, and vice versa. Denote these, respectively, as $p_1(N)$ and $p_3(N)$. We will later be explicit about how $p_1(N)$ and $p_3(N)$ are expected to vary with the number of analysts N, as well as potentially with firm size, but for now we treat them as constant.

Industry momentum occurs when $cov(P^1_{t=1} - P^1_{t=0}, P^2_{t=2} - P^2_{t=1}) = cov(\pi_1, \pi_2 - P^2_{t=1}) > 0$. Expanding this using the factor structure given in Equation 3.1, we have

$$\begin{aligned}
cov(\pi_1 - 0, \pi_2 - P^2_{t=1}) &= cov(\pi_1, \pi_2) - cov(\pi_1, P^2_{t=1}) \\
&= \sigma_I^2 - cov\Big(\pi_1, \frac{\sigma_I^2}{\sigma_I^2 + \sigma_L^2 + \sigma_\varepsilon^2} p_1(N)\pi_1 + \frac{\sigma_L^2}{\sigma_I^2 + \sigma_L^2 + \sigma_\varepsilon^2} p_3(N)\pi_3\Big) \\
&= \sigma_I^2(1 - p_1(N)).
\end{aligned}$$

Regional momentum takes a similar form:

$$\begin{aligned}
cov(\pi_3 - 0, \pi_2 - P^2_{t=1}) &= cov(\pi_3, \pi_2) - cov(\pi_3, P^2_{t=1}) \\
&= \sigma_L^2 - cov\Big(\pi_3, \frac{\sigma_L^2}{\sigma_I^2 + \sigma_L^2 + \sigma_\varepsilon^2} p_1(N)\pi_1 + \frac{\sigma_I^2}{\sigma_I^2 + \sigma_L^2 + \sigma_\varepsilon^2} p_3(N)\pi_3\Big) \\
&= \sigma_L^2(1 - p_3(N)).
\end{aligned}$$

Proposition 1. *The magnitude of industry and regional momentum 1) decreases with the probability that the relevant signal is observed, p, and 2) increases with the variance of the shock, σ.*

For any mispricing (in expectation) to occur at $t = 1$, there must be some probability that investors of firm 2 ignore relevant information conveyed in the profits of its industry ($p_1 < 1$) and/or geographic peers ($p_3 < 1$). High values for these probabilities – i.e., when investors are more attentive – imply a more efficient stock price for firm 2 at $t = 1$, and accordingly, less return predictability between $t = 1$ and $t = 2$. Moreover, shocks arising from a more volatile distribution are associated, in expectation, with stronger predictability. Intuitively, for a given probability that a signal is ignored $(1-p)$, shocks with higher volatility create a larger wedge between prices and fundamental value.

These observations will be useful when we compare the magnitudes of industry and geographic momentum in our empirical tests. We generally expect industry shocks to have more influence on cash flows than geographic shocks ($\sigma_I > \sigma_L$), but the probability that regional shocks are reported by analysts is probably less ($p_1 < p_3$). Consequently, it is an empirical question which effect dominates.

Varying the number of analysts N

To this point, we have taken $p_1(N)$ and $p_3(N)$ as given, so as to simplify the return predictability expressions. We now attempt to be more explicit about their relationship with the number of analysts (N) covering firm 2. In addition to comparing the average magnitudes observed for industry and geographic momentum, an important part of our empirical work will explore the impact of analyst coverage on both types of return predictability. As we now show, the number of analysts following a given stock has potential implications for any lead-lag relation its return may have with respect to industry and/or regional portfolios.

Recall that a report may take on four possible values: $\{\pi_1, \pi_3\}, \{\pi_1\}, \{\pi_3\}, \{\}$. Denote the probability of each, respectively, as x, y, z, and $1-x-y-z$. Let us assume that reports are written independently. Then, with N reports, the aggregate probability that π_1 is reported by at least one analyst, $p_1(N)$, is equal to $1-(1-x-y)^N$. Likewise, the analogous expression for π_3 is $1-(1-x-z)^N = p_3(N)$.

One of our key assumptions is that analysts are unlikely to cover firms operating in fundamentally different sectors, consistent with the patterns observed in Figure 3.2. Applied to the probabilities above, this implies that $x \approx z \approx 0$, which in turn implies that

$p_1(N) \approx 1-(1-y)^N$ and that $p_3(N) \approx 0$.

Two empirical implications follow. First, *industry momentum should decline with analyst coverage.* The intuition is that because analysts tend to specialize by industry, a larger number of analysts increases the probability that π_1 is reported by at least one of them. Consequently, the chance that investors of firm 2 will become aware of firm 1's earnings – allowing them to incorporate this information into prices – increases with N.

The expression allows us to be even more specific. Noting that $\frac{\partial p_1(N)}{\partial N} = -log(1-x-y)^N(1-x-y)^N \approx -log(1-y)^N(1-y)^N$, we can see that the relation between p_1 and N depends crucially on y. When the per-analyst probability of overlap (y) is high, even a small number of analysts will virtually ensure that π_1 is reported, i.e., $p_1 \approx 1$. On the other hand, for moderate or small probabilities of overlap, p_1 continues to increase even for relatively large N. For example, if $y = .15$, then $p_1(10) = 56\%$, but increases to 96% if 20 analysts are involved.

The second implication is that *geographical momentum should be relatively insensitive to analyst following.* If the probability that a given analyst covers both 2 and 3 is sufficiently small, then not only is p_3 similarly small, but is relatively insensitive to changes in N. As the mirror image to p_1, $\frac{\partial p_3(N)}{\partial N} = -log(1-x-z)^N(1-x-z)^N \approx 0$. For example, if $x+y = .01$, then with five analysts (beyond the 90^{th} percentile in the data), p_3 is still less than 5%, and for ten analysts (98^{th} percentile), the probability that π_3 is reported is less than 10%. The lack of sensitivity to N implies that geographical lead-lags may remain significant, even for firms covered by a large number of analysts.

The remainder of the chapter is empirical. After first describing the data in the section immediately following, we document the presence of geographic lead-lag relationships, and benchmark them against lead-lags between industry peers. We then test both empirical implications above, asking how analyst following (or other proxies for investor scrutiny) impact the magnitude of these estimated lead-lags.

3.3 Data and Descriptive Statistics

Firm location. Our analysis pertains to stocks headquartered in, or immediately proximate to, the twenty largest urban centers in the United States. To construct our sample, we begin with the universe of domestic common stocks (codes 10 and 11) traded on the NYSE, NASDAQ, AMEX over the period 1970-2013. Then, we assign to each firm a location variable, based on the zip code (ZIP) corresponding to its headquarter location in the COMPUSTAT database. Because COMPUSTAT lists only the zip code of the firm's *current* headquarters, we will misclassify firms that have relocated, such as Boeing, which moved its HQ from Seattle to Chicago in 2001. Though this introduces measurement error into our analysis, this works against us, i.e., the effects we estimate will be closer to zero than they would be absent headquarter misclassification.[4]

Headquarter locations are grouped by economic areas (EA), as defined by the Bureau of Economic Analysis. EAs are intended to capture local nodes of economic activity, and typically involve a main metropolitan area, along with smaller surrounding regions from which workers may commute. Examples of EAs include San Jose-San Francisco-Oakland (CA), Atlanta-Sandy Springs-Gainesville, (GA-AL) and Houston-Baytown-Huntsville (TX).[5]

Industry Classification. In addition to categorizing firms by headquarter location, we also group them by industry affiliation. Every month, we link each firm to a single Fama-French 12 industry, which groups firms by SIC designations. The indus-

[4] In Section 3.5.1, we perform some robustness checks by showing that misclassifying a small percentage of locations does not affect our results.

[5] Further details on the definition and characteristics of EAs can be found at `http://www.bea.gov/newsreleases/regional/rea/rea1104.htm`.

tries are non-durables (1), durables (2), manufacturing (3), energy (4), chemicals (5), business equipment (6), telecommunications (7), utilities (8), shops (9), healthcare (10), finance (11), and other (12). We intentionally select such relatively broad groupings in order to reduce the extent of overlap between firms classified in different industries.

Summary Statistics. Table 3.1 presents summary statistics for our sample, broken down by decade. Panel A shows the results by city (EA), and Panel B by industry. Progressing from the left to right, we see a steady increase in the number of publicly traded firms, with an average (per city) of 73 in the 1970s to 178 in the 2000s. However, this growth is unequally distributed among both cities and industries.

As shown in Panel B, "old economy" industries have dwindled since the 1970s, with declines in the number of publicly traded firms observed for non-durables, durables, manufacturing, and utilities (chemicals is virtually flat). In contrast, rapid growth is observed in business equipment (341% increase in public companies from the 1970s to 2000s), telecommunications (+273%), healthcare (+585%) and finance (+629%). Many of these same patterns are reflected in Panel A, which indicates stagnation for traditional manufacturing hubs like Detroit, St. Louis, Cleveland and Indianapolis, and a burgeoning among technology centers like Boston, Denver, Seattle, and San Francisco.

The last four columns present average monthly returns by decade for cities (Panel A) and industries (Panel B). Across industries, we observe substantial heterogeneity with, for example, the energy sector having among the highest average return of any industry in the 1970s (2.64%) and again after 2000 (1.47%), the intermediate decades being dominated by telecommunications (2.14% in the 1980s) and business equipment (2.59% in the 1990s). To some extent, these industrial patterns are reflected geographically, e.g., Houston-based firms performed very well in the 1970s and 2000s. However, the data seem to indicate regional differences in stock returns beyond industrial clustering. For example, in the 1990s, monthly stock returns of Washington

D.C.-based firms averaged almost one-half percent higher than those headquartered in Chicago (1.51% vs. 1.09%), despite neither area being heavily concentrated in a single industry. Similar geographical heterogeneity is observed in other decades, e.g., Minneapolis (1.30%) vs. Miami (0.79%) in the 2000s, Los Angeles (1.51%) vs. Atlanta (1.06%) in the 1970s, and St. Louis (1.60%) vs. Boston (0.99%) in the 1980s. Such regional differences are the foundation of our analysis.

Other variables. In our empirical analysis, we also incorporate a number of firm-specific variables, most of which are standard. Summary values are tabulated in Panel C of Table 3.1. The source of variation is at the panel level, creating variation both across firms and over time. The average firm in our sample has been public for 10.4 years, and returns 1.17% per month, with a (monthly) volatility of 18.84%. Averages for size, book-to-market and trading volume are also shown, and are consistent with prior literature (e.g., variables' breakpoints from Ken French's website). Finally, the average firm is covered by 3 analysts.[6]

[6] This is based on the number of EPS forecasts by unique analysts for a given firm-quarter.

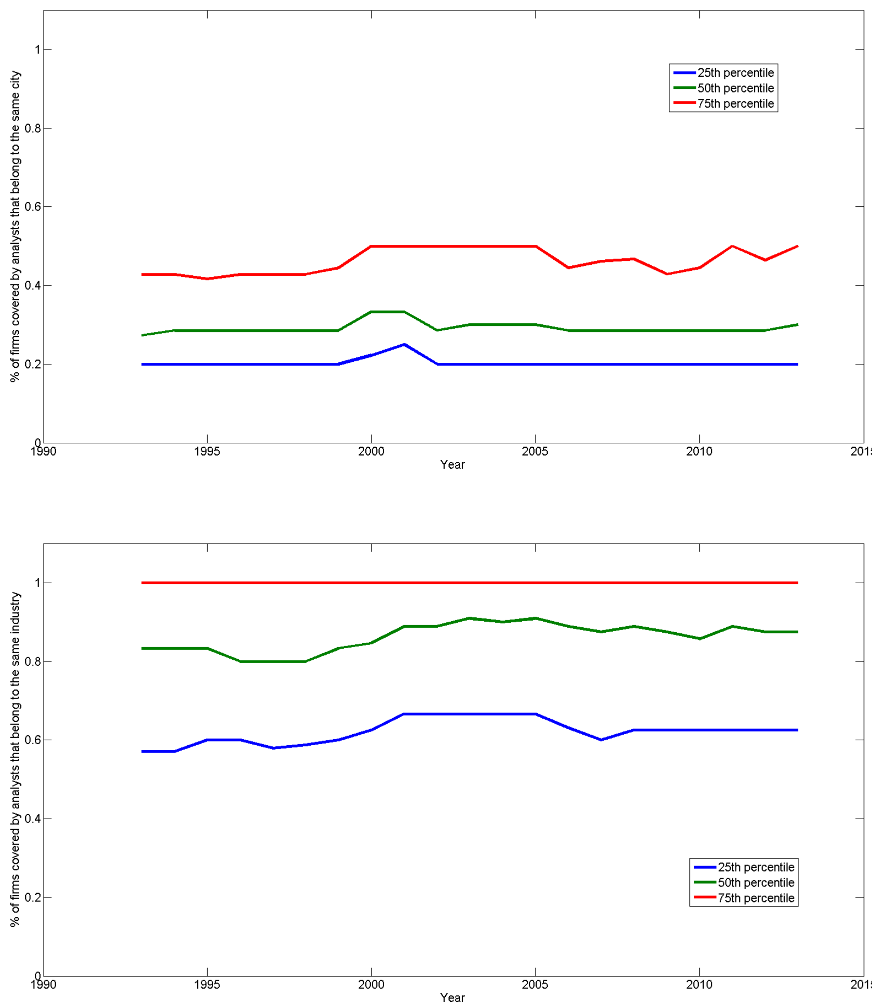

Figure 3.2: Distribution of analyst coverage by cities and industries. These graphs show the time series distribution of city (top panel) and industry (bottom panel) concentration of analyst coverage. For each analyst in every year, we identify the modal (i.e., most commonly represented) industry and city. Then, for each analyst, we identify the fraction of covered firms in these modal industries and cities and sort the analysts according to these fractions. For example, in the top panel, less than 28% of the firms covered by the median analyst in 1995 are headquartered in the same city. As another example, in the bottom panel, at least a quarter of the analysts, every year, cover only firms that belong to the same industry. Sample: 1993-2013.

Table 3.1: Descriptive statistics. **Panel A**: Average number of firms, cross-sectional mean and volatility of monthly stock returns for the twenty largest U.S. cities, by decade. **Panel B**: Average number of firms, cross-sectional mean and volatility of monthly stock returns for the twelve Fama and French (1992) industries, by decade. **Panel C**: Cross-sectional distribution of monthly returns, monthly trading volume (e.g., the total number of shares traded within a month), size (in thousands, in December 2013 dollars), number of unique analysts covering a firm (within a quarter). Monthly data, 1970-2013.

PANEL A: CITIES

	Average # of firms				*Average return (volatility)*			
	1970-1979	*1980-1989*	*1990-1999*	*2000-2013*	*1970-1979*	*1980-1989*	*1990-1999*	*2000-2013*
Atlanta	44	90	163	140	1.06% (13.76%)	1.39% (15.78%)	1.41% (18.28%)	0.89% (18.67%)
Boston	100	210	355	303	1.44% (14.87%)	0.99% (15.10%)	1.75% (18.46%)	1.01% (18.96%)
Chicago	122	173	298	296	1.12% (12.64%)	1.49% (13.50%)	1.09% (14.73%)	0.85% (13.42%)
Cleveland	58	72	93	68	1.10% (13.51%)	1.32% (13.71%)	1.15% (16.85%)	1.31% (15.39%)
Dallas	89	183	243	174	1.60% (14.75%)	0.77% (17.13%)	1.16% (20.10%)	1.14% (18.65%)
Denver	32	122	141	108	2.12% (17.29%)	-0.10% (22.48%)	1.40% (22.00%)	0.96% (19.81%)
Detroit	58	73	92	70	1.22% (13.49%)	1.26% (14.10%)	1.28% (17.15%)	1.07% (19.15%)
Houston	77	144	206	193	1.89% (13.67%)	0.60% (18.98%)	1.04% (19.28%)	1.55% (19.30%)
Indianapolis	17	29	51	40	1.10% (11.76%)	1.12% (13.21%)	1.30% (14.36%)	1.08% (15.62%)
Los Angeles	127	287	403	291	1.51% (16.90%)	1.07% (18.58%)	1.23% (23.89%)	0.93% (21.22%)
Miami	46	113	170	102	1.22% (17.10%)	0.83% (20.02%)	1.15% (24.93%)	0.79% (22.30%)
Minneapolis	45	108	181	114	1.49% (13.91%)	1.27% (16.43%)	1.42% (18.14%)	1.30% (18.07%)
New York	372	674	856	666	1.26% (15.13%)	1.14% (18.07%)	1.27% (20.39%)	0.92% (18.37%)
Orlando	13	33	39	26	1.54% (18.77%)	0.81% (16.60%)	1.80% (24.94%)	0.86% (22.19%)
Philadelphia	74	127	227	240	1.31% (14.07%)	1.38% (15.46%)	1.42% (18.02%)	0.89% (15.38%)
Phoenix	24	48	71	55	1.51% (16.37%)	0.52% (19.90%)	1.48% (21.10%)	1.16% (20.15%)
San Francisco	54	172	338	356	1.61% (14.93%)	0.75% (17.07%)	2.31% (22.94%)	0.86% (18.67%)
Seattle	14	37	63	69	1.90% (15.36%)	1.15% (15.69%)	1.81% (19.98%)	0.94% (22.34%)
St. Louis	31	42	61	49	1.04% (12.27%)	1.60% (15.15%)	1.22% (14.20%)	1.45% (18.71%)
Washington, DC	55	132	214	191	1.27% (15.45%)	1.04% (16.78%)	1.51% (19.32%)	1.17% (20.64%)

PANEL B: INDUSTRIES

	Average # of firms				*Average return (volatility)*			
	1970-1979	*1980-1989*	*1990-1999*	*2000-2013*	*1970-1979*	*1980-1989*	*1990-1999*	*2000-2013*
Consumer Non Durables	166	178	211	135	1.07% (14.07%)	1.62% (15.03%)	0.65% (19.18%)	1.22% (16.45%)
Consumer Durables	66	79	87	55	1.16% (13.86%)	1.20% (16.70%)	0.97% (19.64%)	0.88% (21.03%)
Manufacturing	278	347	357	236	1.43% (14.52%)	1.21% (16.45%)	1.09% (18.29%)	1.36% (17.70%)
Energy	64	172	147	123	2.64% (15.22%)	-0.11% (21.83%)	0.89% (18.68%)	1.47% (17.21%)
Chemicals	63	80	92	68	1.13% (12.89%)	1.35% (14.32%)	0.98% (16.52%)	1.18% (17.70%)
Business Equipment	152	469	736	671	2.01% (18.17%)	0.69% (19.54%)	2.59% (25.35%)	0.80% (24.53%)
Telecoms	26	57	114	97	1.39% (14.00%)	2.14% (16.88%)	2.24% (23.24%)	0.28% (25.38%)
Utilities	61	76	74	55	0.92% (7.45%)	1.62% (8.10%)	1.13% (8.26%)	1.14% (9.64%)
Wholesale and Retail	187	320	414	274	1.10% (14.77%)	1.04% (17.52%)	0.89% (21.23%)	1.17% (18.78%)
Healthcare	55	182	412	377	1.41% (15.64%)	1.13% (20.05%)	1.51% (22.49%)	1.51% (24.92%)
Finance	148	492	1,085	1,079	1.10% (13.80%)	1.02% (13.44%)	1.23% (13.54%)	0.87% (11.32%)
Others	187	414	536	382	1.37% (16.12%)	1.07% (18.65%)	1.10% (22.92%)	0.85% (20.23%)

PANEL C: FIRMS-MONTH

	10th	*25th*	*50th*	*75th*	*90th*	*Mean*	*Stdev*
Returns	-15.98%	-6.66%	0.00%	7.04%	17.65%	1.17%	18.64%
Trading Volume	30,900	119,900	564,000	2,855,800	11,843,200	7,831,805	68,543,163
Size	13,911	41,153	154,983	645,053	2,602,622	1,856,786	11,821,009
Book-to-Market	0.148	0.318	0.603	1.031	1.63	0.786	0.718
Analysts	0	0	0	4	11	3.22	5.87

3.4 Lead-lag Effects: Industry versus Regional Groups

This section describes our main empirical results. Subsection 3.4.1 uses Fama and MacBeth (1973) regressions to establish the presence of lead-lag effects, both between industry peers as well as between regional neighbors operating in different sectors. We then show how these lead-lags can be used to create profitable trading strategies in subsection 3.4.2. In the final subsection 3.4.3, we compare the cross-sectional patterns between industry and geographic momentum. Consistent with the model's predictions, we observe industry lead-lag effects most strongly among small, thinly traded companies, but regional predictability even (and equally) among large, heavily traded firms with substantial analyst coverage.

3.4.1 Fama-MacBeth regressions

Observations defined at the firm-month

Our benchmark specification predicts firm-level monthly stock returns using two predictors: (1) the lagged returns of a portfolio consisting of non-local industry peers, and (2) the lagged returns of a portfolio consisting of non-industry, local peers. The former portfolio is intended to capture lead-lag effects within industries (Moskowitz and Grinblatt (1999)), and the latter cross-industry lead-lag effects within cities.

We estimate the following stock-level predictive regression at the monthly level using the Fama and MacBeth (1973) methodology:

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin c,j,t} + \varepsilon_{i,c,j,t+1}, \tag{3.3}$$

where $r_{i,c,j,t+1}$ is the month $t+1$ excess return of firm i, headquartered in city c, and operating in industry j. There are two predictor variables, both measured at time t. The first is $r_{c,\notin j,t}$, the equally-weighted, lagged return of firms headquartered in city c, but operating outside firm i's industry ($\notin j$). Coefficient β_1 thus estimates the lead-lag effect within cities, but across industrial sectors. The second predictor is $r_{\notin c,j,t}$, capturing the lagged returns of firm i's industry peers (j) located outside its city ($\notin c$). Thus, β_2 measures lead-lag effects between industry (but not local) peers.[7]

Table 3.2 shows the results of estimating (3.3), with successive panels corresponding to increasing horizons of the forecasting variables. In Panel A, both the industry and area portfolios are measured over the preceding month. For example, if the dependent variable is the July 2007 return of Coca-Cola (NYSE:KO), the city portfolio would include the June 2007 returns of such Atlanta peers as Home-Depot, and the industry portfolio would include the June 2007 return of non-local bottlers such as Pepsi-Cola, headquartered in New York City.

Starting with the first column of Panel A, we see that both β_1 and β_2 are significant, with lead-lags within cities being around one-third as strong as those within industry groups. A one percent increase in a firm's lagged industry portfolio is associated with a positive return of twenty basis points the following month (t=8.21), compared to seven basis points (t=5.64) for the same change in a firm's lagged city portfolio over the full sample. Both coefficients are highly significant.

Moving down the table, the horizon over which the forecasting variable is constructed increases. For example, in Panel B, Coca-Cola's July 2007 return would be predicted from Home-Depot's 3-month return from April to June 2007, as well as Pepsi's 3-month return over the same horizon. Results from six- and twelve-month horizons are

[7] A further restriction, which we impose for interpretational simplicity, is that industry portfolios are constructed excluding firms in *any* of the 20 cities we consider in our analysis. This ensures that for firms within a given industry j, β_2 is estimated against an identical portfolio of industry peers. Although we prefer this specification, we note that operationally, this restriction makes almost no difference in our estimates. If industry portfolios are instead constructed from firm i's industry neighbors using only those headquartered in the 19 other cities ($\notin c$), we obtain nearly identical estimated for β_1 and β_2.

shown in Panel C and D, respectively.

Generally, expanding the horizon reduces forecasting ability, reflected in both smaller point estimates and, to a smaller extent, statistical significance. At three months, the sensitivity to lagged returns is less than half that observed for the one-month horizon, both at the area and industry level. However, even using trailing twelve month returns to form our forecasting variable, we still observe statistically significant lead-lags at both the industry and geographic level. But, the magnitudes (0.025 and 0.013, respectively) are relatively modest.

Columns 2 and 3 show the results separately for the first (1970-1990) and second half (1991-2013), respectively, of our sample period. Regardless of the predictive horizon, both the area- and industry-level predictors remain significant in both sub-samples. We note, however, that lead-lags between geographic peers appear to have weakened somewhat over the last two decades, whereas industry momentum has maintained relatively constant. As we will see shortly, industry momentum is driven primarily by smaller firms, whereas geographic momentum is relatively constant for firms across the size distribution. Thus, it is possible that the results in columns 2 and 3 reflect high, persistent limits to arbitrage for small firms (thus allowing industry-level mispricing to persist for decades), but gradually more efficient pricing over time for large, liquid firms.

Observations defined at the industry-city-month

An alternative way of testing for lead-lag relationships is to combine firms within the same industry and city into a single portfolio, rather then consider each firm separately. Doing so allows us to construct observations at the city-industry-month level, and then run a regression similar to equation 3.3, except that now the dependent variable is a portfolio return. Note that this aggregation reduces the number of observations to a little over 100,000, corresponding roughly to the product of the number of industries

(12), the number of cities (20), and the number of months in our sample (527). We lose about 15,000 possible city-industry-month observations to cases when a potential city-industry-month group contains zero firms.

Table 3.3 shows the results. Compared to the firm-level results, city-industry portfolio regressions give similar estimates. Industry-level lead-lags are about 10% smaller across the board, while area-level lead-lags grow by about the same amount. Consequently, when we compare the respective magnitudes, the coefficients on the geographic portfolios varies between being about one-half (panel A) to two-thirds (panel D) the size of the industry coefficients.

It is also worth noting that despite reducing the number of observations by over an order of magnitude, the statistical significance is similar between Tables 3.2 and 3.3. This suggests that the reduction in statistical power – all else equal our estimated t-statistics should decrease by about $\sqrt{\frac{1{,}611{,}9731}{111{,}942}} \approx 3.8$ – is compensated for by portfolio returns measured with less noise. In any event, the results here confirm the firm-level analysis, and provide strong evidence that for at least some firms, area-level information is incorporated into stock prices with a delay.

3.4.2 Trading profits

The results above lead themselves to a trading strategy that exploits cross-serially correlated returns between geographic neighbors. Every month, we rank each firm i not by its own lagged return (as we would in a simple momentum strategy), but by $r_{c,\notin j,t}$, the average lagged return of firms headquartered in the same region, but operating in different sectors. We use a one month horizon both for the sorting criterion (i.e., area-level stock returns are measured over a month) as well as the holding period (i.e., portfolios are reformed at the end of every month).

In Figure 3.3, we plot the value of a hypothetical dollar invested in each of three portfolios. The first, shown in blue, shows the evolution of a dollar invested in the market portfolio. Dividends are assumed to be reinvested. Against this benchmark, we also plot the 20% of firms with the highest lagged 1-month area returns (green), as well as the 20% of firms with the lowest lagged 1-month area returns (red). Note that the y-axis is displayed in natural logarithms. While the market portfolio grows by a (log) factor of over 4 during the four decades in our sample, bringing \$1 invested in the market to around \$70, \$1 invested in the lowest quintile barely exceeds \$20. On the other hand, a \$1 investment in the highest quintile performs almost an order of magnitude better, growing to approximately \$185 by 2013.

Table 3.4 makes these comparisons more formally. Starting with the first row, we see that the average monthly return for the quintile of firms surrounded by the poorest lagged returns is 74 basis points. Regressing the average returns of this portfolio against the market yields a statistically significant intercept of -26 basis points ($t = -3.21$), nearly identical to that obtained from a regression that also includes Fama and French (1993)'s size and value factors (-24 basis points, $t = -3.00$). The resulting Sharpe ratio is about 0.2, less than half what one would obtain by simply holding a market portfolio.

Proceeding down the table, we see that average returns increases steadily. The middle three groups appear fairly representative of the market as a whole, with similar average returns (CAPM alphas are small and insignificant for each group) and Sharpe ratios. However, outperformance is observed for the highest quintile, with raw monthly returns of 116 basis points, and statistically significant alphas relative to both the CAPM (t=3.19) and Fama-French three factor model (t=3.03). The Sharpe ratio for this portfolio is 0.53. (See also nearly identical results in Appendix Table A1, which displays the results when portfolios are sorted into ten, rather than five, groups.)

Foreshadowing results in the following section, the most remarkable aspect of Table 3.4 is the apparent orthogonality to traditional risk factors. To see this, note that the difference in raw returns between the first and fifth quintiles (42 basis points, $t = 3.64$) is nearly identical to the intercept estimated from either a regression against the market (47 basis points, $t = 4.15$) or against the FF-3 factors (45 basis points, $t = 3.99$).[8]

Further evidence against a risk-based explanation can be inferred from the average portfolio characteristics within each quintile, shown in the far right-hand side of the table. Here, too, we observe no trends relevant for the pattern in average returns. Firm-specific volatility is highest among the quintile with the lowest returns, followed by the second-highest quintile, then the second-lowest, highest, and then median. Size is humped shaped, with average market capitalization being highest for the middle group; indeed, we find almost identical results for a trading strategy that focuses only on the largest 20% of firms in each period (see Appendix Table A2). Book-to-market ratios display the opposite patterns, dipping in the center. These results indicate that adjusting for characteristics rather than factor loadings as in Daniel and Titman (1997), tells the same story: a geographic momentum strategy is profitable, but appears unrelated to standard risk factors.

3.4.3 Richness of the information environment

In the last section, we presented evidence of significant lead-lag effects, both at the industry and geographic level. While underreaction to industry news has been recognized since Moskowitz and Grinblatt (1999), finding a significant lead-lags within regions *between* sectors suggests an additional source of information not (completely at least) appreciated by investors. In this section, we attempt to be more precise about the specific reasons why area-level return predictability might persist.

Returning to the stylized model described in Section 3.2, the key assumption is

[8] Adjusting for momentum (Carhart (1997)), not reported in the table, likewise makes almost no difference.

that information disseminates rapidly within industries, and less so within geographic regions. The mechanism we proposed is based on overlapping coverage between equity analysts, i.e., the extent to which two firms are commonly scrutinized by a common set of individuals. Our intuition is that when two firms share multiple analysts, common sources of information are more rapidly reflected in stock prices, compared to when two firms have little if any overlap. In these latter cases, lead-lag effects can persist, whereby some firms react early to information, and others react with a delay.

One of the model's important cross-sectional predictions is that at the industry level, returns are predictable only for firms with low analyst following. The intuition is that, with few analysts following a firm, the chance of overlap – even with industry peers – is small. In contrast, the model predicts that between geographic neighbors operating in different sectors, the magnitude of lead-lags is not expected to depend (much at least) on the number of analysts. As described in subsection 3.2.2, because the chance that any one analyst co-covers two (local) firms in different sectors is so small, then even with a large number of analysts, local neighbors in different sectors are unlikely to be commonly covered. Hence, whatever return predictability exists between local peers is expected to be relatively invariant to firm size, analyst coverage, trading volume, or other proxies for investor scrutiny.

To test this hypothesis, Table 3.5, Panel A presents the results of estimating Equation 3.3, but stratified by the number of analysts. The predictive horizon is one month, so that we are predicting the returns of firm i in month t, using the one-month lagged ($t-1$) returns of either its non-local, industry ($INDUSTRY$) or local, non-industry ($CITY$) peers. Note also that the cross-sectional filters apply only to the left-hand side of Equation 3.3. By holding the explanatory portfolios constant across columns, we can relate any mispricing (reflected in the coefficients) to the variables used to partition firms in consecutive columns.

Starting first with industry-level predictability, we see a strong, declining relation with analyst coverage. Stock returns of firms with either zero (column 1) or one (column 2) identifiable analyst are most sensitive to lagged industry returns, with significantly estimated coefficients of 0.268 ($t = 7.31$) and 0.272 ($t = 6.43$) respectively. The magnitude drops by almost half to 0.158 ($t = 4.77$) for firms with between two and nine analysts, and by over half again for firms with at the ten following analysts (0.064), the latter of which is marginally significant ($t = 1.83$). The last column tests for equality between the coefficient in the first quartile (zero analysts) and that in the fourth (10 or more analysts), rejecting this at far better than the 1% level.

A different picture emerges for lead-lags at the regional level. Firms with zero analyst following actually have the *smallest* magnitude (0.057, $t = 2.79$) of any group, though this is not significantly different from that in any other column. The coefficient on the lagged area portfolio is fairly stable across columns, with sensitivities of 0.063 ($t = 1.76$), 0.071 ($t = 3.75$), and 0.066 ($t = 2.51$) for firms with progressively more analyst coverage. In contrast to the industry-level comparison, the final column indicates a p-value of 0.74, suggesting no statistically significant difference in area lead-lags for firms with low (column 1) and high (column 4) analyst coverage.

We next present cross-sectional results based on firm size (market capitalization) and trading volume.[9] There are three reasons we perform these additional sorts. First, both are strongly related to analyst coverage, but are available for the entire sample period. Second, even after 1980 (when I/B/E/S data begin), we lack a complete account of analysts that may follow and/or publish reports about specific firms. For example, buy-side analysts are not included in I/B/E/S (Cheng et al. (2006), Groysberg et al. (2013)). Finally, the I/B/E/S database is subject to alterations of recommendations, additions and deletions of records, and removal of analyst names (Ljungqvist et al. (2009)). Our hope

[9] Sorting by alternative size proxies such as book value of assets or sales gives very similar results.

therefore, is that size and trading volume capture cross-sectional variation for the general "scrutiny" of the investing community, even when data on analyst coverage is absent or unreliable.

Both Panels B and C reveal a strong, declining relation for industry momentum, similar to what we observed for analyst coverage. Industry lead-lags are very strong for the first three quartiles, with coefficients of 0.243 ($t = 7.55$), 0.232 ($t = 7.61$), and 0.189, ($t = 6.94$) respectively. In the fourth quartile however, the magnitude drops sharply. While still significant, the effect for the largest 25% of firms – representing 94% of the average total market capitalization – is less than half that observed in the first three groups (0.104, $t = 4.36$). As with Panel A, the difference in the industry coefficient between columns one and four is highly significant, as indicated in the final column.

Almost identical patterns are observed for trading volume (Panel C). Within each month, we rank firms by total trading volume, and form quartiles. The least (most) heavily traded firms are presented on the left-hand (right-hand) side of the table. The point estimates are very similar for the first three groups, with magnitudes of about 0.2, and t-statistics above seven. However, the estimate drops by almost half to 0.128 (t=4.73) in column four, indicating that industry lead-lags are weaker when trading volume is higher. This difference, too, is significant at the 1% level.

Lead-lags at the regional level, however, show little relation to either firm size (Panel B) or trading volume (Panel C). With respect to firm size, the largest point estimate is 0.101 (t=4.11), corresponding to the smallest quartile when ranked by firm size. The second largest effect is the largest quartile (0.069, t=4.30), followed by the second (0.068, t=4.03) and third (0.047, t=2.93) groups. As indicated by the last column, the difference between the coefficient estimated for the quartile of largest firms (column 1) and smallest firms (column 4), is not statistically significant (p=0.197).

Likewise, geographic momentum appears to have no discernible relation to trading volume. The strongest effect is observed among the *most heavily* traded firms (0.100, $t = 4.94$), with the second strongest coming from the lowest quartile (0.072, t=3.78). This difference is not statistically significant (p=0.247). We observe significant, though weaker effects in the second (0.056, t=2.62) and third (0.055, t=2.93) quartiles.

Figure 3.4 provides a visual summary of the results in Table 3.5. For all three cross-sectional cuts, the magnitude of the industry coefficient (red bars) declines, with the most pronounced decline coinciding with the highest quartiles. In contrast, the blue bars – representing the geographic coefficients – display no clear relation with the sorting variables. Taken together, the results in this section provide broad support for the model's predictions. While predictability at the industry level is – when it occurs – generally larger than that at the regional level, it is mainly restricted among small firms, and those with low analyst coverage and trading volume. Predictability at the regional level, though smaller on average, seems to apply equally well to firms of differing sizes, trading volumes, and analyst coverage.

Table 3.2: Predictability of individual stock returns by area and industry porfolios (Fama-MacBeth). This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

where $r_{i,c,j,t+1}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). The predictors are the 1-month (Panel A), and cumulative 3-months (Panel B), 6-months (Panel C) and 12-months (Panel D) lagged city and industry portfolio returns. Column 1: 1970-2013 (full sample). Column 2: 1970-1990. Column 3: 1991-2013. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West standard errors (5 lags). Monthly data.

PANEL A: 1-month predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.073***	.092***	.057***
	(5.64)	(4.82)	(3.26)
$\beta_{INDUSTRY}$	.199***	.185***	.212***
	(8.21)	(6.42)	(5.56)
Avg R^2	1.03%	0.79%	1.24%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
PANEL B: 3-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.035***	.047***	.024**
	(4.90)	(4.56)	(2.51)
$\beta_{INDUSTRY}$	.087***	.091***	.084***
	(7.27)	(5.60)	(4.79)
Avg R^2	1.03%	0.92%	1.13%
Observations	1,583,404	539,563	1,043,841
# time clusters	525	249	276
PANEL C: 6-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.023***	.028***	.018**
	(4.37)	(3.66)	(2.57)
$\beta_{INDUSTRY}$	.043***	.041***	.046***
	(5.43)	(4.02)	(3.77)
Avg R^2	1.03%	0.89%	1.15%
Observations	1,541,127	519,667	1,021,460
# time clusters	522	246	276
PANEL D: 12-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.013***	.019***	.008*
	(3.83)	(3.49)	(1.96)
$\beta_{INDUSTRY}$	.025***	.029***	.021***
	(5.13)	(5.09)	(2.80)
Avg R^2	1.09%	0.94%	1.21%
Observations	1,458,783	481,729	977,054
# time clusters	516	240	276

Table 3.3: Predictability of city-industry portfolios (Fama-MacBeth). This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{j,c,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

where $r_{j,c,t+1}$ is the return of industry j, in city c, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). The predictors are the 1-month (Panel A), and cumulative 3-months (Panel B), 6-months (Panel C) and 12-months (Panel D) lagged city and industry portfolio returns. Column 1: 1970-2013 (full sample). Column 2: 1970-1990. Column 3: 1991-2013. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West standard errors (5 lags). Monthly data.

PANEL A: 1-month predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.082***	.081***	.082***
	(5.90)	(3.69)	(4.82)
$\beta_{INDUSTRY}$	.173***	.166***	.179***
	(8.56)	(5.93)	(6.19)
Avg R^2	3.72%	3.12%	4.27%
Observations	111,942	52,001	59,941
# time clusters	527	251	276
PANEL B: 3-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.043***	.051***	.036***
	(4.82)	(3.81)	(3.04)
$\beta_{INDUSTRY}$	.078***	.082***	.074***
	(7.66)	(4.99)	(5.96)
Avg R^2	3.81%	3.79%	3.82%
Observations	109,063	50,160	58,903
# time clusters	525	249	276
PANEL C: 6-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.026***	.031***	.021***
	(4.73)	(3.79)	(2.93)
$\beta_{INDUSTRY}$	.038***	.039***	.038***
	(5.71)	(3.82)	(4.23)
Avg R^2	3.70%	3.68%	3.71%
Observations	106,037	48,137	57,900
# time clusters	522	246	276
PANEL D: 12-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.014***	.021***	.008
	(3.85)	(4.02)	(1.38)
$\beta_{INDUSTRY}$	.022***	.028***	.018***
	(5.15)	(5.02)	(2.72)
Avg R^2	3.84%	3.92%	3.77%
Observations	100,174	44,277	55,897
# time clusters	516	240	276

Table 3.4: Area momentum trading strategy. This table reports the performance of a trading strategy that exploits return continuation at the geographic level. Every month, we rank each firm *i* by the equally-weighted lagged return of firms headquartered in the same city, outside its industry. We then construct quintile value-weighted portfolios of the sorted firms, and hold them for one month. Portfolios are rebalanced every month. Displayed are mean returns, CAPM α, FF-3 α of each quintile portfolio. *Volatility* is the monthly standard deviation of the portfolio returns, *Mkt share* is the proportional market share of the individual portfolios, *Size* is the natural logarithm of the market value of the portfolios (in thousands), *B/M* is the book-to-market ratio of the portfolios. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West (5 lags) standard errors. Monthly data, 1971-2013.

Momentum Trading Strategy - Quantiles

	returns						portfolio characteristics			
	Mean (%)	*CAPM α*	*t-stat*	*FF-3 α*	*t-stat*	*Sharpe Ratio*	*Volatility (%)*	*Mkt share (%)*	*Size*	*B/M*
lowest city return	0.736	-0.257	-3.213	-0.242	-2.997	0.207	5.291	0.181	15.313	0.597
	0.876	-0.077	-1.117	-0.037	-0.525	0.318	4.985	0.212	15.682	0.567
	1.026	0.108	1.425	0.126	1.576	0.452	4.660	0.218	15.777	0.564
	0.949	-0.009	-0.098	-0.001	-0.019	0.366	5.030	0.206	15.625	0.571
highest city return	1.157	0.212	3.193	0.211	3.026	0.526	4.862	0.180	15.264	0.590
5-1 spread	0.421 [3.64]	0.469	[4.148]	0.453	[3.988]		2.628			

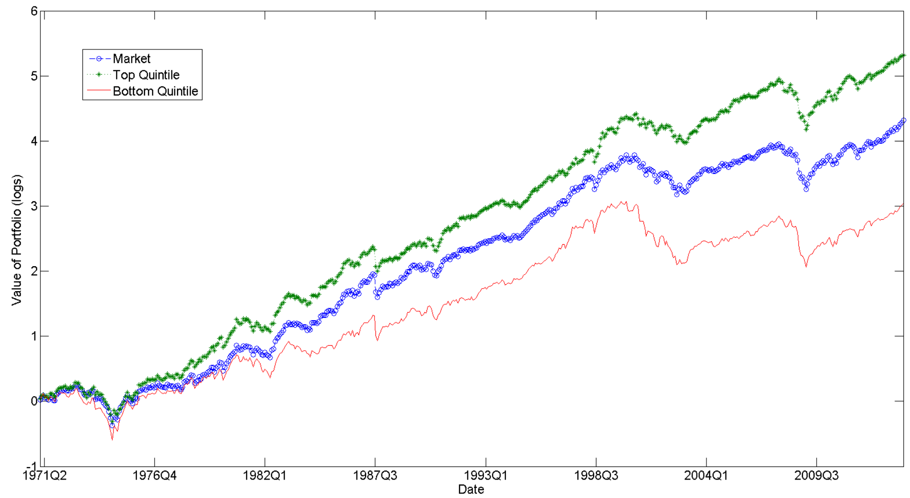

Figure 3.3: Cumulative performance of the trading strategy. This graph shows the time series evolution of a $1 investment in each of three trading strategies. The blue line is a market (S&P500) strategy, where dividends are reinvested in the market. The green (red) line represents a long-only strategy that value-weights the top (bottom) 20% of firms, when ranked by area-level stock returns the prior month. Monthly rebalancing. Numbers are in logs. Sample period: February 1971 - December 2013.

Table 3.5: Predictive regressions with cross-sectional cuts (Fama-MacBeth). This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

conditioning on the number of analysts following a firm post-1985 (Panel A) and quartiles based on firm size (Panel B) and trading volume (Panel C). $r_{i,c,j,t+1}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). Quartiles are estimated within every month. Quartile 1 is the smallest quartile. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West (5 lags) standard errors. Monthly data, 1970-2013.

PANEL A: Number of analysts					
	(0)	*(1)*	*(2-9)*	*($\geq$ 10)*	High/Low Diff
β_{CITY}	.057***	.063*	.071***	.066**	
	(2.79)	(1.76)	(3.75)	(2.51)	[0.7375]
$\beta_{INDUSTRY}$	.268***	.272***	.158***	.064*	
	(7.31)	(6.43)	(4.77)	(1.83)	[0.00***]
Avg R^2	1.09%	1.63%	1.39%	3.00%	
Observations	561,460	127,602	401,220	172,803	
# time clusters	336	336	336	336	
PANEL B: Firm Size					
	(1)	*(2)*	*(3)*	*(4)*	Big/Small Diff
β_{CITY}	.101***	.068***	.047***	.069***	
	(4.11)	(4.03)	(2.93)	(4.30)	[0.1974]
$\beta_{INDUSTRY}$	.243***	.232***	.189***	.102***	
	(7.55)	(7.61)	(6.94)	(4.30)	[0.00***]
Avg R^2	0.83%	1.50%	1.88%	2.52%	
Observations	400,088	403,689	404,079	404,117	
# time clusters	527	527	527	527	
PANEL C: Trading Volume					
	(1)	*(2)*	*(3)*	*(4)*	High/Low Diff
β_{CITY}	.072***	.056***	.055***	.100***	
	(3.78)	(2.62)	(2.93)	(4.94)	[0.2469]
$\beta_{INDUSTRY}$	.216***	.212***	.209***	.128***	
	(8.90)	(7.65)	(7.66)	(4.73)	[0.0035***]
Avg R^2	0.96%	1.21%	1.35%	1.99%	
Observations	380,839	380,929	381,466	381,322	
# time clusters	527	527	527	527	

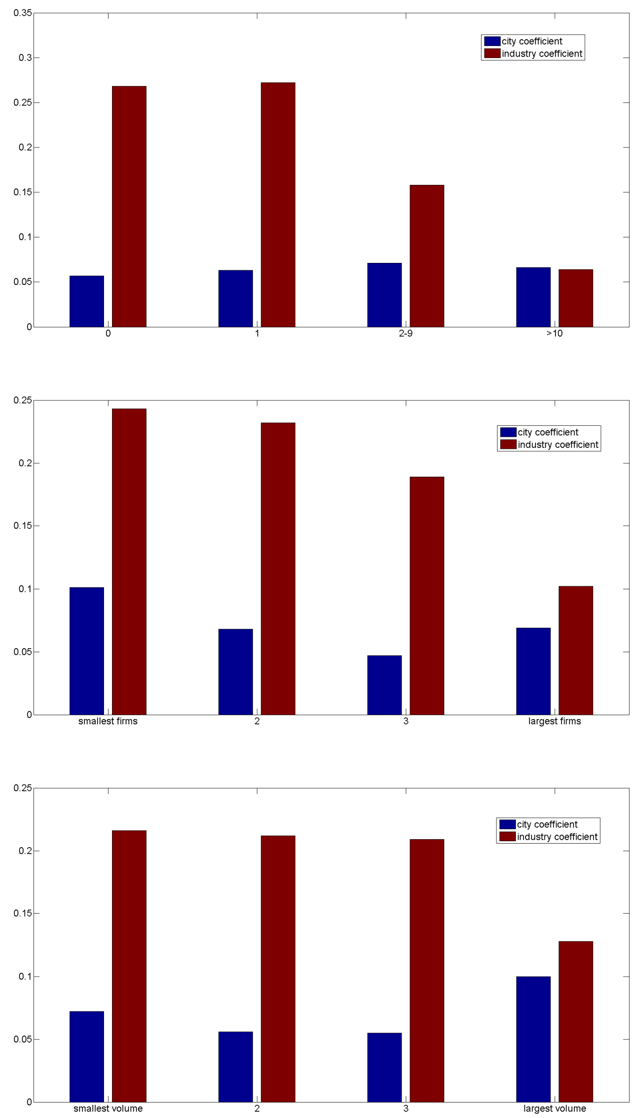

Figure 3.4: Plot of cross-sectional lead-lag coefficients. This graph corresponds to Table 3.5, and plots the lead-lag coefficients for industry (red) and area (blue) portfolios, which are estimated from one-month Fama-MacBeth predictive regressions. Cross-sectional cuts are obtained by splitting the sample into four groups based on number of analysts following a firm (top figure) and quartiles based on firm size (middle figure) and trading volume (bottom figure).

3.5 Robustness and Extensions

In this section, we present the results of a number of robustness checks and specification alternatives to our main results. The first two subsections address the possibility that our measure of a firm's headquarters may be an imperfect proxy for its location, and therefore, its sensitivity to local factors. Subsection 3.5.1 quantifies the potential impact of mis-measured headquarter locations, which may arise when firms relocate, and subsection 3.5.2 expands beyond headquarters to consider, e.g., the location of a companies' operations, manufacturing, etc. Finally, we present our main predictability results under several alternative specifications in subsection 3.5.3.

3.5.1 Misclassified headquarter locations

Our measure of firm location is its headquarters, as inferred by the ADDZIP variable in COMPUSTAT, which reports the zip code of the firm's *most recent* headquarters. Consequently and unfortunately, in most cases, we do not observe when a firm changes headquarters, resulting in a type of look-ahead bias. For example, General Dynamics moved from St. Louis to the Washington D.C. area in 1992, but the ADDZIP variable takes a value of 22042, corresponding to Falls Church, Virginia (near Washington D.C.), both for years prior to its move (pre-1992), as well as afterward (1992 and beyond).

Comprehensive data on firm headquarter changes is conspicuously absent in the finance literature, but audit studies indicate that they are fairly uncommon. Pirinsky and Wang (2006), for example, use news data to track headquarter changes from 1992-1997. Excluding firms that moved as a result of mergers or other major restructuring, as well as those moving within the same MSA, the authors estimate that between 2-3% of firms moved during this five year period, or about 0.5% per year. If we assume that all firms have this rate of relocation, then over 40 years, we expect to for about

$0.995^{43} \approx 80.6\%$ of firms to still be correctly classified in 2013, the last year of our sample. However, because the location reported in 2013 would be correct, on average, half the time (as for General Dynamics post-1992), we should expect error rates by firm-year in the range of perhaps $0.5 * (100\% - 80.6\%) = 9.7\%$. Even this, however, is probably conservative. Because our panel is disproportionately represented by large firms with long histories, and because large firms are less likely to move than small ones (as Pirinsky and Wang (2006) also shows), the percentage of misclassification is likely even lower. Most importantly, bad location data biases against our findings.

Because we have some uncertainty about the true misclassification rate, Table 3.6 presents our one-month predictive regressions under various scenarios. In the first panel, 1% of the headquarter locations are scrambled randomly, followed by successively higher percentages in each panel. For misclassification rates of 1% and 5%, the impact on the area coefficient is trivial, and are only slightly affected by misclassifications of 10%. For 20%, the magnitude is cut by one-third, although it remains statistically significant for the full sample (but not for the latter half). With half the locations assigned incorrectly (Panel E), the result vanishes entirely.

Given Pirinsky and Wang (2006)'s estimates, along with our intuition about the composition of firms throughout the sample, our best guess is an error rate in the 5-10% rate over all firm-years in the panel. If so, this suggests that the reported estimates in our prior tables are not meaningfully affected by misclassifications. On the other hand, if we are wrong by a factor of (10-20%), then our reported results should be grossed up by about 30% to account for measurement error.

3.5.2 Location beyond a firm's headquarters

Throughout the chapter, we have identified a firm's location using its headquarters. While this is both simple and observable for every firm in the sample, it ignores clear differences in the extent to which a firm's facilities, customer base, or labor force are concentrated in a particular geographic region. For example, at one end of the spectrum are retail firms with a national (or even global) presence, e.g., Wal-Mart, Home Depot, Whole Foods, and Costco which have highly dispersed stores, customers, and workers. At the other extreme are companies with most or all their operations conducted at a single location. DTE Energy, a Michigan-based utility company mentioning only Michigan and Indiana in its annual reports, and AutoDesk (mentioning only California) are at other other extreme.

The question we explore in this section is whether regional predictability is stronger for more regionally concentrated firms (e.g., AutoDesk) compared to one with a more disperse presence (e.g., Whole Foods). To obtain a more general measure of a firm's geographical presence, Garcia and Norli (2012) utilize a text-based parsing algorithm that counts the number of unique state names mentioned in the annual reports of publicly traded firms from 1994-2008. As Garcia and Norli describe, state names are often listed when describing/discussing the locations of stores, manufacturing facilities, or other operations. We follow their approach, after downloading the relevant dataset from Diego Garcia's website. For each firm we calculate the time-series average of state names over the available time period (1994-2008), and apply this measure to all years (including before 1994 or after 2008) in which data are available.[10]

Table 3.7 presents the results of our one-month Fama-MacBeth predictive regressions, when sorted by the above/below median level of geographic concentration. The first two columns correspond to the entire sample, with columns 3 and 4 (5 and 6)

[10] Because Garcia and Norli (2012) data are available only for 15 years of our 43 year panel, an extrapolative approach is required in order to apply the concentration measure to our entire sample period. Taking the time-series average of state names for each state, unfortunately, ignores dynamics. However, it is unusual for firms to become dramatically more or less concentrated over time, leading us to believe that the ranking obtained from 1994-2008 provides a good proxy for its ranking across all years. For example, the median time-series standard deviation of state counts for firms in Garcia and Norli (2012) sample is 1.34, suggesting little aggregate time-variation of geographical concentration.

to small (large) firms.[11] Firms above the median list nine states on average, compared to three states for firms below it. Note that because we use the same cutoff (5.46 states) for each of the sub-samples, the corresponding sub-sample averages are similar, but need not be identical to the aggregate sample.

In all cases, the point estimates for the more regionally concentrated firms are somewhat larger compared to their less concentrated counterparts. Small firms are associated with the biggest differential, with highly concentrated firms being twice as sensitive to lagged area returns (0.102, t=3.07) compared to firms that mention more states in their annual reports (0.052, $t = 1.92$). Although this difference is not statistically significant at conventional levels, we have experimented with other specifications and find stronger results. For example, a Fama-MacBeth regression (one month horizon) that interacts the number of states mentioned with the lagged city portfolio returns yields a p-value less than 3%. Given these suggestive results using a fairly coarse measure of regional concentration, we hypothesize that more refined measures – using establishment data from the U.S. Census would seem promising – would give even stronger results.

3.5.3 Alternative regression specifications

Panel regressions with time fixed effects

Our main empirical tests use Fama-MacBeth cross-sectional regressions. By construction, this procedure sweeps out any common "time effect" through a date-specfic intercept applied to each cross-sectional regression. However, as discussed in Petersen (2009), three potential problems remain. First, Fama and MacBeth's procedure does not address serial correlation in residuals. Second, although a unique intercept for each ($year \times month$) controls for a *uniform* effect across stocks, it is possible for sub-

[11] Note that the sample size is reduced by about 500,000 firm-month observations, corresponding to firms not in the Garcia-Norli database.

samples to be more or less sensitive to a given shocks. For example, in September 2001, airline stocks suffered far worse than did, say, stocks of food or defense firms. Third and finally, Fama-MacBeth regressions give equal weight to each cross-sectional regression when calculating the standard error of β_1 (the sensitivity to lagged area returns), thus ignoring the fact that some periods contain more information than others.

Table 3.8 shows the results of re-estimating equation 3.3, but with date fixed effects, and residuals double-clustered by firm and date (Petersen (2009)). When comparing these to the area-level lead lags to the Fama-MacBeth estimates shown in Table 3.2, we observe similar point estimates, but weaker statistical significance. As an example, the full sample coefficient (t-statistic) for our one-month predictive regression is 0.063 (3.66) in panel regressions, and 0.073 (5.64) with Fama-MacBeth. Results are also slightly stronger at the 3-month and 6-month predictive horizon, and nearly identical at the 12-month window. Note, however, that although weaker, geographic lead-lags remain significant at all horizons.

Industry momentum also weakens from a significance perspective. In the double-clustered panel estimation for example, industry lead-lags are profitable only at relatively short horizons (within three months), and all but disappears in the second half of the sample. In Fama-MacBeth regressions however, the anomaly continues to work after 1990 (though with a smaller magnitude), and remains profitable through twelve months after the sort date.

We have experimented with various specifications, in an attempt to better understand which of the three factors mentioned at the beginning of this section are most responsible for the weakened results. It turns out that firm-clustering is relatively insignificant; the standard errors reported in Table 3.8 are almost identical if this cluster is removed. Rather, the additional clustering by time is most responsible for the differences. This exercise thus indicates that accounting for remaining cross-sectional correlation within time may be relevant, and likewise suggests that our panel-generated estimates are considerably more conservative than Fama-MacBeth's methodology.

Delayed portfolio formation

In this section we examine the time it takes the information in one firm's stock price to be incorporated in the stock prices of its industry and location peers. Specifically, we examine whether there is still predictability when we skip a month between when the past returns are measured, and when the strategy is implemented. If information is transmitted relatively quickly, we expect that predictability should be largely eliminated.

Table 3.9 reports Fama-MacBeth regressions that are identical to the one-month predictive regressions reported in first panel of Table 3.2, save for the one-month skip. As the table reveals, area- and industry-level lead-lags are weaker, but they remain statistically significant. The impact of delayed portfolio formation is most severe when the predictor variables are measured over short horizons. For example, the coefficient on the area-level predictor drops from 0.073 to 0.037 (t=3.06), and that on the industry-level predictor drops from 0.199 to 0.118 (t=4.59). At longer horizons, the impact of delayed portfolio formation is less pronounced. For example, when month $t-13$ to $t-1$ returns are used to forecast stock returns beginning at month t, the coefficient on the area portfolio is nearly identical (0.011, t=3.43) to that estimated without the one-month skip (i.e., month $t-12$ to t returns are used as predictors). Industry momentum suffers a similarly modest decline, dropping from 0.025 to 0.019 (t=3.52).

Together, these results suggest that prices remain inefficient for at least a month after portfolio formation, suggestive a fairly long delay in processing industry or area-specific information.

Value-Weighted Portfolios

In the main results of Table 3.8 we construct the local and industry portfolios by equally weighting firms within each group, similar to Pirinsky and Wang (2006). As a robustness check, in Table 3.10 we re-estimate our one-month Fama-MacBeth predictive regression using value-weighted local and industry portfolios. Whereas both industry and geographic momentum remain statistically significant, the magnitudes are 30-40% smaller, depending on the horizon. In retrospect, this result is intuitive. If the goal is to measure local economic fundamentals using portfolio returns, an equally-weighted basket is more likely to be informative, compared to one that puts disproportionate weight on a few large firms (e.g., Dallas's ExxonMobil, Seattle's Amazon, etc.), especially given that they are less likely to be regionally concentrated.

Table 3.6: Misclassified locations. This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

where $r_{i,c,j,t+1}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). In every panel, the predictors are the 1-month lagged "random" city and industry portfolio returns. 1% of the locations are randomized in Panel A, 5% in Panel B, 10% in Panel C, 20% in Panel D, 50% in Panel E. Column 1: 1970-2013 (full sample). Column 2: 1970-1990. Column 3: 1991-2013. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West standard errors (5 lags). Monthly data.

PANEL A: 1% misclassified locations			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.073***	.090***	.057***
	(5.61)	(4.66)	(3.34)
$\beta_{INDUSTRY}$	.199***	.185***	.212***
	(8.21)	(6.41)	(5.57)
Avg R^2	1.03%	0.79%	1.24%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
PANEL B: 5% misclassified locations			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.064***	.076***	.054***
	(5.28)	(4.09)	(3.40)
$\beta_{INDUSTRY}$	.199***	.184***	.212***
	(8.20)	(6.35)	(5.58)
Avg R^2	1.02%	0.78%	1.23%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
PANEL C: 10% misclassified locations			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.057***	.062***	.052***
	(4.71)	(3.56)	(3.13)
$\beta_{INDUSTRY}$	.199**	.185***	.212***
	(8.21)	(6.37)	(5.58)
Avg R^2	1.01%	0.76%	1.23%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
PANEL D: 20% misclassified locations			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.037***	.050***	.026*
	(3.34)	(2.97)	(1.75)
$\beta_{INDUSTRY}$	.198***	.185***	.211***
	(8.21)	(6.37)	(5.58)
Avg R^2	0.99%	0.76%	1.21%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
PANEL E: 50% misclassified locations			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	-.003	.008	-.014
	(-0.28)	(0.56)	(-0.76)
$\beta_{INDUSTRY}$	.197***	.186***	.208***
	(8.29)	(6.35)	(5.64)
Avg R^2	0.97%	0.72%	1.20%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276

Table 3.7: Geographic Concentration. This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

conditioning on geographic concentration. $r_{i,c,j,t+1}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). Geographic concentration is defined as in Garcia and Norli (2012), based on the number of states mentioned in the 10K. The first column of every block ("Low") indicates the least geographic concentrated firms. The second column of every block ("High") indicates the most geographic concentrated firms. Columns 3-4 (5-6) report the geographic concentration results for small (large) firms. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West standard errors (5 lags). Monthly data, 1970-2013.

Geographic Concentration						
	Full Sample		Small Firms		Large Firms	
	Low	*High*	*Low*	*High*	*Low*	*High*
β_{CITY}	.056***	.083***	.052*	.102***	.059**	.076***
	(3.57)	(3.93)	(1.92)	(3.07)	(2.61)	(3.72)
$\beta_{INDUSTRY}$	.161***	.198***	.190***	.233***	.132***	.143***
	(7.40)	(7.66)	(6.30)	(7.14)	(5.66)	(5.28)
Avg R^2	1.38%	1.20%	1.53%	1.32%	2.26%	2.33%
Observations	526,278	528,599	188,966	307,598	337,312	221,001
# time clusters	527	527	527	527	527	527
avg. # of states	9	3	8	3	10	4

Table 3.8: Predictability of individual stock returns by area and industry portfolios (pooled OLS). This table reports the coefficients of the panel predictive regression with fixed effects

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

where $r_{i,c,j,t+1}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). The predictors are the 1-month (Panel A), and cumulative 3-months (Panel B), 6-months (Panel C) and 12-months (Panel D) lagged city and industry portfolio returns. Column 1: 1970-2013 (full sample). Column 2: 1970-1990. Column 3: 1991-2013. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Standard errors are clustered on both time and firm dimensions following Petersen (2009). Monthly data.

PANEL A: 1-month predictors

	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.063***	.101***	.048**
	(3.66)	(4.62)	(2.11)
$\beta_{INDUSTRY}$	.155**	.175***	.147*
	(2.30)	(5.65)	(1.71)
Adj R^2	10.47%	14.27%	8.96%
Adj R^2 w/o time fixed effect	0.58%	0.87%	0.46%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
# firm clusters	12,998	5,978	10,519

PANEL B: 3-months cumulative predictors

	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.027***	.047***	.019
	(2.86)	(3.54)	(1.42)
$\beta_{INDUSTRY}$	.065*	.076***	.062
	(1.80)	(4.42)	(1.30)
Adj R^2	10.45%	14.38%	8.92%
Adj R^2 w/o time fixed effect	0.10%	0.18%	0.06%
Observations	1,583,404	539,563	1,043,841
# time clusters	525	249	276
# firm clusters	12,930	5,942	10,459

PANEL C: 6-months cumulative predictors

	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.018***	.025***	.015*
	(2.87)	(3.23)	(1.67)
$\beta_{INDUSTRY}$	.026	.032***	.023
	(1.24)	(3.21)	(0.85)
Adj R^2	10.36%	14.31%	8.86%
Adj R^2 w/o time fixed effect	0.00%	0.05%	0.00%
Observations	1,541,127	519,667	1,021,460
# time clusters	522	246	276
# firm clusters	12,821	5,828	10,375

PANEL D: 12-months cumulative predictors

	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.014***	.016***	.011**
	(4.11)	(3.51)	(2.51)
$\beta_{INDUSTRY}$	.014	.022***	.011
	(1.41)	(3.42)	(0.82)
Adj R^2	10.27%	14.46%	8.76%
Adj R^2 w/o time fixed effect	0.00%	0.03%	0.00%
Observations	1,458,783	481,729	977,054
# time clusters	516	240	276
# firm clusters	12,480	5,571	10,132

Table 3.9: Predictability of individual stock returns by area and industry portfolios skipping 1-month (Fama-MacBeth). This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{i,c,j,t+2} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+2}$$

where $r_{i,c,j,t+2}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the equally-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the equally-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). The predictors are the 1-month (column 1), and cumulative 3-months (column 2), 6-months (column 3) and 12-months (column 4) lagged city and industry portfolio returns. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West standard errors (5 lags). Monthly data, 1970-2013.

PANEL: Skipping 1-month				
	1-month return	*3-months return*	*6-months return*	*12-months return*
β_{CITY}	.037***	.025***	.017***	.011***
	(3.06)	(3.15)	(3.49)	(3.43)
$\beta_{INDUSTRY}$	.118***	.062***	.025***	.019***
	(4.59)	(4.68)	(2.75)	(3.52)
Avg R^2	0.97%	0.98%	0.97%	1.06%
Observations	1,598,268	1,569,804	1,527,731	1,445,822
# time clusters	526	524	521	515

Table 3.10: Predictability of individual stock returns by area and industry portfolios (value-weighted portfolios). This table reports the coefficients of the Fama-MacBeth predictive regression

$$r_{i,c,j,t+1} = \alpha + \beta_1 r_{c,\notin j,t} + \beta_2 r_{\notin C,j,t} + \varepsilon_{i,c,j,t+1}$$

where $r_{i,c,j,t+1}$ is the stock return of firm i, in city c, industry j, $r_{c,\notin j,t}$ is the value-weighted lagged return of firms located in city c, outside industry j (city portfolio), and $r_{\notin C,j,t}$ is the value-weighted lagged return of firms in the same industry j, but outside of the 20 cities (industry portfolio). The predictors are the 1-month (Panel A), and cumulative 3-months (Panel B), 6-months (Panel C) and 12-months (Panel D) lagged city and industry portfolio returns. Column 1: 1970-2013 (full sample). Column 2: 1970-1990. Column 3: 1991-2013. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West standard errors (5 lags). Monthly data.

	full sample	*1970-1990*	*1991-2013*
PANEL A: 1-month predictors			
β_{CITY}	.036***	.041***	.032**
	(3.97)	(3.34)	(2.38)
$\beta_{INDUSTRY}$	.145***	.177***	.117***
	(7.55)	(7.24)	(4.07)
Avg R^2	0.76%	0.61%	0.89%
Observations	1,611,973	553,127	1,058,846
# time clusters	527	251	276
PANEL B: 3-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.019***	.018**	.021***
	(4.01)	(2.43)	(3.28)
$\beta_{INDUSTRY}$	.059***	.063***	.055***
	(5.53)	(4.70)	(3.39)
Avg R^2	0.80%	0.73%	0.86%
Observations	1,583,404	539,563	1,043,841
# time clusters	525	249	276
PANEL C: 6-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.016***	.018***	.015***
	(4.04)	(2.80)	(2.91)
$\beta_{INDUSTRY}$	.033***	.034***	.032***
	(4.74)	(3.77)	(3.07)
Avg R^2	0.83%	0.70%	0.94%
Observations	1,541,127	519,667	1,021,460
# time clusters	522	246	276
PANEL D: 12-months cumulative predictors			
	full sample	*1970-1990*	*1991-2013*
β_{CITY}	.006**	.011***	.002
	(2.28)	(2.63)	(0.51)
$\beta_{INDUSTRY}$	.025***	.027***	.023***
	(5.45)	(5.85)	(3.08)
Avg R^2	0.93%	0.79%	1.05%
Observations	1,458,783	481,729	977,054
# time clusters	516	240	276

3.6 Conclusions

Analyzing lead-lag effects between related securities provides a useful way to gauge the efficiency of financial markets. Prior research has identified a number of ways to identify "similar" firms including relationships between companies in the same industry, between customers and suppliers, and between focused firms and conglomerates. Such classifications play an important role in the trading strategies of quantitative hedge funds, which exploit lead-lag effects between related stocks, bonds, options, and other derivatives. The underlying rationale is that although similar firms are exposed to common fundamental shocks, there may still exist variation in the rate at which this information is reflected in prices. Naturally, stocks with the highest (lowest) informational efficiency react the quickest (slowest).

This chapter contributes to this literature by identifying geography – using firm headquarters – as a source of fundamental value. We show that under a variety of empirical specifications, regionally-sorted portfolios generate trading profits that are a third to half as large as those using industry sorts. Note that in addition to documenting an apparent pricing inefficiency, these results point to the presence of locally-derived fundamentals that have a meaningful impact on neighboring firms, even those in vastly different lines of business.

Our most important contribution is a non-result. When we analyze lead-lags between industry peers, trading profits decline sharply for highly scrutinized firms – specifically large firms, those with significant trading volume, or with high analyst following.[12] In contrast, we observe no relation between trading profits and the same variables among regionally sorted portfolios. In addition to being relatively unusual among asset pricing anomalies, this finding suggests that a regional trading strategy might be

[12] Likewise, Cohen and Frazzini (2008) find that predictability between firms with economic linkages declines with proxies for informational efficiency.

profitably deployed even with large amounts of capital.

The mechanism we propose is a refinement of limited attention. Our intuition is that when an investor simultaneously monitors two stocks, he/she is more likely to recognize common relevant sources of information, and through simultaneous trading, reduce lead-lag effects. *Within an industry,* it is the largest firms who are more likely to have more analysts in common with other industrial peers, making it difficult to use industry-based information to predict the returns of large firms. On the other hand, *within a geographic region*, the chance of overlap between neighboring firms in different sector is very small, even among (large or heavily traded) firms with substantial analyst followings individually. As we illustrate in a simple model, industry-focused analysts permit regional lead-lags among virtually all firms (which we observe), but industry lead-lags only for those with the least scrutiny (which we also observe).

It should be stressed that our analysis, which suggests that the organization of the analyst community affects the co-movement of securities, takes that organization as given. Of course, the organizational structure of the analyst community is endogenous, and takes into account the synergies associated with analyzing a closely related group of firms. While we cannot conclude that our analysis indicates that analysts should engage in the collection of costly information by location as well as industry, our analysis does indicate that location-based return information, which is virtually free, can be used to supplement the industry information of stock market analysts. Indeed, one implication, in light of our findings, is a potential role for *regionally focused* analysts in the investment community. Given the trend toward urbanization, and the increased importance of spillovers and other city-level dynamics (Moretti (2012)), an institutional shift toward recognizing these factors seem worth exploring.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Parsons, Christopher A.; Sabbatucci, Riccardo; Titman, Sheridan. The

dissertation author was the primary investigator and author of this paper.

3.7 Supplementary Results

Table A1: Area momentum trading strategy (deciles). This table reports the performance of a trading strategy that exploits return continuation at the geographic level. Every month, we rank each firm *i* by the equally-weighted lagged return of firms headquartered in the same city, outside its industry. We then construct decile value-weighted portfolios of the sorted firms, and hold them for one month. Portfolios are rebalanced every month. Displayed are mean returns, CAPM α, FF-3 α of each decile portfolio. *Volatility* is the monthly standard deviation of the portfolio returns, *Mkt share* is the proportional market share of the individual portfolios, *Size* is the natural logarithm of the market value of the portfolios (in thousands), *B/M* is the book-to-market ratio of the portfolios. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West (5 lags) standard errors. Monthly data, 1971-2013.

Momentum Trading Strategy - Deciles										
	returns						**portfolio characteristics**			
	Mean (%)	*CAPM α*	*t-stat*	*FF-3 α*	*t-stat*	*Sharpe Ratio*	*Volatility (%)*	*Mkt share (%)*	*Size*	*B/M*
lowest city return	0.721	-0.253	-2.554	-0.276	-2.809	0.199	5.274	0.087	15.109	0.613
	0.788	-0.206	-1.799	-0.188	-1.709	0.233	5.508	0.094	15.282	0.597
	0.854	-0.097	-0.954	-0.116	-1.029	0.294	5.130	0.102	15.453	0.581
	0.937	-0.024	-0.235	0.046	0.475	0.336	5.348	0.109	15.616	0.568
	0.980	0.050	0.495	0.070	0.639	0.388	5.009	0.110	15.636	0.565
	1.096	0.167	1.548	0.182	1.615	0.468	5.015	0.108	15.586	0.575
	1.008	0.071	0.776	0.042	0.446	0.408	5.003	0.105	15.543	0.573
	0.930	-0.041	-0.394	-0.040	-0.367	0.331	5.356	0.100	15.445	0.579
	1.104	0.138	1.398	0.084	0.759	0.451	5.264	0.094	15.273	0.592
highest city return	1.132	0.190	2.021	0.191	2.136	0.487	5.071	0.086	15.042	0.602
10-1 spread	0.410 [2.88]	0.444	[3.167]	0.467	[3.301]		3.198			

Table A2: Area momentum trading strategy with the 20% largest firms. This table reports the performance of a trading strategy that exploits return continuation at the geographic level only using the top 20% firms by market capitalization. Every month, we rank each firm *i* by its market capitalization and keep the top 20%. We then re-rank those firms by the equally-weighted lagged return of firms headquartered in the same city, outside its industry. We then construct quintile value-weighted portfolios of the sorted firms, and hold them for one month. Portfolios are rebalanced every month. Displayed are mean returns, CAPM α, FF-3 α of each quintile portfolio. *Volatility* is the monthly standard deviation of the portfolio returns, *Mkt share* is the proportional market share of the individual portfolios, *Size* is the natural logarithm of the market value of the portfolios (in thousands), *B/M* is the book-to-market ratio of the portfolios. ***, ** and * denote significance at the 1%, 5% and 10% level, respectively. Newey-West (5 lags) standard errors. Monthly data, 1971-2013.

Momentum Trading Strategy - Only 20% Largest Firms

	returns						**portfolio characteristics**			
	Mean (%)	*CAPM α*	*t-stat*	*FF-3 α*	*t-stat*	*Sharpe Ratio*	*Volatility (%)*	*Mkt share (%)*	*Size*	*B/M*
lowest city return	0.778	-0.198	-2.214	-0.191	-2.210	0.238	5.225	0.169	15.679	0.591
	0.829	-0.114	-1.313	-0.032	-0.396	0.289	4.926	0.213	16.054	0.564
	0.972	0.055	0.667	0.102	1.219	0.408	4.701	0.221	16.131	0.554
	0.902	-0.047	-0.525	-0.001	-0.012	0.336	4.982	0.210	16.043	0.565
highest city return	1.208	0.271	3.187	0.302	3.193	0.557	4.909	0.187	15.770	0.583
5-1 spread	0.429 [3.07]	0.468	[3.366]	0.493	[3.438]		2.963			

3.8 Bibliography

Atchison, Michael D., Kirt C. Butler, and Richard R. Simonds, 1987, Nonsynchronous security trading and market index autocorrelation, *The Journal of Finance* 42(1), 111–118.

Badrinath, S.G., Jayant R. Kale, and Thomas H. Noe, 1995, Of shepherds, sheep, and the cross-autocorrelations in equity returns, *The Review of Financial Studies* 8(2), 401–430.

Becker, Bo, Zoran Ivkovic, and Scott Weisbenner, 2011, Local dividend clienteles, *The Journal of Finance* 66(2), 655–683.

Bernile, Gennaro, Alok Kumar, and Johan Sulaeman, 2015, Home away from home: Geography of information and local investors, *The Review of Financial Studies* 28(7), 2009–2049.

Brennan, Michael J., Narasimhan Jegadeesh, and Bhaskaran Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *The Review of Financial Studies* 6(4), 799–824.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52(1), 57–82.

Cheng, Yingmei, Mark H. Liu, and Jun Qian, 2006, Buy-side analysts, sell-side analysts, and investment decisions of money managers, *Journal of Financial and Quantitative Analysis* 41(1), 51–83.

Chordia, Tarun, and Bhaskaran Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns, *The Journal of Finance* 55(2), 913–935.

Clement, Michael B., 1999, Analyst forecast accuracy: do ability, resources, and portfolio complexity matter?, *Journal of Accounting and Economics* 27, 285–303.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *The Journal of Finance* 63(4), 1977–2011.

Cohen, Lauren, and Dong Lou, 2012, Complicated firms, *Journal of Financial Economics* 104, 383–400.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *The Journal of Finance* 52(1), 1–33.

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33(1), 3–56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81(3), 607–636.

Garcia, Diego, and Oyvind Norli, 2012, Geographic dispersion and stock returns, *Journal of Financial Economics* 106, 547–565.

Groysberg, Boris, Paul Healy, George Serafeim, and Devin Shanthikumar, 2013, The stock selection and performance of buy-side analysts, *Management Science* 59(5), 1062–1075.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2008, The only game in town: Stock-price consequences of local bias, *Journal of Financial Economics* 90, 20–37.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *The Journal of Finance* 55(1), 265–295.

Hou, Kewei, 2007, Industry information diffusion and the lead-lag effect in stock returns, *The Review of Financial Studies* 20(4), 1113–1138.

Jegadeesh, Narasimhan, and Sheridan Titman, 1995, Overreaction, delayed reaction, and contrarian profits, *The Review of Financial Studies* 8(4), 973–993.

John, Kose, Anzhela Knyazeva, and Diana Knyazeva, 2011, Does geography matter? firm location and corporate payout policy, *Journal of Financial Economics* 101, 533–551.

Kadan, Ohad, Leonardo Madureira, Rong Wang, and Tzachi Zach, 2012, Analysts' industry expertise, *Journal of Accounting and Economics* 54(2-3), 95–120.

Korniotis, George M., and Alok Kumar, 2013, State-level business cycles and local return predictability, *The Journal of Finance* 68(3), 1037–1096.

Kumar, Alok, Jeremy K. Page, and Oliver G. Spalt, 2013, Investor sentiment and return comovements: Evidence from stock splits and headquarters changes, *Review of Finance* 17, 921–953.

Ljungqvist, Alexander, Christopher Malloy, and Felicia Marston, 2009, Rewriting history, *The Journal of Finance* 64(4), 1935–1960.

Lo, Andrew W., and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, *The Review of Financial Studies* 3(2), 175–205.

McGahan, Anita M., and Michael E. Porter, 1997, How much does industry matter, really?, *Strategic Management Journal* 18(S1), 15–30.

McQueen, Grant, Michael Pinegar, and Steven Thorley, 1996, Delayed reaction to good news and the cross-autocorrelation of portfolio returns, *The Journal of Finance* 51(3), 889–919.

Mech, Timothy S., 1993, Portfolio return autocorrelation, *Journal of Financial Economics* 34(3), 307–344.

Moretti, Enrico, 2012, *The New Geography of Jobs* (Houghton Mifflin Harcourt).

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do industries explain momentum?, *The Journal of Finance* 54(4), 1249–1290.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *The Review of Financial Studies* 22(1), 435–480.

Pirinsky, Christo, and Qinghai Wang, 2006, Does corporate headquarters location matter for stock returns?, *The Journal of Finance* 61(4), 1991–2015.

Rumelt, Richard P., 1991, How much does industry matter?, *Strategic Management Journal* 12(3), 167–185.

Schmalensee, Richard, 1985, Do markets differ much?, *The American Economic Review* 75(3), 341–351.

Stickel, Scott E., 1992, Reputation and performance among security analysts, *The Journal of Finance* 47(5), 1811–1836.

Tuzel, Selale, and Miao Ben Zhang, 2015, Local risk, local factors and asset prices, *Working Paper* .