Essays in Financial Economics

Citation
Zhang, Fan. 2014. Essays in Financial Economics. Doctoral dissertation, Harvard University.

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:12274634

Terms of Use

Share Your Story

Accessibility

Essays in Financial Economics

A dissertation presented

by

Fan Zhang

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

September 2013

Dissertation Advisor: Professor John Y. Campbell   Fan Zhang

Essays in Financial Economics

Abstract

This dissertation presents three essays. The first essay finds that the

household risky ratio, the ratio of high risk assets over low risk assets directly

owned by households, is a strong negative predictor of the equity premium on the

US stock market. The predictability is robust to definition of the asset classes,

first versus second half of sample, and the finite-sample bias of Stambaugh (1999).

The predictability is stronger than, and not subsumed by popular predictors like

price-earnings ratios, yield spread, equity share of issues, or consumption-wealth

ratios. The main predictive power is decomposed into three similar parts: 1) the

household tilt of risky assets, which is novel and generally orthogonal to known

predictors; 2) a valuation ratio component; and 3) an issuance component of high

risk versus low risk assets.

The second essay uses a regression discontinuity design on Dodd-Frank's

say-on-pay vote to identify the impact of the vote on CEO compensation.

Crossing the vote discontinuity from 51% to 49% drops the level of CEO pay by

59.8% gross for the next year. The effects are robust, statistically significant, and

economically significant. Despite the sharp drop in pay, crossing the vote

discontinuity has no statistically distinguishable impact on firm market value, earnings, dividends, and capital structure.

The final essay, which is coauthored with David C. Yang, begins by noting the issuance of options by market makers induces hedging feedback demand: when the underlying price goes up, market makers are forced to buy, a well-known phenomenon caused by options gamma. This essay empirically measures hedging feedback demand through the gamma of options above the baseline. This residual gamma causes destabilization in the form of additional momentum. The effect occurs quickly (within four days) and does not revert away rapidly (within ten days). The effect is robust to time trends and liquidity effects. The essay further uses three novel instruments for residual gamma. Residual gamma increases 1) near options expiration, 2) when the underlying price happens to be similar to past price levels, and 3) when the price of the underlying equity happens to be near a round number.

Contents

List of Figures

List of Tables

Acknowledgements

In addition to being an academic work, this dissertation acts as the capstone to five years of graduate school. In that time period, I have received support from advisors, friends, and family.

The research presented here has received invaluable help from conversations with members of my dissertation committee. John Campbell, the committee chair, was instrumental in inspiring me to study the field, encouraging me to pursue the Ph.D., and devoting countless hours to review my work carefully. David Laibson mentored me both personally and professionally at Harvard College, in graduate school, and beyond, always providing me pertinent, incisive, and actionable advice. Malcolm Baker's work was the immediate intellectual inspiration for my research, the proverbial giant on whose shoulder much of this dissertation stands.

I also would like to give special thank Robin Greenwood for deep conversations that helped shape each chapter of my dissertation. Chapter 2 received helpful suggestions from Shawn Cole and Victoria Ivashina, and Chapter 3 received helpful suggestions from Andrei Shleifer and Jeremy Stein. I also thank Alexander Chernyakov, Joshua Gottleib, Timothy McQuade, and Carolin Pflueger for helpful comments thorough seminars.

I am grateful for having an amazing group of friends. David Yang and Charles Nathanson had special impact on my graduate school experience, acting

simultaneously as friends, roommates, and classmates. I am grateful to them for making the experience so enjoyably and intellectually interesting. I also thank Allen Cheng for being a great friend and collaborator. He had a tremendous impact on my graduate school experience and beyond.

Above all, I thank my family for their support. My parents, Lihua Zhou and Jiancai Zhang, have been so unconditionally supportive of me and my choices. I thank them for believing in me, and their willingness to do anything within their power to help me when needed. I also thank my grandparents for the important role they have played in my life, especially Qian Huiying for her dedication and hard work.

Inevitably, these acknowledgements fall short. Five years is a substantial time and the people who have helped and influenced me are numerous. Beyond that, lapses of memory cause the omission of many who have played important roles. Even those mentioned are not done justice due to limitations of both space and power of articulation.

I should finally express my gratitude towards those who provided me the resources for graduate school: Harvard University and the National Science Foundation, and thereby the public at large. I am incredibly appreciative of our fellow citizens, who take from their own desires, through gifts and taxes, to advance our understanding of the world.

Chapter 1

1   Household Risky Ratio and Expected Stock Returns

1.1   Introduction

Over the last few decades, a key asset pricing finding is that the equity premium varies over time. Predictors of the time varying equity premium include price-earnings ratios (Fama French 1988) (Campbell Shiller 1988), yield spreads (Fama Schwert 1977, Keim Stambaugh 1986, Campbell 1987), consumption-wealth ratios (cay) (Lettau Ludvigson 2001), or equity share of issuances (Baker Wurgler 2000).

The equity premium can be seen as a type of inverse price on the equity market (Gordon 1962). In understanding time variation of prices, it seems natural

to look at time-variation of quantities. As the difference between expected returns on high risk versus low risk assets, the equity premium motivates the inspection of the quantities associated with high risk versus low risk assets.

Baker and Wurgler (2000) do exactly this, looking at quantities of equity versus debt issued by the corporate sector as a predictor of future equity market returns and find positive results. In particular, the paper finds that when equity prices are high, and future expected returns are low, corporations rationally respond by issuing a higher ratio of equities versus bonds. The results of Baker and Wurgler (2000) suggest that the corporate sector absorbs and optimizes to the demand shocks.

A natural question to ask is, from where do the demand shocks arise? One sector to examine are households, which might be a source of demand shocks due to either behavioral theories like extrapolative beliefs (Greenwood Shleifer 2013) or to consumption-CAPM theories like Campbell Cochrane (1999). Therefore, I then look at the household sector's direct holdings of high risk versus low risk assets.

I draw my data from the Federal Reserve Flow of Funds, which enumerates all the instruments held by the household sector. I find that the four major holdings which compose the vast majority of household direct holdings are equities, risky mutual funds, credits, and deposits. The first two sum up to high risk assets and the bottom sum up to low risk assets.
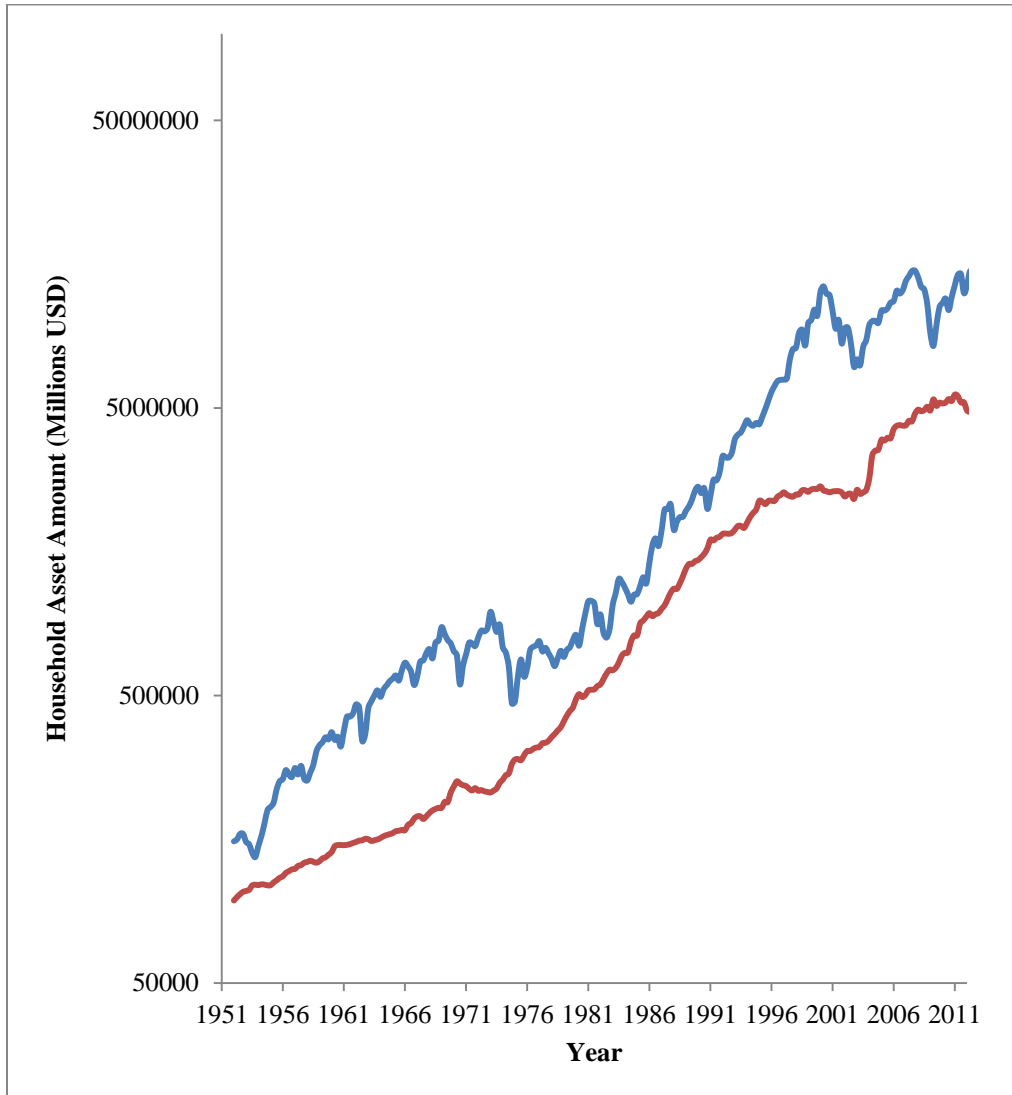
Figure 1.1: Household High Risk and Low Risk Assets, 1951-2012.

High risky (blue) and low risk (red) household holdings are from the Federal Reserve Flow of Funds data. High risk assets are defined as equity held directly or risky mutual funds. Low risk assets include credit instruments and deposits held directly. All dollars figures are nominal.
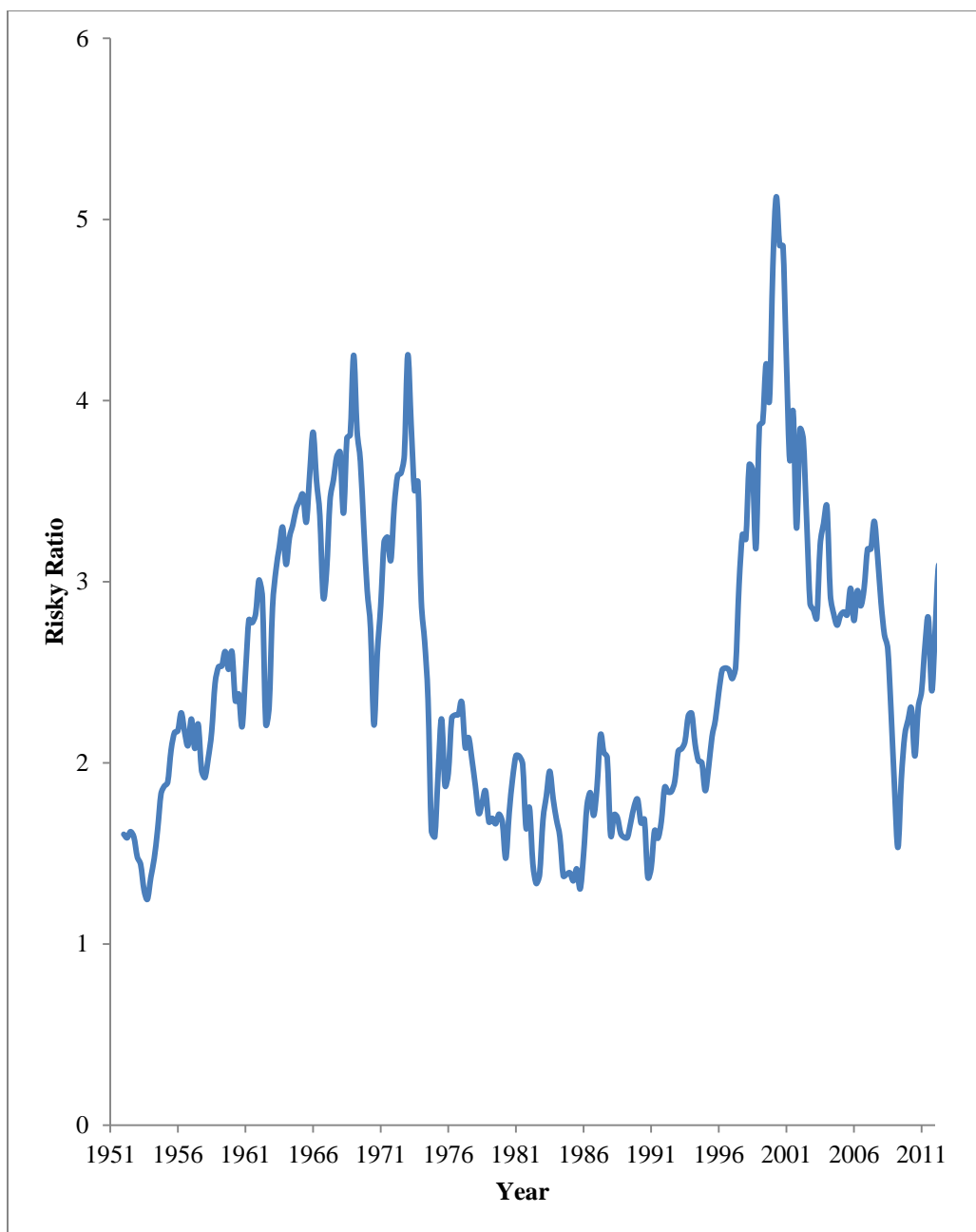
Figure 1.2: The household risky ratio 1951-2012.

The household risky ratio, defined as the value of high risk household assets divided by low risk household assets from the Federal Reserve Flow of Funds data. High risk assets are defined as equity held directly or risky mutual funds. Low risk assets include credit instruments and deposits held directly.

4

Table 1.1: Multivariate OLS Regressions for Predicting Excess Returns.

OLS regressions of one quarter ahead excess returns on the value-weighted CRSP on one period lagged multiple predictors:

$$R_{t,t+1} = \alpha + \beta_1 Risky_{t-1} + \beta_2 Risky_{t-2} + \beta_3 S_{t-1} + \beta_4 CAPE_{t-1} + \beta_5 Term_{t-1} + \beta_6 CAY_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return of the value-weighted CRSP for one quarter forward; $Risky_{t-1}$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; $Risky_{t-5}$ is the household risky ratio lagged one extra year; S_{t-1} denotes the equity share in new issues as defined in Baker Wurgler (2000); $CAPE_{t-1}$ denotes the ten year cyclically adjusted price to earnings ratio, defined as per Campbell Shiller (1988); $Term_{t-1}$ denotes the yield premium of ten year over one month federal government obligations; CAY_{t-1} is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001); t-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods of lags. Regression (8) normalizes all predictors to have unit variance (termed here by the word normed), while all other regressions use non-normalized predictors. N=226.

	Predictors Not Normalized							Normed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept								
$Risky_{t-1}$	-2.49	-2.86	-2.52	-3.76	-2.63	-2.21	-2.39	-1.94
	[-4.16]	[-2.69]	[-4.33]	[-4.20]	[-3.95]	[-3.52]	[-1.97]	[-1.97]
$Risky_{t-5}$		.32						
		[0.27]						
S_{t-1}			-10.49				-7.52	-.71
			[-2.13]				[-1.16]	[-1.16]
$CAPE_{t-1}$				.18			.02	.12
				[1.75]			[0.11]	[0.11]
$Term_{t-1}$					-.28		-.31	-.38
					[-0.73]		[-0.72]	[-0.72]
CAY_{t-1}						49.42	55.94	.95
						[1.51]	[1.24]	[1.24]
R^2	0.054	0.054	0.062	0.060	0.052	0.060	0.060	0.060

Table 1.2: Robustness to First and Second Half.

OLS regressions of one quarter ahead value-weighted CRSP excess returns on one period lagged household risky ratio.:

$$R_{t,t+1} = \alpha + \beta_1 Risky_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return on the value-weighted CRSP for one quarter forward; $Risky_{t-1}$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds. The regression period is for the first half of the sample, the second half of the sample, and then the entire sample. *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 years of lags. N=241

	First Half	Second Half	Entire Sample
Risky	-2.94	-2.15	-2.49
	[-3.34]	[-2.64]	[-4.16]
R^2	.066	.038	.054

Table 1.3: Construction of the Data.

OLS regressions of one quarter ahead value weighted CRSP excess returns on one period lagged of the household risky ratio, defined in various ways:

$$R_{t,t+1} = \alpha + \beta_1 X_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return on CRSP for one quarter forward; X_{t-1} denotes various household risky ratios calculated by dividing various household high risk assets over various low risk assets, as collected from the Federal Reserve Flow of Funds. The numerator may contain equities or mutual funds, and the denominator may contain credit markets or deposits. An X in the table below denotes the inclusion of each variable in the definition of X_t for that regression. t-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods of lags. N=241

	(1)	(2)	(3)	(4)	(5)
X_{t-1}	-2.49	-1.21	-1.05	-3.22	-2.34
	[-4.16]	[-3.40]	[-2.97]	[-3.48]	[-1.58]
Numerator has Equities	X	X		X	X
Numerator has Mutual Fund	X		X	X	X
Denominator has Credit Market	X	X	X		X
Denominator has Deposit	X	X	X	X	
Adj R^2	.054	.041	.034	.035	.008

Table 1.4: Univariate regression under different inference assumptions.

The univariate regression:

$$R_{t,t+1} = \alpha + \beta Risky_t + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return of value-weighted CRSP for one quarter forward; $Risky_t$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds. The OLS method uses standard inferences. The Newey-West method uses heteroskedastic and autocorrelation corrected standard errors with 5 years of lag. The Kendall (1954) correction is a point estimate adjustment. The Kothari-Shanken inferences use the bootstrap methodology to correct for small sample bias as outlined in Stambaugh (1999). Campbell-Yogo (2006) is a hypothesis test whose implementation is outlined in Campbell Yogo (2005). The Lewellen (2004) provides not an unbiased estimate but an upper bound for β and the p-value, and a lower-bound for the t-statistic.

	Newey-West	OLS	Kendall (1954)	Kothari-Shanken	Campbell-Yogo	Lewellen Bound
β on $Risky_t$	-2.27	-2.27	-1.79	-1.89		-.92
95% CI $\underline{\beta}$	-3.43	-3.55	-3.07	-2.89	-4.32	-1.55
95% CI $\bar{\beta}$	-1.11	-.993	-0.51	-1.12	-.03	-.29
t-statistic	-3.85	-3.50	-2.75			-2.88
p value	.0001	.0006	.0065	.0040	.0500	.0062

I then define the *household risky ratio* as the ratio between the two: the household sector's direct holdings of high risk assets over the household sector's direct holdings of low risk assets. I find the household risky ratio is a negative predictor future equity premium between 1951 and 2012, which is the full sample of available data from the Flow of Funds. The negative relationship is obvious in simple binnings of the predictor variable (Figure 1.1), and persists for years beyond the binning date (Figure 1.2).

In a univariate setting, the predictability is strong: the R^2 is 5% quarterly and higher than any single of the following popular predictors: the Campbell Shiller (1988) cyclically adjusted price-earnings ratio (CAPE), the yield spread, the equity share of issues, and the cay (Table 1.1). The predictive power of the household risky ratio is also not subsumed by the above popular predictors: the coefficient remains essentially the same in a multivariate regression alongside the above predictors (Table 1.2).

The predictive power of the household risky ratio is robust to a range of variations. It is robust to construction from Flow of Funds data. Dropping any single component of the four mentioned above does not eliminate the predictive power (Table 1.3). Adding boundary components, components that arguably could be categorized as either inside or outside household discretionary assets, also does not affect the results. It is robust to subsamples: the coefficient in the first half of the sample is almost identical to that of the second half of the sample

(Table 1.4). The effect cannot be explained by Modigilani-Miller effects (Section 5).

Finally it is robust to the small sample correction of Stambaugh (1999). The household risky ratio does experience Stambaugh bias since it has a high autocorrelation (quarterly ρ= .96) and cross correlation with returns (ρ = .86). However, correction for Stambaugh bias via the methods of Kendall (1954) and Kothari Shanken (1997) show that bias to be limited to about 20% of the effect size. This is confirmed by the Lewellen (2004) bounds test and the Campbell Yogo (2006) correction method (Table 1.5). The magnitude of this 20% drop does not change whether the predictor is economically or statistically significant, and is less than other popularly accepted predictors like dividend-price ratios.

I also run overlapping regressions, taking care to address econometric issues of such regressions, and show that the household risky ratio has substantial long-horizon predictability. The R^2 is as high as 22% at the one year level and 40% at the three year level. However, for all other regressions I still set the baseline prediction-horizon to be equal to the sampling-period of one quarter. This follows in the theory of Harri and Brorsen (2009) and avoids the complicated corrections and lack of transparency of overlapping regressions.

Table 1.5: Dickey-Fuller GLS Test of Unit Root in the Household risky ratio.

The table displays the Dickey-Fuller GLS (DF-GLS) test for a unit root in the household risky ratio (Panel A) and changes in household risky ratio (Panel B). The household risky ratio is the ratio of household high risk assets over household low risk assets as reported by the *Federal Reserve Flow of Funds* data. For both the household risky ration and changes in the household risky ratio, the table displays for testing 1 lag and 14 lags (the default maximum tested by the Stata software used) of the DF-GLS test statistic, along with the 1%, 5%, and 10% hypothesis test cutoff values of the test-statistic. N=228.

Number of Lags	DF-GLS Statistic	Rejected?	1% Critical Value	5% Critical Value	10% Critical Value
	Panel A: Household risky ratio				
14	-1.578	No	-3.48	-2.813	-2.534
1	-2.287	No	-3.48	-2.919	-2.630
	Panel B: Changes in Household risky ratio				
14	-3.243 **	Yes	-3.48	-2.811	-2.532
1	-4.841***	Yes	-3.48	-2.922	-2.633

Table 1.6: Comparison of the Household risky ratio, Total risky ratio, and Household Bias as Return Predictors.

OLS regressions of one-quarter forward value-weighted CRSP excess returns on multiple predictors lagged one quarter:

$$R_{t,t+1} = \alpha + \beta_1 \ln(Risky_{t-1}) + \beta_2 \ln(Household\ Tilt_{t-1}) + \beta_3 Valuation\ Ratio\ Part_{t-1}$$
$$+ \beta_4 Issuance\ Part_{t-1} + \beta_5 S_{t-1} + \beta_6 CAPE_{t-1} + \beta_7 Term_{t-1} + \beta_8 CAY_{t-1}$$
$$+ \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess return on CRSP for one quarter forward; $\ln(Risky_{t-1})$ denotes the log household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; $\ln(Household\ tilt_{t-1})$ denotes the log of the household tilt, the quotient between the household risky ratio and *total risky ratio* as calculated by dividing economy-wide high risk assets over economy-wide low risk assets, as collected from the Federal Reserve Flow of Funds; $Valuation\ Ratio\ Part_{t-1}$ is the linear projection of the total risky variable above onto CAPE and $CAPE^2$; $Issuance\ Part_{t-1}$ is the residual from the above projection; S_{t-1} denotes the equity share in new issues as defined in Baker Wurgler (2000); $CAPE_t$ denotes the ten year cyclically adjusted price to earnings ratio, defined as per Campbell-Shiller (1988); $Term_{t-1}$ denotes the yield premium of ten year over one month federal government obligations; CAY_t is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001). *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods lags. N=228.

	Panel A: No Covariates				
	No Covariates				
	(1)	(2)	(3)	(4)	(5)
$\ln(Risky_{t-1})$	-6.47				
	[-4.12]				
$\ln(Household\ Tilt_{t-1})$		-9.26			-14.44
		[-2.31]			[-3.39]
$Risky\ Valuation_{t-1}$			-3.46		-6.17
			[-1.79]		[-3.14]
$Risky\ Issuance_{t-1}$				-7.68	-6.50
				[-2.39]	[-2.25]
S_{t-1}					
$CAPE_{t-1}$					
$Term_{t-1}$					
CAY_{t-1}					
R^2	.0571	.0172	.0113	.0164	.0664

Table 1.6 (Continued)

<div align="center">Panel B: With Covariates</div>

	Covariates				
	(6)	(7)	(8)	(9)	(10)
$\ln(Risky_{t-1})$	-6.44				
	[-2.11]				
$\ln(Household\ Tilt_{t-1})$		-9.49			
		[-2.13]			
$Risky\ Valuation_{t-1}$			-5.94		
			[-2.75]		
$Risky\ Issuance_{t-1}$				-.13	-4.89
				[-.03]	[-1.32]
S_{t-1}	-7.73	-10.45	-13.50	-11.39	-9.70
	[-1.20]	[-1.73]	[-2.15]	[-1.62]	[-1.59]
$CAPE_{t-1}$	.02	-.24		-.20	
	[.15]	[-2.63]		[-2.24]	
$Term_{t-1}$	-.36	-.00	-.32	-.25	
	[-.84]	[.01]	[-.71]	[-.52]	
CAY_{t-1}	53.98	70.01	96.10	94.15	
	[1.26]	[1.67]	[2.39]	[2.20]	
R^2	.0638	.0635	.0568	.0458	.0236

This paper then investigates the source of the predictive power of the household risky ratio. The household risky ratio can be seen as the sum of three parts. First define an auxiliary variable called the *total risky ratio*, the ratio of economy-wide high risk assets over economy-wide low risk assets. Part one is then the *household tilt,* the ratio of the *household risky ratio* over the *total risky ratio*; it measures how much households tilt towards high risk assets above and beyond the rest of the economy. The household tilt variable is seen to be generally orthogonal to known popular predictors (Table 1.6) and hence is novel. Parts two and three explain the total risky ratio. The second part is the *valuation ratio part* of the total risky ratio, and captures the extent to which the total risky ratio can be seen as a price of risky assets, and thus a tracker of historical price changes. For example, if the value of high risk assets doubles overnight, the total risky ratio also doubles. The valuation ratio component captures this doubling. The valuation ratio part is then defined as the projection of the total risky ratio onto the known valuation ratio CAPE.

The final part must mechanically be what remains from the above. In particular, it is the residual from the above projection of the total risky ratio onto the known valuation ratio CAPE. Since asset classes can only grow though capital gains and issuances, and capital gains are captured by CAPE above, the residual can be thought of theoretically as the component of the household risky ratio that arises from issuances. In fact, empirically the final part indeed strongly

14

interacts with Baker-Wurgler issuances (Table 1.6). For these reasons, I term it the *issuance part*.

These empirical results can be related to the theory of demand systems for assets (Brainard and Tobin 1968). Households can be seen as exogenously demanding more risky assets. This causes the quantity of household high risk assets to increase, while at the same time increasing the price of high risk assets and decreasing forward returns. This justifies the observed negative correlation between the household risky ratio and forward returns. The demand system model can also be seen in recent work by Baker Wurgler (2000), Baker Greenwood Wurgler (2003), and Greenwood Vayanos (2010).

These empirical results have welfare implications. Since the household sector buys equities when future returns are lower than average, the return experienced by households will be less than a constant-ratio standard. I calculate that households receive a Sharpe ratio of .267 from holding risky assets instead of .311 under a constant-ratio standard. Both returns are far below the .533 Sharpe ratio attainable by perfectly conditioning on predictive variables.

This seems to be an opportune time to note this dissertation has been adapted by concatenating together the master versions of three papers. Minimal formatting and numbering changes have been made to comply with dissertation format requirements. However, textual changes may not have been made. Therefore, discrepancies may arise. For example, what was called a paper or article is officially called in this version an essay or chapter. Similarly, in the

paragraph below, Section 2 refers to the subpart labeled by 1.2. Subsection C of Section 2 may likewise be 1.2.3. This version sacrifices some clarity for format compliance, and therefore will not be as clear as the master version optimized solely for clarity, which can be requested from the author. Nevertheless, in terms of content this version is nearly identical to the corresponding master copy and fully usable for research.

Section 2 presents the data, definitions, and various constructions of the household risky ratio variable. Section 3 presents the predictive power of the variable for the equity premium. Section 4 decomposes the predictive power of the household risky ratio into components. Section 5 discusses some theory and implications of the results. Section 6 discusses extensions. Section 7 concludes.

1.2   Data

The main data series used to construct the household risky ratio and related variables is the *Federal Reserve Flow of Funds Financial Accounts of the United States*, or simply the Flow of Funds. This data series reports quarterly data starting the fourth quarter of 1951, and continues to present day with data released around 60 days after the end of the quarter. All data is collected from the Flow of Funds using the data download program from the quarterly series (Q series) measuring non-seasonally-adjusted levels (FL series).

The Flow of Funds is organized mainly along two dimensions. The first dimension is the sector, defined as a partitioning of players in the economy.

16

Sectors include households and nonprofit organizations, nonfinancial businesses, state and local governments, and soforth (see Figure 1.3). The second dimension is instrument type, which can be seen as a class of assets or liabilities: home mortgages (liability), total loans (liability), treasury securities (asset), and security credit (asset) are just a few examples. For both the sector and instrument dimension, there are varying levels of aggregation.

To construct the household risky ratio, the numerator and denominator was first separately constructed from data within the household and nonprofit organizations sector. The numerator consists of the sum of equities directly held by households (FL153064105.Q) and the risky mutual funds directly held by households (FL153064205.Q). This corresponds roughly to the total amount of high risk assets that households hold directly with discretion, defined as assets that households can readily and liquidly trade to reflect their preferences and beliefs. The approximation is even better in variation-space instead of levels-space. The denominator consists of household directly held credit market instruments (FL154000025.Q), and deposits which includes low risk money market mutual funds (FL154004005.Q). This corresponds roughly to the amount of low risk assets that households hold directly with discretion.

The household risky ratio then can be seen as approximately the ratio of high risk to low risk assets directly and discretionarily held by households. Two items are to be discussed: what exactly does approximate mean, and what's the significance of assets that are directly and discretionarily held?

Flow of Funds Matrix for 2012

(Billions of dollars; All Sectors -- Assets and Liabilities)

		Households and Nonprofit Organizations		Nonfinancial Business		State and Local Governments		Federal Government	
		A	L	A	L	A	L	A	L
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	Total financial assets	54390.5	--	19560.3	--	2095.4	--	1416.8	--
2	Total liabilities and equity	--	13453.1	--	44080.0	--	3732.7	--	13468.7
3	Total liabilities	--	13453.1	--	19820.5	--	3732.7	--	13468.7

Domestic Nonfinancial Sectors		Domestic Financial Sectors		Rest of the World		All Sectors		Instrument Discrepancy
A	L	A	L	A	L	A	L	
(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
77462.9	--	69869.8	--	19384.4	--	166717.2	--	-7418.8
--	74734.5	--	70743.7	--	13820.2	--	159298.4	--
--	50475.0	--	65763.9	--	9074.5	--	125313.4	--

Figure 1.3: Flow of Funds Excerpt: Size of Sectors.

The above figure is taken directly from the first three rows of table Z.1 of the March 7, 2013 publication of the *Federal Reserve Flow of Funds*. The rows show the financial assets and liabilities of all sectors defined by the *Flow of Funds*. Within each sector "A" specifies assets and "L" specifies liabilities.

The approximation is actually quite good. Within the class of household discretionary and direct low risk assets for example, credits and deposits are by far the largest instruments. The next largest instruments like security credit or miscellaneous assets compose only less than 5% of credits and deposits, and similarly for risky assets. This is improved even more once one considers that approximation in levels is not as important as approximation in variation. For high risk assets even though some local government retirement plans are not measured, the variation of the unmeasured portion likely moves with the measured portion – which is to say with the same variation as the equities market. This robustness and point that the variation is what matters is demonstrated in the later discussion of Table 1.4 which shows that predictability barely drops even as large parts of the above four components are dropped. Overall then, the approximation of the numerator and denominator to household discretionary and direct holdings is quite good.

The second issue to discuss is why the paper limits to discretionary and direct holdings of the household. In some respects, this is a fundamental issue. After all, the entire economy is indirectly controlled by the household sector, so even the definition of the household sector by the Flow of Funds assumes some boundary. This paper is simply assuming a narrower boundary.

The major theoretical reason motivating the limit to discretionary and direct holdings is that only assets that are readily and liquidly traded by households can be seen to strongly and immediately reflect their preferences and

beliefs. To the extent that the household risky ratio has predictive power through capturing preferences and beliefs, this property is essentially. For example, take the largest asset classes not included: real estate, private businesses, and life insurance reserves. The first is moderately illiquid with transaction costs of 6% plus time, preparation, and other opportunity costs. It is also a bundled good that reflects preferences for internal space, external location, commute distance, and school district. Thus, sales and purchases of housing might be seen as a much noisier measure of household risk preferences then, say, holding the S&P index. This is similarly true for private businesses which suffer from heavy adverse selection issues in sales. Life insurance reserves likewise are often managed by a portfolio manager and so less effectively measure household preferences.

The above two points provide theoretical motivation for the construction of the household risky ratio variable. Even without the above two reasons, the household risky ratio can be taken at face value as an empirical construct with predictive power. Concerns of data snooping could be allayed by the fact that the 4.16 t-statistic in univariate regressions corresponds to a Bonferroni correction of testing one thousand independent variables, whereas the Fed Flow of funds has much fewer, especially independent variables. Further assurance can be realized in the robustness of the ratio to construction (Table 1.4).

The other primitive variable constructed from the Flow of Funds is the total risky ratio, which is the exact analogue series of the household risky ratio but for all sectors combined instead of just the household sector.

Compared to other broad economic series, the Federal Reserve Flow of Funds is minimally affected by look-ahead bias. Generally the series is release quickly after a quarter's end: around 60 days on average, and almost never exceeding a quarter. Thus, the one quarter lag used in the baseline is more than sufficient to account for this. The Flow of Funds is revised from time to time. However, revisions are generally limited to one to two years back, and this paper did not find any revisions of the four broad series used to define the household risky ratio.

Figure 1.2 presents the amount of household high risk versus low risk assets on a log scale from 1951-2012. The high risk series, as expected, is more variable over time. The high and low risk series seem to be trending up linearly, perhaps each following a first order autoregressive process with unit root and drift. The two series looks co-integrated, which is justifiable by theory: in the long run they might be expected to be growing at the same rate as the entire economy.

Figure 1.2 shows the household risky ratio, the first series above divided by the second series. Peaks and trends in this series are apparent. The ratio slowly but steadily climbs from 1951 to about 1969, which might signify a trend where equities began to be seen as less a speculative asset from the aftermath of the great depression, into an asset that everyone could own. As will be discussed later, the rise in price may have caused an increase in quantity if households have extrapolative beliefs, and the two effects reinforced each other.

Households then experienced two large price drops in 1970 and 1974,

causing them to be more cautious about stocks for many decades ahead, up until

the mid-90s. It is exactly during this time that equities saw their highest returns.

In the late 90s, as the Internet bubble got underway, households against started

shifting into high risk assets, again to see significant wealth loss as the bubble

burst.

Other data sources have been compiled from sources as standard as

possible. The ten-year cyclically adjusted price-earnings ratio (CAPE) is

calculated as per Campbell Shiller (1988), and the data is collected from Robert

Shiller's website. CAY is defined and supplied by Lettau Ludvigson (2001).

Equity shares of issuances is from Baker Wurgler (2000). Total stock market

returns is the value-weighted series from CRSP. The long rate is the GS10 rate

collected from Robert Shiller's website. The short rate is the 1-month treasury

bill rate provided by Ibbotson and Associates, Inc. and provided on Kenneth

French's website. The equity premium is calculated as the difference between the

stock market returns minus the short rate expressed in percentage points. The

yield spread is calculated as the difference between long rates and short rates.

1.3   Predictive Power

This section outlines the predictive power of the household risky ratio. It begins

with informal graphical analyses and then formalizes these analyses through

regressions. Finally, this section checks robustness through multivariate

regressions and additional tests.

1.3.1 Baseline Predictive Power

Figure 1.4 shows 1-year and 3-year value-weighted CRSP forward equity

premia sorted by the lowest to highest quintile of the household risky ratio

variable. The drop in equity premia from the highest to lowest quintile is

generally monotonic, demonstrating a somewhat continuous relationship between

the two variables. The magnitude of the effect is larger for 3-years, which is not

surprising given the persistence of the household risky ratio predictor variable.

Figure 1.5 shows forward annual risk premia with respect to the lowest,

middle, and highest tercile of household risky ratio. Note that there is no

relationship between the terciles and returns before the tercile formation time.

For the first 9 years, the lowest tercile has significantly greater returns than then

highest tercile. The years of predictability seems to accumulate most at the

beginning but persist for many years. One clear explanation for this is simply the

fact that the predictor variable is quite persistent, so predictability a few years out

can simply reflect this persistence. In this way Figure 1.5 can also be seen as a

cross correlogram between returns and the household risky ratio.

At a glance, it seems clear that there is some negative relationship between

the household risky ratio and forward returns. The relationship seems to persist

for quite a number of years afterwards.
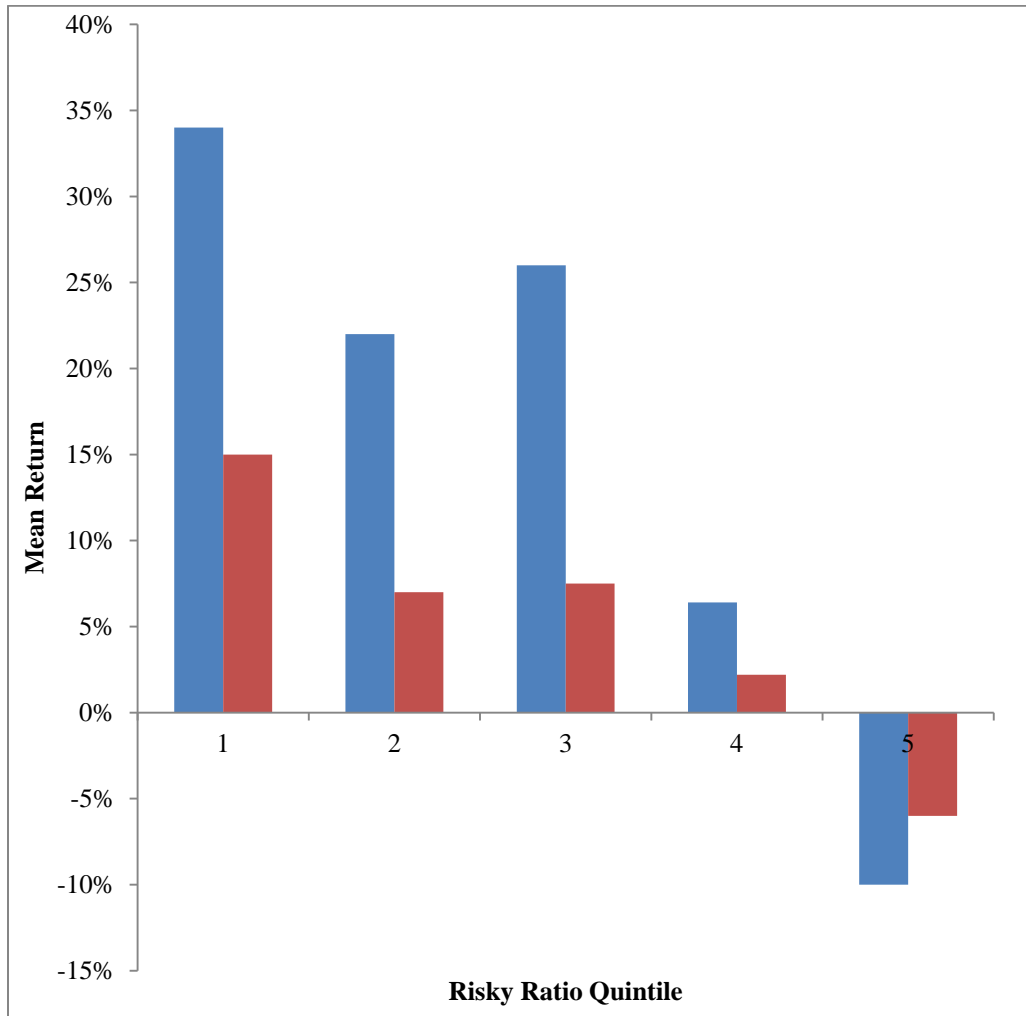
Figure 1.4: Mean 1 and 3 year forward returns by household risky ratio, 1951-2012.

Mean 1-year forward (red) and 3-year forward (blue) value-weighted CRSP excess return by quintile of household risky ratio. The risky-ratio variable (228 observations) were ranked and then binned into quintiles. Quintile 1 below contains the lowest fifth of the value of household risky ratios, and quintile 5 contains the highest fifth. Then the excess return for the next 1 and 3 years on the value-weighted CRSP is calculated and plotted.
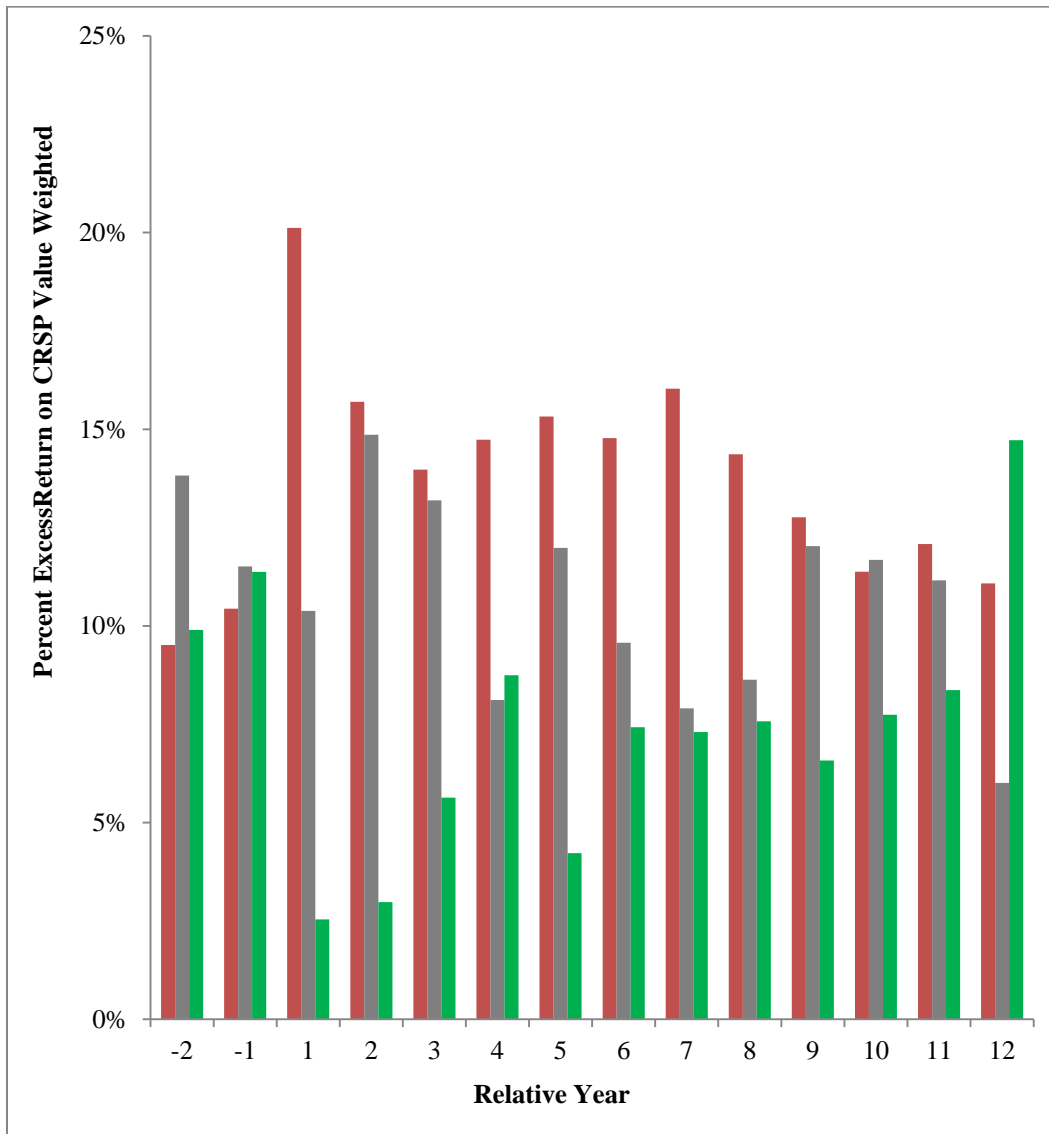
Figure 1.5: Mean past and future annual equity returns from t-2 to t+12 by household risky ratio.

Each quarter is binned by terciles of household risky ratio – low (red), medium (grey), and high (green). Annual excess valued-weighted CRSP returns from 2-years past (t-2) to 12 years in the future (t+12) for each bin are plotted in the bar chart below.

1.3.2 Univariate Regressions

I next show the power of the household risky ratio predictor in a univariate regression setting. I predict one-quarter forward returns using a variety of predictors. All predictors are lagged one quarter, leaving a one quarter minimum gap between the measurement time of the predictors and the start of the return period predicted. As standard in the literature, this gap ensures that the prediction is completely in the future period, and that the conditioning data is available for practitioners. This is especially important because some predictors like the household risky ratio and the Shiller CAPE contain equity prices; common noise in equity price measurement would enter on both sides of the regression and swamp the magnitude of return predictability.

The general regression framework for return predictions following the discussion above is then

$$R_{t,t+1} = \alpha + \beta X_{t-1} + \epsilon_t$$

Table 1.7 shows the results of this regression against a set of regressors. First, notice that the household risky ratio without additional lags is highly significant with a t-statistic of 4.16. As a univariate regression, the significance of this statistic under these relatively transparent settings is very high. Even more transparently, consider that the adjusted R^2 at .054 and the correlation is 0.232. These figures are all substantially higher than almost all popular predictors in the literature.

Table 1.7: Univariate OLS regressions for predicting five-year-ahead market returns.

One-quarter ahead valued-weighted CRSP excess returns are regressed on a variety of predictors:

$$R_{t,t+1} = \alpha + \beta X_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the excess returns on the value-weighted CRSP for one quarter ahead and X_{t-1} variously denotes the household risky ratio, defined as household high risk assets divided by household low risk assets; the household risky ratio lagged an extra quarter; the household risky ratio lagged an extra year; equity shares of issues (Baker Wurgler 2000); Campbell Shiller (1988) 10-year cyclically adjusted price to earnings ratio (CAPE); the 10-year 1-month government obligation yield spread; and the consumption-wealth proxy CAY (Lettau Ludvigson 2001). *t*-statistics are heteroskedastic and autocorrelation robust (Newey-West) with 5 periods of lag.

Predictor	β	$t(\beta)$	α	$t(a)$	R^2
Household risky ratio	-2.49	[-4.16]	8.05	[5.42]	.054
Household risky ratio (lagged extra quarter)	-2.40	[-4.35]	7.84	[5.79]	.050
Household risky ratio (lagged extra year)	-1.94	[-2.94]	6.72	[4.20]	.032
Baker-Wurgler Equity Share	-7.32	[-1.34]	3.02	[2.83]	.003
Campbell-Shiller CAPE	-.13	[-1.68]	4.36	[2.81]	.011
Term Spread	.32	[0.87]	1.08	[0.99]	.000
Lettau-Ludvigson CAY	79.73	[2.51]	1.82	[3.49]	.022

To examine the degree of predictability at different points in time, Table 1.7 includes results for the household risky ratio at one quarter and one year of additional lag. The coefficient noticeably decays from -2.49 to -2.40 to -1.94, but magnitude of the predictor even at an additional year out is only 22% less than originally. The t-statistics and R^2 also remain highly significant but noticeably decay. As will be seen later, this is largely caused by persistence in the household risky ratio, and not the independent ability of different lags of the household risky ratio at predicting future returns.

The next few rows of Table 1.7 examine the univariate regressions of other popular predictors of future equity premium. The paper chose four of the most popular variables in the literature as benchmarks, following Campbell Thompson (2008): the equity share of issues from Baker Wurgler (2000), the Shiller 10-year cyclically adjusted PE (CAPE) from Campbell Shiller (1988), the 10 year minus 1 month yield spread, and the cay proxy for consumption/wealth by Lettau Ludvigson (2001).

The coefficients for all these four other predictors are significantly less than that for the household risky ratio. The Lettau-Ludvigson cay is most significant with a t-statistic of 2.51 and an R^2 of .022. In second is CAPE with a t-statistic of 1.68 which is marginally significant. The lower significance of CAPE in these regressions compared to the literature can mainly be attributed to the shorter data series. For example, one of the most convincing works for the

28

predictive power of CAPE is Campbell Shiller (1988) which uses data back to 1871 versus this paper's 1951, and a specialty of Robert Shiller's research is general marshaling past data to gain higher significance.

The equity-share of issuance then has a t-statistic of around 1.34. There are quite a few differences between this analysis and Baker Wurgler (2000) that may give rise to lower significance. Baker Wurgler (2000) note their effect has most significance and their test logically has most power in equal-weighted CRSP, while this paper uses the value-weighted CRSP equity premium. Baker Wurgler (2000) also uses different prediction periods from this paper, with predicted periods of one-year. The significance does increase if the duration of the issuance is increased from one quarter to four quarters. Finally, the time period matters again: the regressions in this paper are from beginning of the Flow of Fund data (1951) forward whereas Baker Wurgler (2000) used data from 1929. The yield spread in this analysis barely has univariate predictability at all.

Overall the result of this analysis is that the household risky ratio is a powerful predictor of future equity premium. This is not only the case in Table 1.7 which compares four popular variables under this setting, but also by direct examination of t-statistics and R^2 of the household risky ratio predictor, and comparing against analogous values across a wide variety of popular predictors in the setting of their seminal papers.

1.3.3   Default Inferences for Regressions

This subsection discusses and justifies the inference assumptions used by default for the regressions above and the rest of the paper. In particular, it discusses the focus on the one-quarter prediction horizon, as well as the choice to use Newey-West standard errors with five periods of lag.

By default in this entire paper, the prediction period is chosen to be one quarter due to that being the unit of observation in the *Federal Reserve Flow of Funds*. The later section on long-term predictions discusses more in depth the results of different periods of predictions, and econometric implications of having a prediction period that is longer than the data period. As a general preview, the results are more or less the same with quarter returns, one year returns, or three year returns, but the one quarter returns are most natural and least dependent on model corrections.

Also by default in this entire paper, regressions are done with Newey-West with five periods of lags for robustness. The Newey-West procedure encompasses the Eicker Huber White (EHW) heteroskedastic robust error correction (White 1980). The household risky ratio regression is not especially hetereoskedastic, but inferences are often affected 10% or more using the EHW method. Often the EHW standard errors are *less* than OLS due to return variance being concentrated near the center of the predictor.

Newey-West also additionally takes into account time series correlation of error terms assuming a triangle (Bartlett) kernel for time series correlation structure. The household risky ratio predictor has somewhat autocorrelated errors,

and inferences are often affected 40% or more using Newey-West procedures

instead of EHW. Often the Newey-West standard errors are less than EHW due

to negative residual correlation. The negative residual correlation can be verified

by adding lagged returns in the household risky ratio regression. However, the

paper does not put lags into the baseline specification because the negative

correlation is weak. In a sample regression with ten lags, only one coefficient

exceeds 1.5 and none are significant at the 5% Bonferroni-corrected level.

The lag period was chosen using the procedure suggested by Newey

(1993), and set at $\frac{3}{4}T^{\frac{1}{3}} \approx 5$, with T=240. In reality lags between one and ten

periods generate almost the same result everywhere. While the paper does not

show simple EHW standard errors or OLS standard errors, neither of these two

less-robust alternate procedures generate a univariate t-statistic of less than 3.5 for

the household risky ratio. More generally, econometric surveys (Harri Brorsen

2009) have shown that Newey-West corrects many substantially critical errors,

and almost never give estimates substantially worse than other estimators in the

space of heteroskedastic and autocorrelation robust estimators or less, including

EHW and OLS. For both general and situation specific reasons this paper then

universally uses Newey-West with five periods of lag as default.

1.3.4 Multivariate Regressions.

The univariate regressions are useful in examining the prediction power of

each variable, as well as cross comparisons. However, to more fully understand

31

the structure of the household risky ratio, the source of interactions with other

variables, the source of the prediction power, it is imperative to examine

multivariate regressions in Table 1.1. As above, the paper uses a one-quarter

prediction period with a one period lag for the prediction, and Newey-West

standard errors with 5 lags by default.

Table 1.1 begins its multivariate regression by regressing returns on

household risky ratio and one year of additional lag. This regression gives insight

into the time series structure of the predictability of household risky ratio. Do

lags of household risky ratio have independent predictive power for future returns,

as is the case for cay? This might arise if there is period-specific noise in

measurement of the household risky ratio. Do the two lags knock each other out,

as might be expected if the predictability of the household risky ratio is due to low

frequency components? Looking at Figure 1.2 it might be tempting to suspect that

predictability is driven completely by a few periods in which the household risky

ratio was particularly high or low. Also, this regression lets us see whether it's

the level of household risky ratio or its change that matters more, and whether

longer lags might have more predictive power for some reason.

Looking at Table 1.1 regression (2), and comparing against the univariate

regression (1), a few things stand out. First, the coefficient on the household risky

ratio is near identical, going from -2.49 to -2.86. However, the t-statistic drops

significantly from 4.16 to 2.69. Second, the coefficient on the one-year-lagged

household risky ratio is near zero, with a t-statistic of 0.27. That the t-statistics

are lower in regression (2) does suggest that household risky ratio and lagged

household risky ratio knock each other out somewhat. Some portion of the

predictability then can be due to a low frequency component of the household

risky ratio. However, the significance of the leading household risky ratio term

and its being equal to the univariate coefficient suggests that high frequency

changes on the order of less than a year matter, and have the same effect on future

returns as low-frequency changes. The high frequency variation in the household

risky ratio is not simply noise.

The R^2 of regression (1) compared to (2) is the same, in line with the fact

that the marginal t-statistic of the one-year-lagged household risky ratio is near

zero. Adding an extra term does not increase predictability, so there is no

independent prediction power by lags. (This is verified but not shown with other

periods of lags). In a Markov-chain sense, the most recent household risky ratio

could be seen as a state variable or a sufficient statistic for the household risky

ratio process. This also means it is assuredly the level of household risky ratio

that matters, versus changes, in contrast to other predictors like the equity shares

of issuance (Baker Wurgler 2000).

The next regressions run in Table 1.1, regressions (3) through (6), examine

bivariate regressions of the household risky ratio against various other predictors

commonly used to forecast equity premia. The paper looks to see if any

predictors knock the other out – a sign that one predictor is subsumed by the other

better, less noisier, predictor. To the extent both predictors still maintain size and significance, it is a sign that the two predictors have somewhat independent predictability.

Regression (3) looks at the household risky ratio versus the equity share of issuances. These two variables have some relationship because companies issuing equity versus debt should be expected to mechanically increase or decrease the household risky ratio over time. Also, in the next section it will be seen that one component of the household risky ratio can be thought of as issuances. Some confounds might be expected, but instead the comparison of regression (3) against the univarate regressions show that significance and magnitude of both regressors increases slightly.

The household risky ratio coefficient increases from -2.49 to -2.52, significance increasing from -4.16 to -4.33. The equity share of issues increases much more from -7.32 (t-statistic of -1.34) to -10.49 (t-statistic of -2.13). The adjusted R^2 also increases from .054 and .003 in the univariate regression to .062 together. This suggests not only that the two predictors have independent predictive power, but also that the prediction is strengthened when the two are put together. This might occur if companies issue stock for two reasons: one is a mechanical response to the demand side of the economy, captured by the household risky ratio; another could be CFOs knowing to market time beyond demand signals.

Regression (4) looks at the bivariate regression of the household risky ratio against the Campbell-Shiller CAPE. One immediately notes that the coefficient on the household risky ratio actually becomes significantly higher in magnitude, jumping from -2.49 to -3.76. The coefficient on CAPE is also significant but now in the opposite direction: moving from -.13 to .18. This shows that not only is the household risky ratio not subsumed by the popular valuation ratio predictor CAPE, but that it actually may be a better valuation ratio than CAPE. Another interpretation might be that the household risky ratio captures more predictive components than CAPE, causing CAPE to act as a negative control. The decomposition of the household risky ratio in the next section lends some credence to this idea.

Campbell-Shiller (1988a) gives an accounting relationship between PE ratios, earnings growth, and expected returns. Given this opposite response in CAPE along with the household risky ratio, it may be interesting to see whether CAPE along with the household risky ratio can finally predict earnings changes.

Regression (5) runs the household risky ratio against the term spread. Neither variables are affected significantly. The term spread is not particularly significant, having both univariate and bivariate t-statistics of less than one. This seems consistent with a story in which the term spread is not a particularly powerful variable. Regression (6) runs the household risky ratio with cay. Both variables are reduced somewhat in magnitude and significance. The bivariate R^2

is also somewhat less that the summed R^2, showing that two variables are picking up on some common predictability. This is not surprising given Lettau Ludvigson (2001) theory of cay being a consumption wealth valuation ratio.

Regression (7) uses all variables and covariates, except the lagged household risky ratio which was seen before to be near collinear with the household risky ratio itself. Because many of the popular predictors tend to be different variations on each other, the significance and size of predictability is generally less than in the bivariate and univariate regressions. This is true of the household risky ratio as well. But what is striking is that the household risky ratio variable is still quite significant (p<.05). Also, across all the regressions from (1) through (7), the magnitude of the coefficient on the household risky ratio is more or less the same at around -2.5.

This remarkable consistency across the bivariate and multivariate regressions show that the household risky ratio is not subsumed by other popular predictors, or even the space spanned by other popular predictors. Finally regression (8) standardizes each predictor to have unit variance. The coefficients then are identically proportional to the square of the correlation and square root of R^2, and give a sense of the strength of each predictor on the same scale.

The two main takeaways from Table 1.1 is both the empirical robustness of the household risky ratio as a predictor, and the point that there is something novel about the household risky ratio. Further evidence will be given below that

there is particular significance to the household portion of the household risky ratio.

1.3.5   Robustness

One concern in regressions over large timespans is parameter stability. How much of the predictability at persists from the start of the sample to the end of the sample? Alternately, is predictability completely isolated to one portion of the sample, or even worse, a few years? If the household risky ratio process or the returns process has underlying long persistence, then normal inferences could be incorrect. All of the above issues can be heuristically addressed by doing a split regression on the first and second half of the data (Table 1.2). Such a procedure is model free, and while not the most powerful, provides a clear glance at whether the predictive power is stable. The paper finds that the lineup between the first and second half of the sample is remarkable: the coefficients are quite close: -2.94 versus -2.15. This gives evidence that the result isn't driven just by a few years, and the inferences are not spuriously caused by persistence.

There may be concern over what exactly is classified as components of the numerator and denominator of the household risky ratio. In particular, there might be concerns that one component is the main driver of the result, and so the household risky ratio is not robust to definition. As categorized by the Flow of Funds, there are four major categories in the paper's definition of the household risky ratio. The high risk component is composed of equities and mutual funds.

The low risk portion consists of credit market securities and deposits. Table 1.3

uses a jackknife-like methodology to test robustness of the household risky ratio.

The paper runs regressions removing one component at a time, and sees whether

predictability drops. It seems that predictability remains in almost all case. The t-

statistics generally remain significant, and the coefficient varies from -1.21 to -

3.22. The robustness of regressions to removing components shows that the

household risky ratio is robust to definition.

1.3.6   Small Sample Bias of Stambaugh (1999)

Stambaugh (1999) and Nelson and Kim (1993) note that in regressions in

which predictors X_t are autocorrelated, there is a possibility for a small sample

bias. In particular, using notation of Kothari Shanken (1997), suppose returns

r_{t+1} is being predicted with a univariate autocorrelated variable x_t.[1]

$$r_{t+1} = \alpha + \beta x_t + u_{t+1} \ (1)$$

$$x_{t+1} = \gamma + \phi x_t + v_{t+1} \ (2)$$

Correlations between in the error term u_{t+1} and v_{t+1} will cause a small

sample bias. In the language of Lewellen (2004), the source of this bias is the

well-known downward bias of OLS in estimating (2) equation for positive ϕ, in

particular $E[\hat{\phi}_{OLS} - \phi] < 0$. This transmits over to β in through the correlation.

In particular:

---

[1] The theory of Stambaugh bias is naturally presented without a one period gap between the predictors and predicted. This does not pose a problem for us as we simply define $x_t :=$ $Risky_{t-1}$

$$\hat{\beta} = \beta + (x'x)^{-1}x'u$$

$$\hat{\phi} = \phi + (x'x)^{-1}x'v$$

Which translates into, as Stambaugh (1999) notes, defining $\gamma := \frac{Cov(u,v)}{Var(v)}$

$$E[\hat{\beta} - \beta] = \gamma E[\hat{\phi} - \phi] \neq 0$$

Since the household risky ratio does include a significant price component, it is very close to the original canonical example used by Stambaugh (1999) to demonstrate the bias. In particular, the predictor is quite autocorrelated with $\hat{\phi}_{OLS} = .956$. The error terms are of (1) and (2) are also notably correlated at $\widehat{corr}(u,v) = .869$. All correlations are taken at the one-quarter level.

This paper corrects for the error below using a few methodologies: Kendall (1954), Kothari and Shanken (1997), and a more recent approach by Campbell and Yogo (2006). The correction shows that the household risky ratio is indeed subject to small sample bias about 20% of the magnitude of the coefficient. This bias is not enough to affect qualitatively the results in the paper, especially many of the baseline t-statistics are above 3 or 4. The paper follows the literature and does not by default correct for Stambaugh bias in order to maintain regressions that are multivariate, heteroskedastic and autocorrelation consistent, and standardized to well-known methodologies. However, it is important to keep this relative magnitude of 20% in mind when reading other tables, which affects not only the household risky ratio, but other popular price predictors that have a valuation ratio component like CAPE.

The first correction is a point-estimate correction suggested by Kendall (1954), who notes that analytically the bias in ϕ, $E[\hat{\phi} - \phi] \approx \frac{1+T\hat{\phi}}{T-3}$ which holds generally for T>50 as in this case, giving a value of -0.01598. Then, $E[\hat{\beta} - \beta] = \gamma E[\hat{\phi} - \phi] \approx \frac{\widehat{cov}(u,v)}{\widehat{var}(v)}\left(-\frac{1+3\hat{\phi}}{T}\right)$ again for T>50 can be estimated, in this case 30.00 x -.01598 = -.479. Thus the Kendall correction reduces the OLS estimate of β from -2.27 to -1.79, a move of about 21.1%. The Kendall method does not change the standard errors, leading to a t-statistic of 2.75 and p-value of .0065. The reduction in significance comes solely from the move in the point estimate.

Kothari-Shanken (1997) extends the Kendall (1954) method through bootstrap re-estimation of the standard errors. In particular, Kothari-Shanken simulates the data series by taking the Kendall-adjusted values for ϕ, β and drawing error terms u, v. For each bootstrapped series, a beta is estimated. The distribution of bootstrapped betas then simulates the sampling distribution of β and lets us do inferences. The process results in a p-value of 0.004.

The Kothari-Shaken methodology does not generate wider confidence intervals from the OLS ones. In fact, the confidence interval is actually lower than OLS due to the same reason mentioned previously of EHW standard errors being less than OLS: the middle of the predicted variable has more residual variance than the edges. Like the Kendall methodology the loss in t-statistic comes mostly from the point estimate moving closer to zero versus the standard

40

error increasing. The Stambaugh bias again is about 20% of the OLS coefficient magnitude.

Campbell and Yogo (2006) propose a different Stambaugh bias correction methodology, motivated by the fundamentals of statistical hypothesis testing. The idea behind Campbell Yogo, in relation to its use in this paper, is as follows. First, Campbell and Yogo arrive at an intuitively powerful and rigorous Q-test that accounts for Stambaugh bias. The Q-test depends on knowing ϕ above, which in reality must be estimated rendering the original test infeasible. Campbell Yogo cleverly patch this problem by using the Bonferroni procedure to merge the infeasible Q-test with an estimate of ϕ through the DF-GLS estimator.

The rationale for the Q-test is as follows. Suppose the autocorrelation ϕ of the predictor x_t was known beforehand. The Neyman-Pearson lemma proposes a likelihood ratio test (LRT) as the most powerful test for hypothesis of $\beta = \beta_0$ vs $\beta = \beta_1$. The LRT can be conditioned on an ancillary statistic and be considered inside the space of invariant tests, tests that do not change in response to invariant changes to x_t or r_{t+1}. Then this LRT is also uniformly most powerful (UMP) for β inside the above space of tests. Call this UMP test statistic the Q-statistic, which can be estimated easily in an OLS regression setting. The test based on cutoffs of the Q-statistic is the Q-test.

Kendall provides an intuitive first order estimate of the Stambaugh bias and Kothari Shanken (1997) provide intuitive first order inference corrections.

Neither however have foundations in rigorous statistical testing, which is what is provided by Campbell Yogo (2006). Further, the construction procedure of the Campbell Yogo (2006) is designed towards ensuring high power while being rigorous. As such, this paper considers the Campbell-Yogo correction to be the gold standard for comparison.

The test is implemented via Campbell Yogo (2005). The DF-GLS statistic on the household risky ratio with one period if lag is calculated at -2.287 (Table 1.4), corresponding to a 97.5% confidence interval for c of [-21.87, 1.47] from Table 1 of Campbell Yogo (2005). This leads to a 97.5% CI over phi of [.910, 1.006]. This leads to point estimates of $\hat{\beta}(\phi)$ of -3.65 and -.76 respectively. For each of the two above point estimate, this paper takes another 97.5% confidence interval around the estimates to arrive at a 95% Bonferroni estimate for β: [-4.32, -.03]. An approximate 90% CI would then be [-4.06, -.30].

As noted in Campbell, due to the Bonferroni methodology, the actual confidence interval coverage is guaranteed to be 95% or above. Thus, the estimation above is conservative. Even with the most conservative methodology, it is seen that the household risky ratio is still significant. As a baseline, Campbell Yogo (2006) calculate 90% confidence intervals for D/P, E/P, T-bill rate, and Yield spread, and find many of their variables have 90% confidence intervals that cross 0, and almost all are very close to 0 with respect to the length of the confidence interval. It seems likely that if 95% confidence intervals were

42

calculated, all of the other above univariate predictors would cross zero. Keep in mind that much of the Campbell Yogo (2006) data set extends back to 1926 or even 1880. With this as a baseline, the [-4.32,-.03] confidence interval is remarkably strong for the household risky ratio.

Finally, to verify significance, this paper does one last check in the style of Lewellen (2004). It seems reasonable to assume that the household risky ratio is not an explosive process. Looking at Figure 1.2, the household risky ratio looks likely even stationary as household high risk assets and low risk assets ought to be cointegrated with the size of the economy. Although it is possible it has a unit root, it makes little sense for the household risky ratio to be explosive. Assuming $\phi \leq 1$, the first step to re-run is the Campbell Yogo (2006) bounds. No longer is the 97.5% confidence interval for ϕ [.910, 1.006], but rather it is instead [.910,1]. This leads to a 95% confidence interval for β of [-4.06,-0.217].

Following Lewellen (2004) more directly, consider again $E[\hat{\beta} - \beta] = \gamma E[\hat{\phi} - \phi]$. Rearranging this leads to the estimator that eliminates the bias: $\hat{\beta}_{adj} = \hat{\beta}_{OLS} - \gamma(\hat{\phi} - \phi)$. Assuming the largest correction possible (being overly conservative) under the above assumptions sets $\phi = 1$, yielding $\hat{\beta}_{adj} = -.92$. This is significantly less than the OLS estimate or any other estimates really because it is a lower bound in magnitude. The advantage of the Lewellen method is that the standard errors are much reduced, from .65 in OLS to .32 here. This

yields a much smaller relative confidence interval than Campbell Yogo (2006) and gives a larger t-statistic of 2.88.

The Lewellen (2004) method does not provide any unbiased point estimates of β, but as an advantage, it does have more rigorous foundations than Kendall (1954) or Kothari Shanken (1997), and since the process studied here has an autoregressive root close to unity, it is more powerful than Campbell Yogo (2006).

1.3.7 Persistence and Long-Horizon Predictions

Table 1.5 analyzes the persistence properties of the household risky ratio predictor. The analysis of this is similar to that of the persistence of D/P. From a theoretical perspective, the household risky ratio should not have a unit root: as the economy expands it seems that the fraction of high risk assets versus low risk assets ought to vary around some constant. There certainly is no theoretical reason to believe it has a greater-than-unit root: no model seems to justify an explosive process. From an empirical perspective, if it is believed that expected excess returns do not have a unit root, then it makes no sense to predict it with a series with a unit root. Table 1.5 Panel A shows the DF-GLS test on the household risky ratio: while a unit root cannot be ruled out, this may be due to the power of the test with only 240 quarters as input. Table 1.5 Panel B shows the DF-GLS test on changes of household risky ratio: a unit root is clearly rejected. Thus, the household risky ratio is shown to be integrated maximally of order one.

In this paper, the predictive regressions are generally run on one-quarter forward returns, which also is the frequency of the data as provided by the Federal Reserve flow of funds. The literature often considers longer-horizon returns (Baker Wurgler 2000, Campbell Shiller 1988) to increase the R^2 of the regression as well as to demonstrate the predictive properties over the long term. This section runs longer-dated returns using the household risky ratio and discusses the longer-dated results in the context of econometric theory.

A main advantage of long-horizon returns is that the coefficients estimated are directly interpretable as the long-term impact of the predictor variable. If the exact estimate of interest is the size of the one-year or three-year return as a result of a fixed change in the household risky ratio, then the most natural regression to run would be a long-horizon regression.

Another purported reason for using overlapping regressions is the higher R^2 of the regressions. However, especially for persistent predictors, the R^2 increase is mechanical. Valkanov (2003), Hjalmarsson (2006), and Boudoukh et al (2008) show that increasing the prediction period does not increase the power of tests. In fact, in monte-carlo simulations, the estimated coefficients of one period disaggregated returns and N-period aggregated returns are often correlated by .99 or more.

Other reasons often cited for using overlapping regressors include errors in the predictor variable (Cochrane Piazzesi 2005), missing observations, and higher

45

prediction accuracy. However, Harri and Brorsen (2009) show that for a vast majority of statistical reasons given for using overlapping regressions, better statistical predictors are possible. They also show that the usual inference correction methods for overlapping regressions are quite problematic, which is borne out in this paper in Table 1.8.

Turning to this paper's specific results, in Table 1.8, regressions (1) and (2) report disaggregated returns, returns with one-quarter prediction periods, replicating the univariate and multivariate results from before. Regressions (3), (4), and (5) run OLS no overlap (OLSNO) regressions with prediction horizons of one year or three years. OLSNO represents the process in which intermediate observations are dropped and OLS inferences are used. OLSNO is not perfectly efficient due to the dropped observations, but it preserves the validity of the inferences.

Table 1.8: Long Horizon Regressions.

OLS regressions of k-period ahead value-weighted CRSP excess returns on multiple predictors lagged one quarter:

$$R_{t,t+k} = \alpha + \beta_1 Risky_{t-1} + \beta_2 S_{t-1} + \beta_3 CAPE_{t-1} + \beta_4 Term_{t-1} + \beta_5 CAY_{t-1} + \epsilon_t$$

Where $R_{t,t+k}$ denotes the excess return of the CRSP value-weighted holding return for k quarters forward as reported by CRSP; $Risky_{t-1}$ denotes the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; S_{t-1} denotes the equity share in new issues as defined in Baker Wurgler (2000); $CAPE_{t-1}$ denotes the ten year cyclically adjusted price to earnings ratio, defined as per Campbell Shiller (1988); $Term_{t-1}$ denotes the yield premium of ten year over one month federal government obligations; CAY_{t-1} is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001).

Regressions are divided into three categories. For Disaggregated Returns, the prediction period k is always one quarter and so by construction there is both no overlap and no data dropped. *t*-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 periods of lags. For OLS No Overlap, the prediction period k is always ranges between one year (k=4) and three years (k=12). There is no overlap but data is dropped. *t*-statistics are shown in brackets using Eicker-Huber-White standard errors. For Overlap Corrected via Newey-West, there is no data dropped but overlap. *t-statistics* are calculated by Newey-West standard errors with 4+k periods of lags, a generalization of the number of lags used in the Disaggregated Returns method.

	Disaggregated Returns		OLS No Overlap – Intermediate Data Dropped			Overlap Corrected via Newey-West		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period	1-Qtr	1-Qtr	1-Year	3-Year	1-Year	1-Year	3-Year	1-Year
$Risky_t$	-2.49	-2.39	-10.22	-23.93	-8.58	-9.21	-18.42	-9.25
	[-4.16]	[-1.97]	[-4.31]	[-3.36]	[-1.48]	[-4.84]	[-5.91]	[-2.40]
S_t		-7.52			-29.75			-26.76
		[-1.16]			[-1.03]			[-1.69]
$CAPE_t$		.02			-.22			.07
		[0.11]			[-.39]			[0.17]
$Term_t$		-.31			-1.74			-.44
		[-0.72]			[-1.05]			[1.06]
CAY_t		55.94			265.62			153.04
		[1.24]			[1.33]			[1.08]
R^2	0.054	0.060	0.223	0.401	0.031	0.192	0.365	0.236

Regressions (3) runs a prediction with a one-year horizon. The coefficient

and R^2 is exactly four times that of the quarterly regression. However, note that

the power of the regression in terms of t-statistic does not increase, in line with

the theory of Harri and Brorsen (2009). Similarly, regression (4), which predicts

a three-year horizon has about double the coefficient and R^2 as regression (3), has

slightly lower t-statistics. That the coefficient and R^2 does not triple, and that the

power goes down, is all indicative of the fact that the horizon of the prediction

power does start breaking down sometime between one and three years. Three

years is also sufficient to see substantial decay in the persistence in the household

risky ratio predictor.

Regression (5) runs a one-year-horizon multivariate regression. As

expected from theory, the coefficients and R^2 of the one-year regression is about

four times the one-quarter numbers. The loss in efficiency from the dropped

observations is also clearer here as the power tends to be somewhat lower.

Regressions (6), (7), and (8) replicate regressions (3), (4), and (5)

respectively, except instead of dropping intermediate observations and using OLS,

it uses Newey-West regressions with 4+k periods of lag, where k is the horizon of

the prediction expressed in quarters. The coefficients are roughly the same as the

OLSNO regressions since Newey-West uses OLS for point estimates; the small

deviations are due to the fact that intermediate observations are no longer dropped.

The t-statistics however all seem higher than OLSNO. This is due to two reasons:

the efficiency is higher since no observations are dropped, and less desirably, the Newey-West standard errors are biased downwards (Harri and Brorsen 2009). This is evident in the fact that the t-statistic actually increases between one-year and three-year regressions, whereas theory dictates that it should decrease.

To conclude, Table 1.8 demonstrates the extension of the household risky ratio predictor to longer-horizons. The R^2 does increase from .054 for one quarter to .223 for one year and .401 for three years, which as an aside is quite high for known predictors. While the R^2 and coefficient increases, the power of the tests and t-statistics are not better. Sometimes power is worse due to dropped observations in the OLSNO regressions. The Newey-West regressions illustrate the complexity of correcting the inferences of overlapping regressions correctly. To maximize power then and minimize inference assumptions, this paper chooses to use disaggregated quarterly regressions, in line with the recent theory of Harri and Brorsen (2009)

1.4   Empirical Decomposition of the Household Risky Ratio

This section analyzes the empirical sources of predictability that the household risky ratio has for forward returns. This section decomposes the household risky ratio into components related to household tilt, valuation ratios, and issuances. Then, I run univariate and multivariate regressions against each of these components to see what the source of predictability is.

1.4.1   Define the components theoretically

In the section below, I will decompose the household risky ratio into the following three components, with terms to be defined below:

$$\log(Risky\ Ratio_t)$$

$$= \log(Household\ Tilt_t) + Valuation\ Ratio\ Part_t$$

$$+ Issuance\ Part_t$$

$$Valuation\ Ratio\ Part_t \perp Issuance\ Part_t$$

Consider again the household risky ratio, the ratio of household holdings of high-risk assets to low-risk assets. The first component of this ratio is the portion that is specific to the household, which is termed the *household tilt*. This can be defined as the portion of the household risky ratio that is not common to the economy. To define the household tilt, this paper then first defines an economy-wide risky ratio, termed the *total risky ratio* as follows:

$$Total\ Risky\ Ratio = \frac{Total\ Risky\ Assets}{Total\ Assets}$$

And calculate the difference between the total risky ratio and household risky ratio as:

$$Household\ Tilt = \frac{Household\ Risky\ Ratio}{Total\ Risky\ Ratio}$$

Note that this also gives:

$$Household\ Tilt = \frac{\%\ Risky\ Assets\ Held\ by\ Households}{\%\ NonRisky\ Assets\ Held\ by\ Households}$$

And

$$\log(Household\ Risky\ Ratio)$$

$$= \log(Total\ Risky\ Ratio) + \log(Household\ Tilt)$$

$Household\ tilt$, is then a measure of the household's holding of high risk over low-risk assets relative to the entire economy. Thus, $household\ tilt$ is cleansed of all economy-wide effects, effects that will be discussed later on that should affect high risk and low risk assets in general, having nothing to do with households in particular.

In this section, preference is also given to the log version of the decomposition due to the additivity of the components. Other sections use the raw non-log risky ratio for simplicity and transparency of construction. As shown in Table 1.9, the predictive properties of the log household risky ratio are almost identical to the household risky ratio. This is a natural consequence of the fact that the variation of household risky ratio is small relative to the neighborhood in which the log function is locally linear, and so log household risky ratio is nearly an affine transform of the household risky ratio.

The second component of the household risky ratio, encompassed entirely within the log total risky ratio, is the idea of a valuation ratio or past price changes. If the market value of all high risk assets homogenously doubles overnight then the total risky ratio would double as well. Note that this effect is completely subsumed in total risky ratio part of household risky ratio: there is nothing special about households. Many authors in the literature (Poterba Summers 1988, Fama

French 1988) have demonstrated that past prices changes, especially at the three to four year level, are good negative predictors of future returns. Looking at Figure 1.5 the predictability of different horizons, which can be seen as a type of cross correlogram, it is seen that indeed much of predictability of the household risky ratio becomes strong within the last three to four years of lags. To the extent then that the total risky ratio is capturing price changes, the total risky ratio has predictive power.

More fundamental than price changes are valuation ratios. Past returns are only weak predictors of future returns, and really the fundamental predictability comes from valuation ratios like D/P, CAPE, and B/M (Cochrane 2008). When run in a multivariate regression, the predictive power of past price changes is almost always subsumed by valuation ratios. As exposited by Campbell Shiller (2005), the predictive power comes from the fact that if earnings growth is difficult to predict, which is empirically the case, then CAPE changes must be attributed to future changes in expected return.

Table 1.9: Regression of forward CRSP return on various transformations of the household risky ratio.

OLS regressions of one quarter future equity market CRSP returns on multiple transformations of the household risky ratio:

$$R_{t,t+1} = \alpha + \beta_1 X_{t-1} + \epsilon_t$$

Where $R_{t,t+1}$ denotes the one quarter forward excess return as reported by CRSP. X_{t-1} variously denotes $Risky_{t-1}$ the household risky ratio calculated by dividing household high risk assets over low risk assets, as collected from the Federal Reserve Flow of Funds; or X_{t-1} denotes $1/Risky_{t-1}$ the multiplicative inverse of $Risky_{t-1}$; or X_{t-1} denotes $\ln Risky_{t-1}$ denotes the natural log of $Risky_{t-1}$. t-statistics are shown in brackets using Newey-West heteroskedastic and autocorrelation robust standard errors with 5 quarters lags. N=228

	(1)	(2)	(3)
Intercept	8.05	-4.87	7.43
	[5.42]	[-2.68]	[5.53]
$Risky_{t-1}$	-2.49		
	[-4.16]		
$\frac{1}{Risky_{t-1}}$		15.13	
		[4.12]	
$\ln(Risky_{t-1})$			-6.47
			[-4.12]
R^2	0.054	0.059	.057

A similar mechanism is at work thinking of the total risky ratio as a valuation ratio. Consider assets in the economy to be paying off a safe stream securitized into the present through low risk assets, and a high risk stream securitized into the present through a high risk asset. Then as the required risk premium on high-risk assets decreases, this will push up the valuation of the high risk assets with respect to low risk assets, increasing the total risky ratio. Thus, the total risky ratio acts as a valuation ratio in predicting future returns in the same way that CAPE does.

To capture this portion of the household risky ratio, termed the *valuation ratio part,* this paper projects the total risky ratio onto a standard valuation ratio: the CAPE of Campbell Shiller (1988):

$$\log(Total\ Risky_t) = \beta_0 + \beta_1 CAPE_t + \beta_2 CAPE_t^2 + \epsilon_t$$

$$Valuation\ Ratio\ Part := \log(\widehat{Total\ Risky_t})$$

Note then risky valuation as defined here is just an affine transform of CAPE and its square.

To arrive at the final component of the predictive power of household risky ratio, consider that the valuation ratio analysis above considers only the part of total risky ratio with a fixed quantity of assets. In reality, issuances and redemptions play a role in both the composition of the total risky ratio and its predictive power. In particular, Baker Wurgler (2000) studies corporation issuance of assets. They show that when corporations issue a higher fraction of

equities in a given period, there are lower future real returns on the equity market. After this seminal work, many other papers confirm that issuance, especially by corporations, are biased towards assets that are overvalued and will return less in the future.

Then in the projection above, a final component *risky issuance* can be defined:

$$\log(Total\ Risky_t) = \beta_0 + \beta_1 CAPE_t + \beta_2 CAPE_t^2 + \epsilon_t$$

$$Issuance\ Part := \epsilon_t$$

The *issuance part* reinforces the *valuation ratio part* even further. During good times, high-risk assets are overvalued and thus already have a large total valuation. Exactly during this time, corporations are issuing even more high-risk assets that are overvalued. The high-risk asset valuation measure then receives two reinforcing shocks during this time: a positive price shock and a positive quantity shock.

Putting all three components together, the accounting relationship below holds:

$$\log(Household\ Risky\ Ratio_t)$$

$$= \log(Household\ Tilt_t) + Valuation\ Ratio\ Part_t$$

$$+ Issuance\ Part_t$$

$$Valuation\ Ratio\ Part_t \perp Issuance\ Part_t$$

Thus, the log household risky ratio can be thought of as composed of three components: a household tilt variable, a valuation ratio part, and an issuance part. The last two are guaranteed to be orthogonal by mechanical construction. All three variables turn out to be decently orthogonal empirically as well.

1.4.2   Estimation of Components

In this subsection, I estimate the predictive powers of the components as defined above, which allows me to decompose the predictive power of the log household risky ratio. Table 1.6 displays the results of predicting future equity premium using the log household risky ratio and its decomposition above: the log household tilt, the valuation ratio part, and the issuance part.

Regression (1) replicates the baseline univariate regression, this time with the log household risky ratio. Regression (2) runs the same regression with log household tilt variable. The predictive power of the log household tilt variable is quite significant (t=-2.32) and has decent predictive power ($R^2 = .018$). The predictive power of log household tilt is actually quite high on an absolute scale (compare to Table 1.7), and is about a third of the of the entire log household risky ratio.   Also, note that the coefficient on the household risky ratio is -6.47 while the coefficient on log household tilt is -9.26. Thus the log household tilt is a relatively stronger driver component of the log household risky ratio. Unlike the entire log household risky ratio, the log household tilt component is much more orthogonal to the other predictors, both in theory and also can be seen by the lack

56

of interaction between regression (2) and (7). Thus, the household tilt variable is that theoretically novel part of the household risky ratio predictor. It is the component of household risky ratio's power that cannot be explained empirically by previous variables.

The next regression (3), looks as the predictive power of the valuation ratio part of the household risky ratio. The t-statistic is moderately significant (t=1.79), and the R^2 is decent at .0113, about a fifth of the predictive power of the log household risky ratio. The coefficient is -3.46, which is about half that of the log household risky ratio. This all suggests that the valuation ratio part is a relatively weaker driver of the predictive power of the household risky ratio. Also by construction, the valuation ratio is a linear combination of CAPE and $CAPE^2$, so all the predictive properties of CAPE carry over. Therefore, even though the t-statistics are not large, it is known from myriad of studies with longer term data (Campbell Shiller 1988) and rolling predictions (Campbell Thompson 2008) that the valuation ratio part must be a good predictor of the equity premium.

Regression (4) looks at the predictive power of issuance part. This variable is significant with a t-stat of 2.39, and an R^2 of .0164, again about a third of the predictive power of the log household risky ratio. The coefficient at -7.68 is also slightly higher than that on the log household risky ratio, demonstrating this is an important component of the predictive power of the household risky ratio. By construction, it is also orthogonal to the valuation ratio part. In the

above dimensions, the issuance part is quite similar to the log household tilt variable.

However, unlike log household tilt, the issuance part is not orthogonal to known predictors. Regression (9) adds the equity share of issuances to a regression with the issuance part. Both the magnitude and t-statistic of the issuance part are halved by equity share of issuances, showing the close interaction of the two terms. Also, by adding all other covariates in regression (10), it is seen that the predictive power of the issuance part is nearly driven to zero. Thus, the issuance part can be seen to be picking up the predictive power of already known predictor variables, chiefly equity shares of issuance.

The valuation ratio part by definition is a function of CAPE. However, compare univariate regression (3) with a multivariate regression with all covariates besides CAPE, regression (8). It is seen that adding covariates does not reduce the predictive power of risky valuation. If anything, it is increased somewhat. This is to be expected from the fact that none of the other covariates are valuation ratios.

Finally, consider regressions (6) and (7), which are multivariate regressions with a full set of covariates. Regression (7) only contains the log household tilt portion of the household risky ratio, while regression (8) contains the entire log household risky ratio, but note that the adjusted R^2 is nearly identical, as well as the t-statistics. This provides evidence once again that the

58

marginal predictive power of the log household risky ratio on top of the four covariates is isolated mainly to the log household tilt.

In summary, it is possible to decompose the log household risky ratio into three components: log household tilt, a valuation ratio part which is a function of CAPE, and the remaining issuance part. In a univariate sense, all three components are important and significant. With respect to known predictors, the predictive power of the log household risky ratio arises marginally from the log household tilt.

1.5   Discussion

A.   Ruling Out the Modigliani Miller Explanation

One possible explanation of the household risky ratio predicting future lower equity premium could be simple Modigliani Miller, as pointed out by Baker Wurgler (2000). In particular, Modigliani Miller posits that weighted average cost of capital is the same regardless of how corporations fund themselves. Therefore, as the amount of high risk assets in the economy increase with respect to low risk assets, the high risk assets effectively become less risky and command a lower return. Thus the household risky ratio could predict negative future equity returns just as an accounting artifact.

However a rough calibration shows that Modigliani Miller cannot explain anywhere near the size of the effect observed. Between the lowest and highest terciles, the household risky ratio doubles. Modigliani Miller would predict a

59

halving in excess returns. However, in reality excess returns drop about ten times. This order of magnitude difference is similar found in the equity share of corporate issuances in Baker Wurgler (2000).

1.5.1 Welfare Effects

The empirical result that the household risky ratio negatively predicts the equity premium does not take a stand on whether this predictability reflects a rational risk factor or a misoptimization. Assume the latter case. Then what is the Sharpe ratio lost to the household sector from sizing out of the equity market exactly when it is performing well? What is the loss compared to a constant-fraction-hold benchmark, or the optimal conditioning on predictors benchmark? If it is also assumed that households have log utility and no exogenous sources of time-varying utility then return loss can also be calculated.

I calculate three investment possibilities. The first is the actual returns and Sharpe ratio realized by the household sector assuming the fraction of high risk assets is the fraction households invest in the market index at any given moment, and the fraction of low risk assets in the fraction households invest in treasury bills in any given moment. This gives a Sharpe ratio of .267 and annualized log excess returns of 3.26% assuming log utility. The next possibility assumes that households always hold 70% equities and 30% treasury bills, the unconditional average amount of high risk assets versus low risk assets held by the households. This results in a Sharpe ratio of .311 or excess log returns of 3.76%. This

represents an increase of 50 bps over the baseline, or 14% increase over the base

amount.

If households instead were to scale in optimally assuming log utility, then

formula (14) of Campbell Thompson gives an increase of about 3x the current

baseline rate.

$$2.9 = \left(\frac{R^2}{1-R^2}\right)\left(\frac{1+S^2}{S^2}\right)$$

With $S^2 = (.31)^2 = 0.0967$. $R^2 = .206$.

This represents a return of 11% assuming log utility and a Sharpe ratio

of .533. Compared to actual outcomes, in case of optimal conditioning on

predictors, the returns are more than triple, and the Sharpe ratios are more than

double. While 11% may seem high, this is not out of line with the strength of the

predictor. Also, the actual amount that could be realized by a real agent

estimating out of sample would be less than this optimal ideal (Welch Goyal 2008,

Pastor Stambaugh 2009). Thus, mistiming represents a substantial welfare loss

for the household sector.

1.6 Extensions

1.6.1 A theory for demand systems

This paper has examined the empirical phenomenon of the prediction

power of the household risky ratio and decomposed the prediction power into

various empirical components. Here, I explore some ideas behind why the household risky ratio might have such predictive power in theory.

One approach to the theory is to see high risk assets and low risks assets as being cleared in markets or demand systems with different players, each with a net demand curve. Market clearing happens at the price and quantity that sets total net demand to zero. Such demand system view goes back to the seminal work of Brainard and Tobin (1968).

In this view, the household sector has a demand curve for high risk assets. The demand curve receives shocks that are exogenous, the source of which I will examine below. After a positive demand shock, households demand a higher quantity of the high risk asset for the same price. Assuming the remaining sectors' demand curves remain constant, this translates to higher prices on the high risk assets, or lower future expected returns. In this way, higher quantities and valuations of high risk assets held by the household translate to a negative relationship to future returns.

Recent literature like Baker Wurgler (2000) speaks to this demand system view of high risk versus low risk assets. They show that corporations' supply curve of high risk versus low risk (equity versus debt) rationally responds to prices. When equity prices are high, and hence future expected returns are low, corporations supply more equities. Baker Greenwood Wurgler (2003) show a similar phenomenon between maturities in the debt market: when the yield of a maturity is particularly low, and hence the price particularly high, corporations

tend to issue at that maturity. Greenwood Vayanos (2008) is more along the lines of this paper, in showing that government issuances of bonds are supply shocks: when the government exogenously issues excess bonds at a certain maturity prices go down, and yields are higher in the future.

It is important to note that in such a system a demand shock generally should cause both price and quantity responses, but the relative amount of each must be determined by the elasticity of the supply curve. That the response is not purely in price as evidenced here and in Baker Wurgler (2000) shows that securities are not in perfectly inelastic supply. A theoretical basis for this is apparent in that corporations can always start new real projects that are funded. That the response is not purely in quantities suggests that the supply is not perfectly elastic. Corporations need time to put new projects online, and new projects have aggregate diminishing returns in the economy as in the model of Solow (1956).

The fact that Baker Wurgler (2000) and this paper examine similar sets of markets explains why in the tables the issuance share of equities is so related to the issuance component of the household risky ratio. They examine flows by corporations, a subset of the supply side, while this paper examines stocks by households, a subset of the demand side.

On a first order then, some theoretical sense can be made out of the empirical phenomenon present in this paper by using a theory of demand systems and price pressure.

1.6.2   What Drives Demand Shocks?

Following the above idea that variation in household risky ratio is caused by demand shocks, this motivates the question of what causes the demand shocks.

One possibility is rational time variation in risk premiums and preferences. The underlying cause could be the same as that behind business cycles. A proximal model might include Campbell Cochrane (1999) habits. In particular, as stock returns receive a positive shock, excess consumption increases and effective risk aversion decreases. This justifies both a quantity shift from low risk to high risk assets as well as price increase as future required returns decrease. To test this model, a model of habit and surplus consumption could be calibrated, and habit can be correlated with the household risky ratio to see if a relationship exists.

Another strand of thought might explain the demand shocks as arising from sentiment. It is known that perfect optimization often does not describe individuals and even the firm (Laibson et al 1998; Zhang 2013). This may be seen as untestable as if the causes are fundamentally from outside the economy, then no immediate predictors are available, besides perhaps survey evidence, to validate such model. However, many behavioral models accept economic factors as the driver of psychology. For example Greenwood Shleifer (2013) posit that people have extrapolative beliefs: they believe that stocks will go up more following a move upwards. This case is testable as past prices can be used to see whether they relate to the household risky ratio.

1.6.3 Further Data

A central point of this paper is the empirical predictive power of the household risky ratio. A clear and transparent way to extend the empirical power would be to extend the duration of the data series as much as possible, in the style of Robert Shiller. Currently the results are based on more than 60 years of data, so if the data series were doubled, the series would extend back to around 1890. A first-order advantage of such a dataset extension is that it serves as a true out-of-sample test of the hypothesis above, since this paper is uncontaminated by observation of data before 1951. The true out-of-sample test can be used to validate the predictor in a way immune to any claims of data snooping, as well as test the stability of the coefficient estimated. The predictive properties can be understood much better by using the entire expanded series with rolling out-of-sample predictors.

Of course, data expansion has limitations. As is generally the case, data further back in history are noisier due to less advanced data collection technologies and more data that has been lost through time. Even common price series such as equity returns and price-earnings ratios become significantly lower quality before the 1920s. Quantity data such as that used in this paper would be even rarer. Further, for dates far enough back, one must question even the existence of equity markets accessible to households. As a raw method of increasing power, historical data extension is less fruitful, especially with the already high t-statistics observed here.

Another possible extension is to extend the predictability results here to other datasets. If the value of household holding of risky assets predicts future returns on risky assets, then it might stand to reason that the household holding of other asset classes might negatively predict future returns on those areas as well. Preliminarily, this seems to hold with government bonds and corporate bonds. In this way, the ideas and evidence presented in this paper can be developed into a general theory of household tilt.

Finally, this paper can be seen as an extension of the usual price predictors into quantity space. One way to generalize this more is to look at how all the price and quantity variables flow into each other economy-wide. High risk and low risk assets are just two components of what is the household sector's savings stock. The savings stock is affected by investment flows, which is known to be highly procyclical. Tracing investment back to output gives the GDP as the source of this flow. In this space itself, GDP depends on factor prices paid to labor and capital, the latter of which household total assets forms a component. On the capital markets side, there are corporations who are issuing the securities and financial assets being used in the household risky ratio. These corporations translate funds raised to real investment and real projects with payoffs and risk profiles.

Financial economics often centers around theories of prices in the economy and how they relate to each other, especially theories involving the efficient market. Household high risk and low risk assets then is a first step of a

journey towards looking at more quantity-type predictors and looking at the entire economic system to provide macroeconomic foundations for finance.

1.7   Conclusion

This paper shows that the ratio of high risk assets to low risk assets held by the household sector, termed the household risky ratio, is a negative predictor of future equity premium. The predictive power is robust and strong: the univariate t-statistics are above 4, and the annualized R^2 is above .20. The predictive power remains even after variation in construction of the variable, first/second half of the time series, and adjustment for the bias of Stambaugh (1999). The predictive power also is not subsumed by popular predictors like CAPE, equity shares of issuances, term spread, and the cay.

The paper empirically decomposes the predictability into three roughly orthogonal components. First is a household tilt component representing the preference of households for high risk assets above and beyond the entire economy. The second is a valuation ratio part that is a function of the Campbell-Shiller CAPE. And the third is an issuance part that is the residual from the decomposition above. All three components play important roles in the predictability of the household risky ratio: the R^2 is divided generally evenly between them and the coefficient size is the same order of magnitude. The second and third components reflect known predictors in the literature, while the first, the household tilt, seems orthogonal to known predictors.

67

This paper adds to the literature understanding time variation in equity premia by looking at Federal Reserve Flow of Funds data. It follows the footsteps of Baker Wurgler (2000) in going beyond price predictors to quantity data. Additionally, this paper looks at economy-wide household sector quantity data in a first step at connecting the variation in equity premium to economic fundamentals.

Chapter 2

2 Does Say-on-Pay Affect the Firm? Causal Evidence from a Regression Discontinuity Design

2.1 Introduction

The Dodd-Frank Say-on-Pay provisions require practically all significant public companies from 2011 onwards to hold regular say-on-pay votes, votes in which investors express general approval or disapproval of the compensation package proposed by the board for the executive team. This paper provides evidence on the causal effects of Dodd-Frank Say-on-Pay vote by drawing on a novel source of investor censure of executive pay. Specifically, I exploit variation in investor censure that results from the Say-on-Pay vote-fraction discontinuity. That is, the vote passes discretely above 50% and fails discretely below 50%.

69

Random selection of firms at the boundary into the passing or failing group provides identification.

The field of corporate governance, particularly executive compensation, has burgeoned in the last two decades (Murphy 1997). However, relatively few studies have focused on causal identification. Studies often just rely on regression control variables, and the few that use instruments often use ones that are not clean, and leave doubt on the validity of the exclusion restriction. This paper aims to provide causal evidence through a regression discontinuity design (RDD) with excellent internal validity properties.

Specifically, this paper identifies the impact of say-on-pay vote fraction using a regression discontinuity design, in the style of Lee and Lemieux (2009). The say-on-pay votes pass statutorily at 50%, and this cutoff can be used for identification. The usage of corporate votes as the discontinuity variable is in line with previous work by Cunat Gine Guadalupe (2012b).

To understand this identification, suppose the executive compensation vote contains random voter noise, at least some of which is independent of all other variables in the system. The random noise arises from a variety of sources: a voter forgetting to vote or missing an airplane to the meeting due to a weather delay. Due to such random noise, companies at the 50% cutoff are essentially randomized into being approved at 50.1%, and being disapproved at 49.9%. Outcomes from the two groups can be compared, and the difference in outcome can be attributed to random assignment.

The RDD above provides a strong and internally valid method for identifying the causal impact of passing versus failing say-on-pay. This estimate can directly answer the question of whether say-on-pay is causally effective at all at reining in CEO pay. The point estimate from the RDD for total CEO pay change is directly interpretable as the average causal effect of passing the vote for a company who would have received around 50% vote share. This provides a lower bound of the causal effect of say-on-pay, whose degree of censure monotonically increases as vote share decreases from 100%. This paper finds that say-on-pay does indeed causally rein in CEO pay.

As a second step, this paper identifies the causal impact of the vote cutoff on other company characteristics, such as financing policy, payout policy, and firm performance. This can be seen as the causal result on CEO behavior of receiving censure from failing say-on-pay and having his or her pay reduced. This paper finds no significant impact on company characteristics as a result of the vote cutoff.

2.1.1 Background on Say-on-Pay and Executive Compensation

The Dodd-Frank Act of 2010 is a wide set of reforms targeting the financial industry after the crises in 2008. One part of Dodd-Frank is mandating an advisory say-on-pay for all public companies above a certain size. This mandate was not solely motivated by the crises. Its inclusion in Dodd-Frank was

71

an opportunistic way of remedying perceived longstanding agency issues amongst companies' investors, boards, and executives.

A common view in this line is that the executives capture the board. Executives wield their influence over the board to increase their own pay beyond the optimum (Yermack 1997; Bertrand and Mullainathan 2001; Bebchuck and Fried 2004; Kuhnen and Zwiebel 2006). For example, Core, Holthausen, and Larcker (1999) show that measures board-of-director control and ownership structure has high R^2 for CEO pay levels, and signs are consistent with a story of board capture. Say-on-pay provisions allow investors to express dissatisfaction with excessive pay to signal investor preferences and to apply soft pressure to reduce pay (Bebchuk 2007).

In the efficient markets first-best optimum, the board should already pay executives the net present value maximizing sum. Even with agency costs and other structural issues, the company should endogenously structure their board and voting rules to maximize the second-best optimal value (Fama 1980). Under these efficient theories, mandating say-on-pay should not impact the firm.

Even some theories of inefficient pay also predict that say-on-pay would not have an impact. Kaplan (2007) and Bainbridge (2008) argue that shareholder advisory votes are likely to be bypassed by a compromised board. Cunat Gine Guadalupe (2012a) finds that when firms start holding say-on-pay votes, this by itself has no causal impact on CEO pay.

72

On the other hand, Bebchuck (2007) supports more corporate controls, arguing that making investors' opinion more salient to the board induces the board to represent investors better. (Bebchuk, Fried, and Walker 2002) suggest that say-on-pay would increase the outrage costs of overpaying the CEO, and thus reduce pay. Further, reputational concerns and enhanced shareholder voice, formalized by say-on-pay votes, will cause boards to overcome barriers to negotiating with CEOs on behalf of shareholders.

This paper also looks at the causal response of firm characteristics as CEO pay drops at the boundary. If performance suffers, then the argument can be made that wages were efficient before, and this meddling by investors is non-optimal. Many papers (Murphy 1999, Baker Jensen Murphy 1990, Joskow et al 1996, DeAngelo 1991) argue that due to political pressure, CEOs may be underpaid and under-incentivized. In such under-compensation theories, a reduction on CEO compensation should result in a reduction of firm value. Alternately, a CEO might be overcompensated, perhaps due to the inability of boards to evaluate the true cost of pay (Hall and Murphy 2003; Jensen, Murphy and Wruck 2004). In this case, firm value should not decrease and generally be close to zero[2] – which is what this paper observes.

---

[2] Under the weak assumption that incentives are non-decreasing in CEO compensation, then a reduction in pay should result in firm value increasing an amount weakly between 0 and the NPV of CEO compensation itself, both a small percentage of firm value.
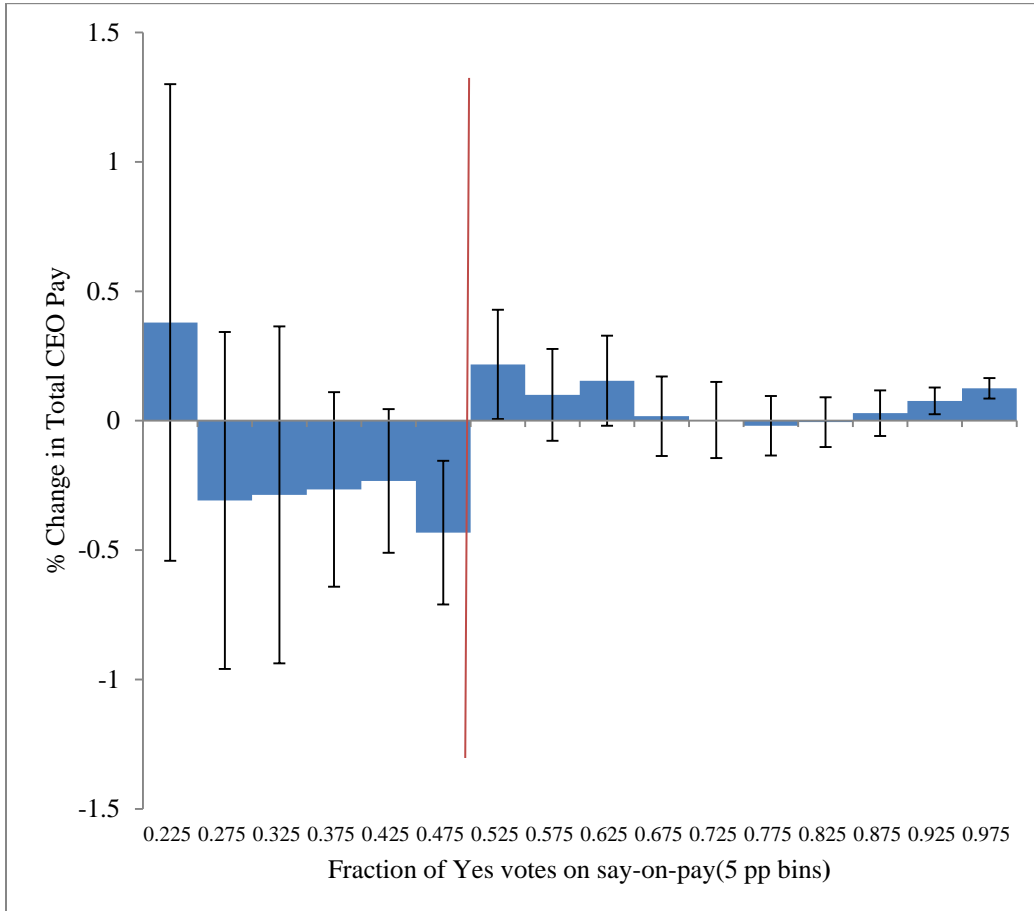
Figure 2.1: Change in CEO Total Compensation Binned by Vote fraction.

The change in Total CEO Pay from fiscal year 2010 to 2011 versus the 2011 Dodd-Frank Say-on-Pay vote fraction. Firms are grouped into 5 percentage-point bins. Votes that just pass are assigned to the 0.525 bin and votes that just fail are assigned to the .475 bin. 95% confidence intervals based on ordinary standard errors for the bin mean are given by the error bars. 0.50 is the pass-fail transition for say-on-pay votes. The red vertical line indicates the cutoff at 50%.

2.1.2 Preview of Results

The paper finds that say-on-pay has significant impact on subsequent executive pay. Barely failing the Dodd-Frank say-on-pay vote in 2011 has a -59.8% gross impact on next-year total CEO pay level, a jump that is clearly visible in Figure 2.1. This amount is equal to 1.3 times the standard deviation of the annual CEO pay change. Thus, boards do react to say-on-pay votes through the natural mechanism of changing CEO pay.

In order to reach this identification, this paper compiles a novel data set of Dodd-Frank say-on-pay executive compensation vote results from raw SEC data. The data was manually compiled from over 5000 SEC filings with no systemic format.[3] Riskmetrics also complies say-on-pay data, but there are large differences. The Riskmetrics data are about shareholder proposals for a firm which starts having say-on-pay votes. Riskmetrics currently has no data on the outcome of say-on-pay itself. These votes are from before 2009 and are sparse: only hundreds of such votes have occurred over all firms in a decade. These votes are also not mandatory: investors from a firm chose to propose starting say-on-pay votes, causing selection bias in the votes. Cunat Gine Guadalupe (2012a) run a regression discontinuity on the Riskmetrics vote and find shareholder proposals to start having say-on-pay votes have no impact on pay but do affect firm value through announcement returns.

---

[3] Not only did the data format vary wildly, but even the reporting of vote results varied. Some companies preferred tables, others lists, and still some narrative paragraphs or images of vote results.

This paper uses the regression methodology to further examine the impact of the vote cutoff on other firm characteristics. First, the announcement-day return is slightly positive but not significantly different from zero. A 95% confidence interval for the price impact of failing say-on-pay is [-1.4%, 1.6%]. This is a surprisingly small impact on the value of the firm given the substantial impact on CEO pay. It also provides evidence against the theory that at the margin CEOs are undercompensated due to political constraints. If that were the case, the large lowering of pay caused by say-on-pay would be expected to destroy significant value and result in a significant negative market reaction.

This paper contributes in two ways: first by providing causal identification in the field of corporate governance and executive compensation by using an interesting RDD vote cutoff. Second, this paper contributes by creating a novel dataset of Dodd-Frank say-on-pay vote results from the primary source manually.

Section II provides a literature review. Section III reviews data collection and dataset construction. Section IV looks at a descriptive OLS analysis of the data. Section V describes the RDD methodology both generally and the specific case that is implemented in this paper. Section VI provides argument that a treatment effect exists. Section VII provides the main result of the impact of vote cutoff on CEO pay. Section VIII discusses extensions. Section IX concludes.

2.2   Literature Review

Seminal work in CEO compensation include Roberts (1956), Baumol (1959), and Lewellen and Huntsman (1970). However, many empirical papers in this literature do not apply rigorous identification strategies to isolate causal inferences. Agency theory views of CEO pay arose around the 1980s. Holmstrom (1979) highlight the tradeoff between risk aversion of CEO versus the need to reward them with high-powered contracts. Holmstrom and Milgrom (1987) arrive at a linear solution to the problem. Jensen Meckling (1976) formalizes the agency issues and analyzes it in terms of financing needs, optimal incentives, and monitoring. Murphy (1997) reviews the empirical literature of CEO pay and provides a thorough description of pay practices, compositions, and setting procedures. Murphy also highlights the rapid increase in CEO pay in the 1990s, which Gabaix and Landier (2008) explain using extreme value theory.

While Fama (1980) espouses the theory that CEOs should have optimal pay set already, many papers also argue for a "skimming" view of CEO pay (Yermack 1997; Bertrand and Mullainathan 2001; Bebchuck and Fried 2004; Kuhnen and Zwiebel 2006). That there are inefficiencies in the CEO labor market may not be a surprise given inefficiencies in thick, public, and relatively arbitragable financial markets. Shleifer (2003) reviews divergence of paid price from fundamental value. In price divergence, Lamont Thaler (2003) provide a general overview and Zhang (2013) provides a specific example on options. Calvet Campbell and Sodini (2006) and Zhang (2013) suggest that households do make mistakes in asset allocation. Baker Wurgler (2006), Fuster Laibson Mendel

(2010), and Greenwood Shleifer (2013) demonstrate investor sentiment and psychology plays an important role beyond fundamental value.

Many of the original empirical studies on say-on-pay come from United Kingdom data, which was the first country to require say-on-pay in the form of a Directors' Remuneration Report resolutions. Conyon and Sadler (2010) find that most shareholders are generally satisfied with executive pay in this dataset. Ferri and Maber (2009) fail to find any across-the-board differences in pay policy comparing before and after the UK mandate. They do find that the few firms that received negative say-on-pay votes subsequently adjust down the magnitude of the pay and shift pay towards higher powered incentives. Alissa (2009) find that low say-on-pay vote predicts future reductions in excess compensation and increases CEO turnover.

In the United States Say-on-Pay evidence is more recent, but analogous studies of public and shareholder opinion stretch back further in time. Core et al. (2008) find that firms generally do not respond to negative press, while Ertimur et al (2011) find that firms do respond to shareholder pressure by reducing pay. The difference between public versus shareholder opinion seems to be sharp. Balsam and Yin (2012) use Dodd-Frank Say-on-Pay data and show that firms that receive more disapproval reduce pay and increase the performance sensitivity of that pay. Cai and Walkling (2009) show that the market reaction to the House of Representatives passing Dodd-Frank say-on-pay was positive for firms with high abnormal CEO pay, suggesting say-on-pay creates value for some firms. The

general consensus of these non-causal studies seems to be that say-on-pay changes pay when a firm receives especially negative votes, but say-on-pay does not change pay across the board. There does not seem to be any causal studies of the impact of say-on-pay votes on compensation.

Thisltlewaite and Campbell (1960) set out the framework the regression discontinuity design (RDD) strategy. Lee and Lemieux (2010), Imbens and Lemieux (2008), and Hahn Todd Van Der Klaauw (2008) have advanced the theory of RDD into the current best practices used today, including the use of a non-parametric kernel regression as optimal. Imbens and Kalyaranaman (2009) provide a method for optimal bandwidth selection for that kernel.

Due to the sharp cutoff often seen in votes of all types, voting has been a common application of RDD, with some notable examples being electoral competition on policy (Lee Moretti, Butler 2004), the effect of gender on legislator behavior (Rehavi 2007), the effect of mayoral party identification on policy (Ferreria and Gyourko 2009). The application of RDD to corporate voting is more novel. Cunat Gine Guadalupe (2012b) use RDD to show that firm values casually increase by 2.8% when the firm implements proposals that increase shareholder control.

This paper is most similar to Cunat Gine Guadalupe (2012a), which looks at the causal impact of firms that decide to start holding say-on-pay votes. They attain causal identification since many firms only start holding say-on-pay as a result of a shareholder proposal. They use the vote cutoff on these proposals in

Riskmetrics data. They mainly conclude that only starting to hold say-on-pay is similar to implementing other positive shareholder control provisions: there is no executive compensation change and firm value increase around 2%.

This paper differs from Cunat Gine Guadalupe (2012a) in a few important dimensions. Most critically, Cunat Gine Guadalupe 2012a uses shareholder proposal votes on *whether to start* holding say-on-pay, while this paper collects novel data and uses the say-on-pay votes *themselves*. Thus, Cunat Gine Guadalupe 2012a causally identifies the overarching impact of implementing say-on-pay, while this paper identifies the specific impact of the say-on-pay vote itself.

This fundamental difference gives rise to a host of other differences. Cunat Gine Guadalupe (2012a) uses shareholder proposals about implementing say-on-pay from 2006-2010, while this paper uses votes mandated by Dodd-Frank from 2011, reducing selection bias. Cunat Gine Guadalupe (2012a) uses Riskmetrics data on shareholder proposals, while this paper compiles a novel data set manually across a wide variety of formats. Since the vote is different, the results are also different. Cunat Gine Guadalupe (2012a) finds no impact on CEO compensation, while this paper finds large differences. Cunat Gine Guadalupe (2012a) finds a positive impact on firm value, while this finds no impact on firm value.

2.3   Data Collection

Data was collected from a variety of sources. Most of the data collection effort went into manually collecting vote data. Observations in the data set are generally firm-year pairs, although the bulk of the study focuses on the 2011 vote year and so some variables are associated with firms instead of firm-years. Table 2.1 provides a summary of the data. The universe consists of all members of the Russell 3000 subject to Dodd-Frank Say-on-Pay[4]. The data is restricted to Dodd-Frank say-on-pay firms because the main dependent variable studied is Dodd-Frank say-on-pay vote, and the restriction to Russell 3000 is not significantly limiting: the vast majority of firms subject to Dodd-Frank say-on-pay lie inside the Russell 3000. The restriction is done to minimize data absence due to covariates, as here limiting the inference to the Russell 3000 is preferable to possible selection bias at the margin.

Table 2.2 provides a timeline overview of data sources. Returns data for individual stocks was collected from CRSP. Daily frequency data from CRSP is used directly for announcement day returns. Monthly frequency data from CRSP is rolled up to the annual level for yearly return variables. All returns in this paper are in log terms.

CEO compensation data is collected from EXECUCOMP. This paper examines only CEO compensation, and so filters for the CEO flag. Data is collected both for the current year and the past year to track changes. The leading

---

[4] For the first year of Dodd-Frank Say-on-Pay this consists of all firms with more than $75 million in market capitalization.

compensation figure this paper uses is SEC total compensation, which refers to SEC Form DEF 14A compensation, giving us the advantage of being able to cross validate the numbers manually and being free from model assumptions used to calculate other EXECUCOMP pay variables. Salary, options, and other compensation components are also selected to match the SEC Form DEF 14A.

The compensation variable is always selected to be the grant-day value instead of realized value. There is robust discussion encompassing Murphy (1999) to present day about whether grant or realized value is more appropriate in different situations. For example, for studying inequality, realized pay is quite relevant. This paper choose grant value both because it is easier to collect, being available contemporaneously, as well as because it is the more appropriate value for this study. After all, this paper studies how the board changes their package in response to votes, and how the CEO responds to the granted package. Using realized value would subject the study to unneeded noise due to stock market movements.

To minimize look-ahead issues, the compensation data from EXECUCOMP is always from the contemporaneous year's DEF 14A. The paper does not use data that is later revised for consistency. Data on company characteristics is collected from COMPUSTAT, which is ultimately from the SEC 10-K filing forms.

Table 2.1: Summary Statistics for Vote Data, CEO Pay, and Firm Characteristics.

The summary statistics below describe the 2010 and 2011 data for all 2188 firms of the Russell 3000 which has 2011 vote data. Pay data and firm characteristic data are collected for both 2010 and 2011. Variables including pay are reduced to the circa 1,300 firms for which EXECUCOMP pay data exists. Panel A summarizes variables associated with each firm in the dataset, while Panel B summarizes variables associated with each firm-year. Dodd-Frank Say-on-Pay vote fraction is collected manually from SEC 8-K forms. A high vote fraction indicates more approval. All pay relates to the CEO and is computed from EXECUCOMP: Total Pay as reported on SEC form DEF 14A (TOTAL_SEC), base salary paid to CEO (SALARY), bonus paid to CEO – usually tied by rule to performance metrics (BONUS), value of options awarded to the CEO (OPTION_AWARDS_FV). Accounting variables are obtained from Compustat Annual: earnings before interest and tax (EBIT), dividends paid that year (DVC), an indicator for whether the firm pays dividends I(DVC>0), book value of equity (CEQ), book value of assets (AT). Institutional Ownership is pulled from Yahoo Finance. Returns data in logs and market cap are from CRSP: Stock Return in 2010 (termed R here), Market Cap (Shares Outstanding x Price on Jan 1, 2010, termed M here), and Announcement Day Return is computed from market close before the announcement to the next day's market close.

All monetary amounts are in current year dollars. K stands for thousands and M for millions. All changes between two years use the average of the two years as the base. Note that the number of observations changes due to missing values for certain data sources and variables. *SD* denotes standard deviation, and *Corr w/ Vote* denotes correlation with Vote %. For correlations, OLS significance is calculated with significance at the 10%, 5%, and 1% levels indicated by *, †, and ‡, respectively.

Panel A: Firm Variables				
Variable	N	Mean	SD	Corr w/ Vote
Vote fraction (%)	2,188	89	12	1.00‡
Total CEO Pay Change 2010-2011 (%)	1,310	7	47	.10‡
Institutional Ownership 2010 (%)	2,091	69	24	-.10‡
Return in 2011 (%)	2,181	-8	35	.08 *
Return in 2010 (%)	2,162	22	31	.18‡
Return in 2009 (%)	2,056	34	50	.03 .
Return in 2008 (%)	1,994	-59	61	.10 †
Announcement Day Stock Return %	2,189	0	6	.02

Table 2.1 (Continued)

	2010				2011			
Variable	N	Mean	SD	Corr w/ Vote	N	Mean	SD	Corr w/ Vote
Total Pay ($K)	1,325	5,726	6,101	-.25‡	1,311	6,153	6,732	-.16‡
Salary ($K)	1,325	807	451	-.16‡	1,311	846	432	-.15‡
Bonus ($K)	1,325	272	1,109	-.09‡	1,311	272	1,159	-.09‡
Options Value ($K)	1,325	1,191	2,676	-.17‡	1,311	1,188	2,361	-.12‡
EBIT ($M)	2,175	635	2,559	-.08‡	2,172	717	3,003	-.07†
Market Cap ($M)	2,174	5,605	19,670	-.02	2,172	5,725	20,779	-.05
Dividend ($M)	2,175	107	488	-.07†	2,172	118	538	-.07†
Paying Dividends (%)	2,175	48	50	-.01	2,172	50	50	-.05
Book Equity ($M)	2,175	2,579	9,804	-.09‡	2,172	2,730	10,309	-.09‡
Book Assets ($M)	2,175	11,481	81,102	-.05*	2,172	12,155	82,394	-.05*

Panel B: Firm-Year Variables

Table 2.2: Timeline of Data Collection and Votes: Exxon Mobile Corporation.

The below shows the timeline of release of various SEC Forms. The table lists the year name as used in the paper, the release date of the form, the subject period of the form, and the subject matter of the form for Exxon Mobile Corporation.

This paper uses the convention that years always refer to year of the *direct* subject period and not the release date. For example, note that the Dodd-Frank Say-on-Pay vote released on 5/31/2011 is referred to as the 2011 say-on-pay vote and not 2010, despite the vote being about 2010 pay. This is because the vote itself occurred in 2011 and is the direct subject of the data release.

Year Name	SEC Form	Release Date	Subject Period	Subject Matter
2010	10K	2/25/2011	FY 2010	Accounting Data
2010	DEF 14A	4/13/2011	FY 2010	Proposed Executive Compensation
2011	8-K	5/31/2011	5/26/2011	Reporting the Results of the say-on-pay held a few days ago on FY2010 Proposed Executive Compensation

Data on votes were manually collected from SEC 8K filings on the
EDGAR databases online, and collecting this data is one contribution of this
paper. Companies have a wide variety of filing formats. Some file the data
within multiple tested Tables in a highly structured HTML file, while others write
the vote results narratively in a flatter text-like file, and formats exist in the entire
spectrum in between. Despite the inability of such disparate data to be
automatically processed, the vote results and interpretations are almost always
straightforward.

For each company participating in Dodd-Frank say-on-pay in that year, I
record the number of FOR votes, AGAINST votes, ABSENTIONS, and
BROKER NON VOTES. I calculate the vote passage percentage as the number
of FOR votes over the total of FOR, AGAINST, and ABSENTIONS. With the
majority of firms, the nominal vote fraction required for passage is 50% of FOR
votes over FOR, AGAINST, and ABSTENTIONS. For votes around the cutoff
(40-60%), I manually check the company's bylaws or voting statements to ensure
this is the case. For the minority where the denominator does not include
ABSTENTIONS in the denominator, the ratio is updated to reflect that. No
company in the data set crosses the pass/fail boundary due to this adjustment, as
the adjustment is on average very small at a few basis points. I also note the date
of record (the vote date) and the date of filing/report. The date of filing is used as
the announcement return date as it is the date on which vote results are made
public and can be publically traded on.

Table 2.2 provides intuition about the sequence of events and data. Consider one observation: Exxon Mobile Corporation (XOM) for year 2011. The lagged subject year then is fiscal year (FY) 2010[5] and the current year is XOM's FY2011. Both fiscal years coincide with the calendar year, with FY2010 closing 12/31/2010. The SEC requires XOM to first produce an annual report of FY2010 financial activities, which XOM does on 02/25/2011: this provides all the lagged-year company characteristic data we collect through COMPUSTAT. Then the SEC requires XOM to hold an annual shareholder meeting, the details of which are announced in Form DEF 14A on 04/13/2011. Form DEF 14A is perhaps the one most affected by say-on-ay, as it contains the compensation committee's proposal for the lag year 2010. This is the data that we collect through EXECUCOMP.

Observing performance and compensation, shareholders then vote in 2011 with respect to the FY2010 pay package. The voting may happen via proxy (mail or online), or during the in-person meeting on 05/25/2011. The results of the meeting are then announced a few days later on 05/31/2011 in Form 8-K. It is this Form 8-K data that we manually collect to form 2011 vote data.

The two SEC Forms, 10K, DEF 14A are released in calendar year 2011 but speak to FY2010 data. By convention, this paper associates data with the year it directly describes. Thus both these forms contain 2010 data. The final SEC

---

[5] Fiscal years are numbered after the year in which they end. For example, years ending on the following dates would all be called FY2010: January 15, 2010, June 30, 2010, and December 31, 2010.

Form, the 8-K, reports a vote that occurred in 2011 and so contains 2011 data. Table 2.2 provides an outline of the year naming conventions.

To get changes, this paper needs to collect data for the next year. For XOM, the 10K for 2011, released in 02/24/2012, provides information on changes in company characteristics and performance. The DEF 14A for 2011, released on 04/12/2012, provide data on compensation response to the 2010 say-on-pay vote.

Note here that for year 2011, the penumbra of our data actually spans three years. The lagged year for accounting and compensation is 2010, while the current year 2011 data is needed to obtain changes. Note that 2011 data is released well into 2012, giving three relevant years. Even though Dodd-Frank was enacted as a law in 2010, only the 2011 year is complete, with 2012 data still arriving well into 2013, the writing of this paper. For the 2011 data year, the say-on-pay vote is held in 2011 and is actually for shareholder approval of the FY2010 pay package, which is generally already fixed. Rarely is the compensation committee recalled to determine 2010 pay again. Rather the model here is that the board receives feedback on the investor's dissatisfaction of the past year's compensation, and then adjusts the next year's compensation accordingly.

2.4   Data Overview and Ordinary Least Squares (OLS) Results

This paper begins its analysis by examining the data descriptively in an ordinary-least-squares setting. OLS gives up good identification for higher power,

which is a large advantage and worth considering. First, determinants of CEO

pay are examined, particularly in the context of its correlation with Dodd-Frank

Say-on-Pay vote fraction. The paper finds that while CEO pay is quite

predictable, there is serious reverse causality issues in an OLS regression of CEO

pay against vote fraction. Then, determinants of CEO pay change are examined

to minimize reverse causality issues. Vote fraction is shown in this setting to

have the same correlation as in regression discontinuity design (RDD) case.

Finally, determinates of vote fraction are examine, and this paper finds that pay,

firm characteristics, and firm performance contribute similar magnitudes to this.

2.4.1   Magnitude of CEO Pay

Table 2.3 presents log total CEO pay as the left hand side variable, and

examines its covariates. The paper examine 2011 pay only. Three standard

groups of covariates are selected from standard CEO pay predictors in the

literature. The first group consists of size measures like market value, total assets,

and total sales. The second group consists of stock performance measures

including last one and three year returns. The final group consists of firm

performance measures like book to market, equity ratio, and ROA. Further, the

paper examines lagged pay and vote fraction as special covariates. These

variables are chosen to be the standard predictors in the literature and their

variations, as set out originally in Murphy (1997), and described in more detail in

Balsam and Yin (2011).

Table 2.3: OLS Determinants of CEO Pay.

OLS regressions of Log 2011 total CEO pay (TOTAL_SEC) on Log 2010 total CEO pay , vote fraction (DFSOP), Market Cap (M), Book Asset (AT), Sales (SALE), 1-year and 3-year past log returns, Book to Market (AT/M), Equity ratio (CEQ/AT), and ROA (EBIT/AT).

$$\log(TOTAL\_SEC_{2011})_i$$
$$= a + \log\big(TOTAL\_SEC_{2010,i}\big) + c\, DFSOP_i + dM_i + eAT_i + fSALE_i + gR_{2011,i}$$
$$+ h\big(R_{2011,i} + R_{2010,i} + R_{2009,i}\big) + i\left(\frac{AT_i}{M_i}\right) + j\left(\frac{CEQ_i}{AT_i}\right) + h\left(\frac{EBIT_i}{AT_i}\right) + u_i$$

All variables are for the year 2011 year unless otherwise specified. All variables in the formula above are defined as in Table 2.1. Log 2010 total CEO pay is the natural log of the 2010 SEC reported total pay. Vote fraction is the Dodd-Frank Say-on-Pay vote fraction expressed as a fraction. Market cap is market value at the start of the year in millions of dollars. Book Assets is total book assets in millions. Sales is annual sales in millions. 1-year return is defined as 2011 year returns for a stock, while 3-year returns is 2009 to 2011 returns inclusive for a stock. Book to market is total book assets over total market value expressed as a fraction. Equity ratio is book value of equity over book value of total assets as a fraction. ROA is EBIT over total assets as a fraction. T-statistics displayed in parenthesis below each estimate and are heteroskedasticity robust. Bold coefficients are significant at the 10% level.

	Panel A: Regressions (1) to (6)						
	Standard Literature Covariates					Lagged Pay	
Variable	(1)	(2)	(3)	(4)		(5)	(6)
N	1117	1290	1281	1150		1309	1116
R^2	.2487	.2532	.0065	.0475		.6739	.7043
2010 Pay						**.939**	**.957**
						(16.62)	**(9.50)**
Vote Fraction							
Market Cap	.029	**.129**					-.027
	(0.29)	**(1.98)**					(0.39)
Book Asset	**.171**	.072					.038
	(2.48)	**(2.23)**					(0.70)
Sales	**.203**	**.204**					-.013
	(6.27)	**(7.80)**					(0.48)
1-Year Return	-.135		**.225**				**.255**
	(1.57)		**(2.41)**				**(3.13)**
3-Year Return	**.139**		.061				-.023
	(2.12)		(1.06)				(0.59)
Book to Market	-.154			.011			.006
	(1.47)			(0.15)			(0.06)
Equity ratio	-.048			**-.473**			-.051
	(.29)			**(8.08)**			(1.05)
ROA	.043			**.148**			.009
	(1.26)			**(3.99)**			(0.40)

90

Table 2.3 (Continued)

Variable	Panel B: Regressions (7) to (10)			
	Including Vote Fraction			
	(7)	(8)	(9)	(10)
N	1309	1116	1310	1117
R^2	.6747	.7045	.0293	0.2751
2010 Pay	**.947**	**.961**		
	(16.37)	**(9.19)**		
Vote Fraction	**.258**	.125	**-1.540**	**6.850**
	(1.67)	(0.56)	**(6.55)**	**(11.93)**
Market Cap		-.029		.053
		(0.41)		(0.54)
Book Asset		.039		.152
		(.71)		(2.29)
Sales		-.012		.183
		(.47)		(5.89)
1-Year Return		**.256**		-.121
		(3.13)		(1.42)
3-Year Return		-.027		.177
		(0.66)		(2.78)
Book to Market		.009		-.183
		(.09)		(1.78)
Equity ratio		-.048		-.059
		(1.02)		(0.91)
ROA		.009		-1.654
		(0.40)		(7.42)

Looking at Table 2.3 regression (1), the R^2 of CEO pay on all three groups of variables is .25. Regression (2) shows that almost all the predictability can be accounted for by size measures – which have an R^2 of .25 as well. This is not surprising given the importance of firm size to CEO pay. Most literature on CEO pay include firm size as a components, and all analysis find strong and robust relationship between the two variables, including the seminal work by Murphy (1997) . To reword Gabaix and Landier (2009), one reason firm size affects CEO pay so much is because CEO ability can be seen as a multiplier on the firm's market cap. Therefore, larger companies gain from skillful CEOs an amount proportional to their size.

In the size regression (2), it is seen that all measures of size positively predict pay. This suggests that market value, total assets, and sales can all be seen as noisy measures of a fundamental size variable, which increases pay.

Table 2.3 part (3) regresses CEO pay against stock performance. There generally seems to be a positive relationship between performance and subsequent year pay, but comparing (1) and (3) shows these figures are heavily affected by covariates. This suggests that regression (3) is interacting heavily with size variables, that much of the explanatory power of returns is coming through the mechanical impact of returns on size. Regardless, the predictability is low with an R^2 of .01 showing that returns are not great predictors of CEO pay.

Table 2.3 part (4) regresses CEO pay on firm characteristics. The R^2 is

moderate at .05, with a significant negative loading on the equity-ratio and

positive loading on ROA. One interpretation of this result is that ROA is a

measure of accounting performance, and CEOs who do well on this metric are

rewarded more. In fact, Murphy (1997) mentions ROA as an explicit bonus term

in many CEO's compensation schedules. This doesn't explain the equity ratio

though. Also, these effects are much lower in the total regression (1), suggesting

another explanation: omitted variable bias for firm size. It seems likely that only

large firms can obtain low equity-ratio, and large firms tend to be run better in a

way that yields high ROA. Thus, equity ratio and ROA are simply weak proxies

for firm size in the absence of an explicit size variable. This omitted variable

story is strongly supported by comparing the R^2 of (1), (2), and (4) and noting

that firm characteristics have negligible residual R^2 on top of firm size.

Up to this point, this section has only reported some standard determinants

of CEO pay. To get towards the question of the impact of say-on-pay on CEO

pay, consider regression (9), a univariate regression of CEO pay on the most

recent say-on-pay vote. Here the coefficient on vote fraction is negative. Taken

causally, this would seem to imply that receiving a low say-on-pay vote actually

causes CEO pay to be higher – opposite to the intuition that say-on-pay is a

disciplining mechanism. However, this fails to take into account an important

factor: reverse causality. High CEO pay in 2011 is likely very related to high

93

CEO historical pay in 2010. It is high historical pay that causes vote fraction to be so low in 2011, explaining the negative sign. In other words, high 2011 pay is likely a correlated of 2010 overpay, which causes low 2011 vote fraction.

Regression (10) adds the three groups of covariate back in along with vote fraction. The sign on vote fraction is still negative. This is not surprising as well – the same critique as regression (9) applies. Namely, 2011 CEO pay is a positive residual predictor of historical overpay. The historical overpay generate reverse causality on vote fraction.

Regression (8) attempts to solve this problem by adding lagged CEO pay. Lagged pay clamps down on the reverse causality channel by controlling for pre-existing overpay. Sure enough, the sign on vote fraction flips signs to the intuitive forward causal one: high vote fraction results in higher pay next year controlling for last year's pay. Regression (7) shows again the importance of controlling for historical pay, especially in comparison to regression (9): just controlling for the reverse causality channel by adding historical pay to the univariate regression causes vote fraction to flip signs.

The previous few regressions show that OLS regressions provide poor identification for the question of the causal effect of vote fraction on CEO pay. Reverse causality is shown to be a substantial issue here. Even with reverse causality eliminated, there are problems of pre-existing differences between firms that receive high and low vote fractions. As Table 2.1.2 amply shows say-on-pay

vote fraction is quite predictable on previous data, highlighting serious omitted variable bias issues.

To wrap up the analysis on Table 2.3, note from regression (5) that last year's pay alone has incredibly high predictability for this year's CEO pay, with an R^2 of .67. This is substantially more than all the other factors combined in regression (1). Further, once a regression includes previous pay, all other variables matter little. The R^2 of all other variables plus past pay in regression (7) is only .70. Of course, past pay is not included in the standard literature on CEO pay because the standard literature is attempting to predict pay from other variables. However, for the purposes of this paper, past pay is a legitimate covariate that demonstrably helps control for omitted variable issues. Historical pay is so important that in the next OLS analysis, this paper explicitly differences out historical pay by looking at changes in pay.

Table 2.4: OLS Determinants of CEO Pay Change.

OLS regression of change in CEO pay (PAY_CHANGE) on Dodd-Frank Say-on-Pay vote fraction (DFSOP), 2009 to 2011 log returns (R), and 2010 and 2011 values of: market cap (M), book asset (AT), sales (SALE), book-to-market (AT/M or B/M), equity ratio (CEQ/AT), and return on asset (EBIT/AT or ROA).

$PAY\_CHANGE_{2010-2011,i}$

$$= a + b\,DFSOP + c\,M_{2011,i} + d\,AT_{2011,i} + e\,SALE_{2011,i} + f\left(\frac{AT_{2011,i}}{M_{2011,i}}\right)$$
$$+ g\left(\frac{CEQ_{2011,i}}{AT_{2011,i}}\right) + h\left(\frac{EBIT_{2011,i}}{AT_{2011,i}}\right) + i\,M_{2010,i} + j\,AT_{2010,i} + k\,SALE_{2010,i}$$
$$+ l\left(\frac{AT_{2010,i}}{M_{2010,i}}\right) + m\left(\frac{CEQ_{2010,i}}{AT_{2010,i}}\right) + n\,R_{2011,i} + o\,R_{2010,i} + p\,R_{2009,i}$$
$$+ q\left(\frac{EBIT_{2010,i}}{AT_{2010,i}}\right) + u_i$$

Change in CEO pay (PAY_CHANGE) is defined as the difference between 2010 and 2011 total SEC reported pay, divided by the average total SEC reported pay over the two year as a fraction. All other variables are defined the same as in Table 2.1 and 2. T-statistics are heteroskedasticity robust. Bold coefficients are significant at the 10% level.

Variable	(1)		(2)		(3)	
N	1020		1020		1310	
R^2	.0502		.0520		.0104	
	b	[t]	b	[t]	b	[t]
Vote Fraction			.170	(1.05)	**.3717**	**(3.25)**
2011 Market Cap	.152	(1.29)	.152	(1.29)		
2011 Book Assets	**.340**	**(3.03)**	**.337**	**(2.99)**		
2011 Sales	.022	(0.14)	.018	(0.12)		
2011 B/M	.037	(0.28)	-.137	(0.73)		
2011 Equity ratio	-.032	(0.58)	-.033	(0.61)		
2011 ROA	.004	(0.15)	.004	(0.20)		
2010 Market Cap	-.191	(1.47)	-.189	(1.46)		
2010 Book Assets	**-.308**	**(2.45)**	**-.306**	**(2.43)**		
2010 Sales	-.016	(0.10)	-.009	(0.06)		
2010 B/M	-.145	(0.78)	-.137	(0.73)		
2010 Equity ratio	.072	(1.06)	.075	(1.12)		
2010 ROA	-.041	(0.34)	**-.043**	**(1.81)**		
2011 Return	.061	(0.52)	.053	(0.44)		
2010 Return	-.015	(0.24)	-.027	(0.42)		
2009 Return	**-.085**	**(2.37)**	**-.085**	**(2.38)**		

Table 2.5: OLS Determinants of Vote Fraction.

OLS regression of Dodd-Frank Say-on-Pay vote fraction (DFSOP) in 2011 on last year (2010) values of total CEO pay (TOTAL_SEC), change in total CEO pay (PAY_CHANGE), market cap (M), book assets (AT), sales (SALE), book-to-market (AT/M), equity ratio (CEQ/AT), ROA, and institutional ownership (IO), as well as 2008 – 2010 inclusive annual log returns.

$$DFSOP_{2011,i} = a + b\,TOTAL\_SEC_{2010} + c\,PAY\_CHANGE_{2009-2010} + d\,M_{2010,i} + e\,AT_{2010,i}$$
$$+ f\,SALE_{2010,i} + g\left(\frac{AT_{2010,i}}{M_{2010,i}}\right) + h\left(\frac{CEQ_{2010,i}}{AT_{2010,i}}\right) + i\left(\frac{EBIT_{2010,i}}{AT_{2010,i}}\right) + j\,R_{2010,i}$$
$$+ k\,R_{2009,i} + l\,R_{2008,i} + u_i$$

Dodd-Frank Say-on-Pay vote fraction is a vote taken in 2011 regarding 2010 pay. All regressors are chosen to be the latest possible data that is known by the time of the vote. Change in CEO pay (PAY_CHANGE) is defined as the difference between 2009 and 2010 total SEC reported pay, divided by the average total SEC reported pay over the two year as a fraction. IO is institutional ownership percent as a fraction defined in Table 2.1. All other variables are defined the same as in Table 2.1 and 2. T-statistics are heteroskedasticity robust. Bold coefficients are significant at the 10% level.

Variable	(1)		(2)		(3)		(4)	
N	1077		1198		1920		1988	
R^2	.1841		.0641		.0592		.0637	
	b	[t]	b	[t]	b	[t]	b	[t]
Total CEO Pay	-.028	(2.53)	-.028	(5.10)	.372	(3.25)	.372	(3.25)
Change in Pay	-.021	(2.82)	-.010	(1.47)				
Market Cap	-.016	(1.33)			-.013	(1.32)		
Book Assets	.023	(1.92)			.007	(0.72)		
Sales	-.002	(0.44)			-.002	(0.65)		
B/M	-.065	(2.21)			-.063	(3.01)		
Equity ratio	.000	(0.01)			.005	(0.57)		
ROA	.141	(1.84)			.097	(4.02)		
IO	-.060	(3.01)			-.078	(6.49)		
2010 Return	.109	(5.97)					.085	(8.52)
2009 Return	.052	(4.41)					.031	(4.82)
2008 Return	.054	(5.86)					.042	(7.67)

2.4.2 Changes in CEO Pay

This paper proceeds by examining predictors of CEO pay changes. As argued previously, CEO pay change is the more appropriate variable on which to measure the impact of say-on-pay: it suffers much less from reverse causality issues.

Regression (1) runs a baseline regression of the change in CEO pay on both the 2011 and 2010 values of the standard predictors of CEO pay levels. These predictors were examined in the previous subsection and Table 2.3. This includes 2011 and 2010 values for size variables, firm characteristics, and stock performance. Two items to note is that the 2011 variables have generally the same signs as in the regression (1) from Table 2.3, while the 2010 variables have the opposite sign. This is expected as the left hand side variable here is CEO pay variables difference across two years: 2011 minus 2010.

However, the R^2 is significantly lower on this regression than regression (1) from Table 2.3. This suggests that differencing CEO pay between two years eliminated most of the predictable components of CEO pay. This could be expected from realizing that most of the predictor variables like Total Asset or Market Value do not significantly change over the period of a year.[6] Further, this result is not a surprise given the low residual predictive power demonstrated between regressions (5) and (6) in Table 2.3.

---

[6] While a firm might easily halve or double in value over a given year, this change in market cap is a small fraction of the distribution of market caps across our universe of the Russell 3000.

In regression (3) of Table 2.4, this paper examines the straight up univariate regression of vote fraction on CEO pay change. The sign here is significant and in the positive direction, showing that high vote fraction is correlated with positive pay changes. Taken directly, this seems to say that firms do react to low vote fraction by reducing pay. Regression (2) adds on the controls to vote fraction, which results in the same sign on vote fraction, although a smaller magnitude. The smaller magnitude is not completely a surprise because as Table 2.5 will show much of the variation on vote fraction is predictable using the same covariates as those added in Regression (2). The difference between (2) and (3) suggests that much of the predictable variation on vote fraction is still affecting pay. This would be in line with the role of say-on-pay in reminding the board of the true cost of pay, and increasing the salience of their dissatisfaction (Hall and Murphy 2003; Jensen, Murphy and Wruck 2004).

2.4.3   Dodd-Frank Say on Pay Vote fraction

Finally, this paper uses OLS regressions to descriptively explore vote fraction and how it correlates to other variables. Table 2.5 includes as regressors the standard regressors as described in Table 2.3 for CEO compensation. In addition, it includes as regressors past pay, motivated by its use in standard say-on-pay studies (Cai and Walkling 2009). The pay variables include pay levels as well as pay changes. Institutional ownership is also added as it generally is highly affected by voting.

This results in three natural a priori categories for covariates. The first category consist pay measures: both levels and changes. The second category consists of firm characteristics, including size, book-to-market, equity ratio, return on assets, and institutional ownership. The final category consists of past returns for three year (See Table 2.5).

Regression (1) predicts vote fraction using all variables, while (2), (3), and (4) regress on each individual group described above. Note that there are minimal interactions between all four regressions. Comparisons of any regression with the all-in regression (1) do not show large changes in any variable's value. This signifies that the groups of variables are jointly orthogonal with each other. This makes sense – returns are generally orthogonal with everything else. Further, pay seems to be not strongly related to firm characteristics. This is verified in Table 2.3 where the only real predictors of pay is firm size, and even that had an R^2 of .25.

Further, note that the R^2 of the total regression (1) is .18, and the regressions of each of the groups of variables are .06. That they add up is a consequence of the orthogonality discussed above. That they're about the same in magnitude demonstrates that these factors almost equally contribute to vote – each component is similarly important.

Total pay negatively predicts vote fraction – this makes sense as historical overpay upsets investors and decreases their say-on-pay approval. Note that this

was the reverse causality result of Table 2.3 regression (9). Similarly, increase in pay in the last year predicts a lower vote fraction the following year. Investors may see the increase in pay as excessive or arbitrary.

Firm size variables don't seem to have a strong relationship with say-on-pay. Book to market is negatively related, which would be expected if a high B/M indicates a low market value from management leading the firm astray. ROA is similarly positively related to vote fraction, perhaps as a measure of CEO ability. The equity ratio itself does not affect vote fraction. Institutional ownership is negatively related to vote fraction, not surprisingly. It is a well-known result that institutional owners tend to wield more influence and express more displeasure with firms.

Finally, past returns are surprisingly good, robust, and consistent predictors of vote fraction. Each year's performance is positively related to vote fraction, and the magnitudes are nearly equal amongst the years. Each year is highly robust as well. The textbook goal of investors trying to maximize net present value seems right (Brealey Myers 2006). Investors do normatively and positively see firm market value as the key objective for executives to maximize, insofar as they consistently censure CEOs that fail on that metric.

Overall, the OLS regressions shown in Table 2.1, Table 2.2, and Table 2.7give significant insight into the relationship of the variables of each other. Primarily, it is seen that CEO pay levels are predictable mainly by size, but suffer from reverse causality if run against say-on-pay vote fraction. CEO change levels

are less predictable and have a positive sign with respect to vote fraction. Vote

fraction itself is somewhat predictable, with equal predictability arising from past

pay behavior, firm performance characteristics, and past returns of the firm.

2.5   Regression Discontinuity Design Methodology

This paper follows the theory of Van der Klaauw (2002) for regression

discontinuity (RDD), and implements estimation of the discontinuity using local

linear regressions due to its optimal rate properties (Porter 2003). The bandwidth

of the local linear regression is selected using the algorithm designed for this

purpose in Imbens and Kalyanaraman (2009). This RDD strategy is the gold

standard implementation in the Stata software used.

2.5.1   Basic Theory of Regression Discontinuity Design

Let $S \in [0,1]$ be scoring variable, in the case of this paper, the Dodd-

Frank Say-on-Pay vote fraction, a continuous variable. Also let $T \in \{0,1\}$ be a

binary treatment variable be completely determined by the scoring variable. In

particular, define $T = I(S \geq .5)$. Here, the interpretation would be statutory

passing the say-on-pay vote. Let Y be any treatment variable. In this paper, the

treatment variable would include pay change, announcement day returns, change

in company characteristics, and so forth. The leading case will be percent year-

on-year change in compensation.

The causal variable that this paper will try to estimate is the impact of treatment on outcome. A starting point would be a straight-up OLS estimator:

$$Y_i = \beta + \alpha \cdot T_i + u_i \quad (3)$$

The variable of interest would then be α. However, T_i is nonrandom and endogenous, and generally $E[u|T] \neq 0$. If the treatment effect depends on individuals, then the coefficient has no causal meaning (versus being $E[\alpha_i]$ in the randomized case).

In this paper, the discontinuity is sharp, meaning that T is identically 0 on [0,0.50) and 1 on [.50,1]. It is not fuzzy, with just a jump in treatment probability at .50 but not necessarily from 0 to 1. This is a direct consequence of the definition of treatment here as statutory passing of say-on-pay vote. The identification strategy will be based on the observation that a sample of firms within a small interval around the cutoff is similar to a randomized experiment at the cutoff. One way to see this would be to interpret the randomization as the independent noise that might arise from voter noise such as a block voter missing a train to the company meeting.

However, such independent randomization is not the only possible interpretation of RDD. Firms right above and below the cutoff of .50 have essentially the same vote S, so on average the firms right above are very similar to the firms below when receiving treatment T. Thus, as this interval decreases, there are vanishing confounds between the group to the left and group to the right

other than treatment. The treatment effect then can be isolated by comparing the difference in average of the two local groups.

This logic leads to strong internal validity of the estimate. The difference between the two groups can be highly attributed to the causal impact of treatment T. For treatments that vary between individuals, the estimate will converge to the average around the treatment $E[\alpha_i | T = S]$. The conditional operator indicates that since identification occurs with S around T, inference outside this neighborhood will not necessarily be valid.

However, reducing the estimation neighborhood comes at the cost of increasing variance., as it decreases the sample size. An alternative is to model the background impact of the vote fraction S on outcomes Y directly. The weakest possible assumption is that the direct impact of S on outcomes Y is continuous (if it were not, the intuition above for comparing borderline groups would not be correct). Thus the theory makes the following assumption:

Assumption 1: The conditional mean function $E[u|S]$ is continuous at the cutoff .50.

This immediately gives that the causal effect of passing the vote is identified by:

$$\lim_{S \downarrow .50} E[Y|S] - \lim_{S \uparrow .50} E[Y|S]$$

As a regularity assumption, since the cutoff is defined as being on the left-hand-side of .50, it is also required that the theory has

Assumption 2: The mean treatment effect $E[\alpha_i|S]$ is right continuous at .50.

However, in practice the entire premise of the RDD estimation is problematic if the treatment effect depends heavily on S. Thus, an implicit underlying practical assumption here needed for economical meaningfulness of the estimate is that the average treatment is valid for a wide neighborhood around the cutoff.

Stronger assumptions about the conditional mean function, termed the "control function" $E[u|T,S]$ increases the power of the test by allowing more data. If the control function $k(S_i)$ could be known perfectly, then the estimating equation

$$Y_i = \beta + \alpha \cdot T_i + k(S_i) + \omega_i \quad (5)$$

would identify the causal treatment effect α efficiently. However, the control function itself is not known and needs to be estimated. Original methods of estimating the control function include a simple linear function, suggested by Goldberger (1972) and Cain (1975). Concerns about differential linear effects on both sides of the cutoff motivate differential estimation on each side. Concerns about higher order effects further motivate higher order polynomial estimation.

In cases where control functions are estimated, misspecification is a serious concern. Components of the control function that is orthogonal to estimation may load heavily on the discontinuity. For example a cyclic control

function estimated with polynomials would show a discontinuity where none existed. The discontinuity estimate varies strongly with specification, opening serious data snooping concerns. The state of the art technique, which this paper uses, is non-parametric estimation, which minimizes specification error. In particular, this theory uses a local linear regression, with optimal bandwidth for this purpose as proposed by Imbens and Kalyaraman (2009).

Due to the local linear regression used to estimate the control function, we impose a stricter **Assumption 1':** the control function $k(S_i)$ is continuously differentiable. In the equation above, differential treatment effect $E[\alpha_i|S]$ is rolled into the control function, so this paper imposes a stronger **Assumption 2':** the treatment effect function $E[\alpha_i|S]$ is continuously differentiable. For power in practical estimation using locally linear regressions, it is important that the functions not only be continuously differentiable, but that the continuity be uniform. Ideally if we define the modulus of continuous differentiability as:

$$\omega := \sup_{x,y} \frac{|k'(y) - k'(x)|}{|y - x|}$$

Then the modulus should be low with respect to the number of data points present. The magnitude of this modulus directly affects the power of discontinuity estimator.

More intuitively, the baseline effect k not at the discontinuity $.50$ should be decently linear and well behaved. For this paper's leading case of looking at the effect of say-on-pay on change in pay, refer to Figure 2.1. The linearity of

106

bins heights before and after the discontinuity provides suggestive evidence that the control function is well-estimated by a locally linear function. Proceeding with this estimation, this paper arrives at a graph in Figure 2.2 of a discontinuity. With these assumptions, this paper goes ahead and uses estimator (5) with nonparametric locally linear functions for $k(S_i)$ to estimate many other treatment effect at the discontinuity.
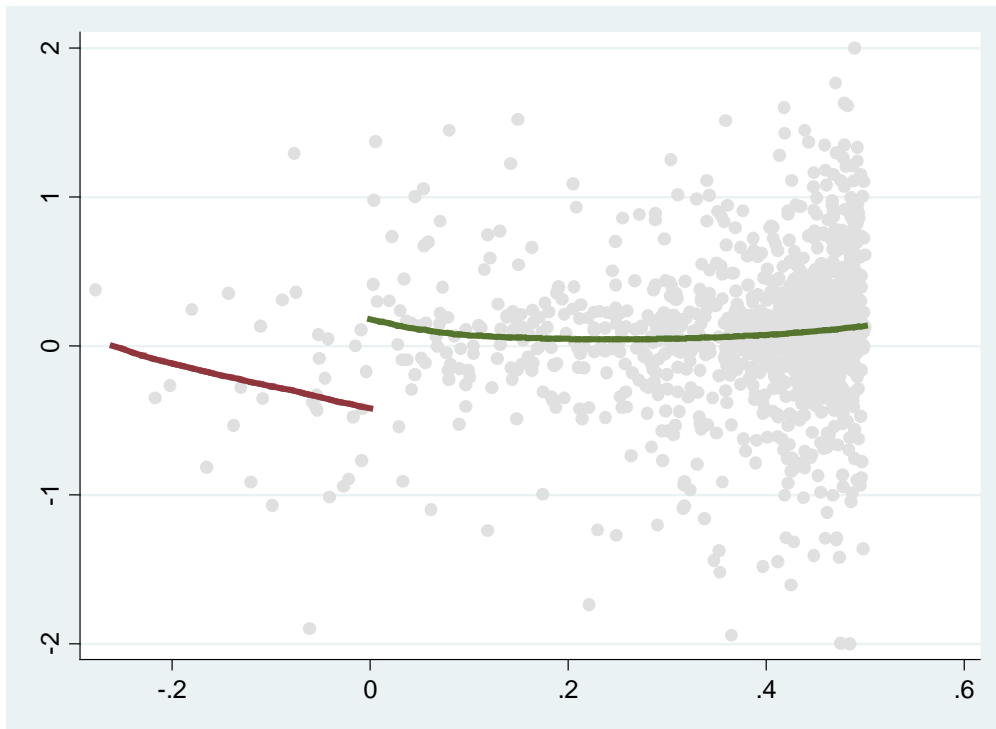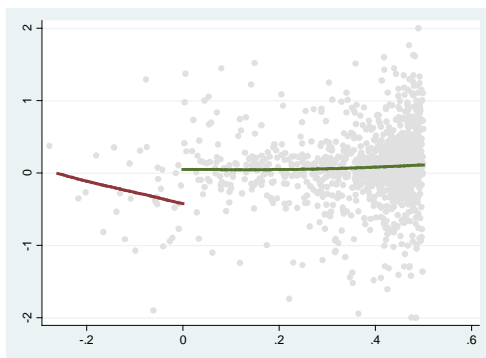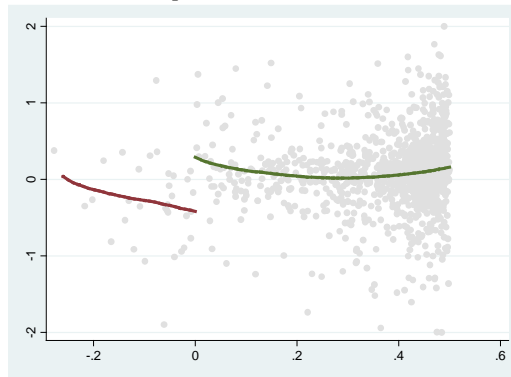
Figure 2.2: Kernel Regression Discontinuity Plot of Total CEO Pay Change vs Vote Fraction.

Plots of Total CEO Pay Change (%) between 2010 and 2011 versus 2011 vote fraction. The grey scatter plot show the raw data, while the superimposed lines show a non-parametric kernel regression above (green) and below (red) the cutoff. The kernel regression allows a discontinuity at 0, and the optimal bandwidth is defined as in Imbens and Kalyaranaman (2009).

2.6   Defining the Treatment

The treatment in our paper, the binary variable corresponding to T_i in our theory above, is whether a firm receives approval in its Dodd-Frank say-on-pay vote. The vote is advisory, so unlike a shareholder vote for a director or a bylaw, there is no mandatory action that occurs between passing and failing. Failing a vote mandates neither a reconvention of the compensation committee nor any docking of CEO pay.

Therefore, the paper first sets out to establish that a discrete effect might indeed be expected to occur at the cutoff. The evidence for this will come from language on the SEC form, the discrete effect on publicity from failing, and the voluntary emphasis of firms themselves on the passage or failure of say-on-pay.

First, we show that there is indeed a statutorily discrete change that occurs at the 50% cutoff. SEC Form DEF 14A serves as the official authoritative and detailed instructions for investor voting. DEF 14A presents material information on which investors may rely for their votes, such as the compensation package being offered to the CEO for the past year. DEF 14A also defines the exact statutory meaning of the vote. As demonstrated in Figure 2.3 Panel A, most firms have language precisely stating that for passage of the advisory vote, "a majority is required". In the context of DEF 14A this is precise language meaning 50% exactly or above, the same language used for binding votes like voting in board members. Some firms further state whether abstentions count in the denominator

109

when calculating this precise 50% Figure. In practice, almost no firm's say-on-pay passage is affected by abstentions. A copy of the proxy vote form is shown in Figure 2.3 Panel B.

Firms then report the results of the vote in SEC Form 8-K. The wording of the results has very wide latitude, with some firms opting to use data tables, and others describing the results. Many firms use a standardized table which shows either "Approved" or "Not Approved" as a standard column (Figure 2.3 Panel C). This format of report again emphasizes the discreteness of passing versus not. Even firms that do not use the column format will mentioned their votes were approved somewhere, perhaps in paragraph form. Even though few firms emphasize disapproval, the inability to say the vote was approved is a discrete change.

On the media front especially, disapproval has special meaning. The Wall Street Journal hosts a "shame list" of all firms that have failed say-on-pay. Many law firms, such as Semler Brossy and Sullivan Cromwell, have lists of all failing companies on their website. More than a dozen lawsuits have been filed against companies relying on the finding that the company failed its say-on-pay vote (Hickok and Rainville 2013). While these lawsuits often do not succeed due to the advisory nature of the votes, that the suits mainly rely on say-on-pay failure as the basis is significant. One example of a USA Today article that named and emphasized all failing firms within a certain scope is shown in Figure 2.4 Panel A.

Panel A: Vote Instructions from SEC Form DEF 14A

Proposal 11 (advisory "say on pay"): This proposal requires the approval of a majority of votes cast at a meeting at which a quorum is present.

Panel B: Proxy Vote Form from SEC Form DEF 14A

```
                                                        ##10113
                                                                    FOR   AGN   ABS
10.*- ADVISORY VOTE ON NAMED EXECUTIVE OFFICER COMPENSATION ------------->>>    FOR  --->>> 11.    0    0    0
       (THE BOARD RECOMMENDS A VOTE FOR THIS PROPOSAL)                   #029640
```

Panel C: Dodd-Frank Say-on-Pay Vote Results from SEC Form 8K

Advisory Vote on Executive Compensation

Vote Results	% Votes For	For	Against	Abstain	Broker Non-Votes
Approved	94.12%	5,563,078,796	347,351,143	19,688,679	1,317,035,730

Figure 2.3: Vote Instructions, Forms, and Results.

This figure shows three facsimile excerpts of various Dodd-Frank say-on-pay voting documents from Exxon Mobile Corporation (XOM). Panel A excerpts vote instructions written by a firm for its investors on SEC Form DEF 14A. Panel B excerpts the proxy vote form itself. This form is contained in SEC Form DEF 14A and is physically used by investors to vote on Say-on-Pay. Panel C excerpts the firm reporting Say-on-Pay vote results through SEC Form 8K

Panel A: USA Today Names Companies that Failed Say-on-Pay

USA TODAY | Money   Subscribe | Mobile   Google USA TODAY stories, photo

Home | News | Travel | Money | Sports

Money: Markets | Economy | Personal Finance | Stocks | Mutual Funds | ETFs | Cars | Real Estat
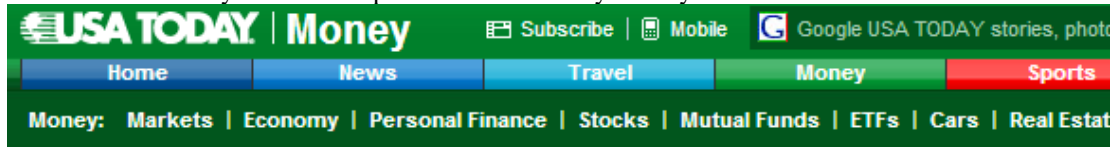
CEO pay soars while workers' pay stalls

By Matt Krantz and Barbara Hansen, USA TODAY   Updated 4/4/201

were tallied, shareholders at four companies voted their opposition to the pay strategies, says Mark Borges, principal of executive pay firm Compensia.

The companies where shareholders rejected pay plans include Beazer Homes USA, Jacobs Engineering, Shuffle Master and Hewlett-Packard, based on Borges' analysis. Borges expects the number of rejected CEO plans to increase dramatically as votes come up in shareholder meetings in April.

Panel B: USA Today Names Companies that Failed Say-on-Pay

"We held our first advisory stockholder vote on executive compensation at our annual stockholder meeting in 2011. We received an **affirmative advisory vote**, with over 53% of the shares voted casting votes in favor of our say-on-pay proposal. While we received a **positive vote**, the Compensation Committee strives to receive as high a vote as possible from the stockholders with respect to advisory approval of executive compensation. The Compensation Committee considered whether changes needed to be made mid-year to 2011 compensation, and decided **not to make changes in 2011 in light of the affirmative vote.**" – Rigel Pharmaceuticals (DEF 14A 2012/04/12)

"Following the annual shareholders' meeting in 2011, the Board agreed that simply rethinking and redesigning our compensation programs was not enough to address shareholders' obvious concerns about lack of profitability and the Company's executive compensation policies. Instead, the Board decided it was crucial to start at the Company's very foundation." After 48% approval rate, Stewart Information Services. Following this announcement, the firm fired a CEO, created an independent board, and made the Chief of HR independent from the CEO.

Figure 2.4: Citations of Approval or Rejection of Say-on-Pay.

The facsimiles below show citations of when a company's Dodd-Frank Say-on-Pay vote is approved versus rejected. Panel A shows a USA Today story naming companies that failed the vote. Panel B shows companies citing the vote approval or rejection as reason behind certain actions. The companies in Panel B were selected to be right near the cutoff.

Finally, the companies themselves cite approval or failure as explicit reasons in changing their compensation policy. Figure 2.4 Panel B shows a particularly striking example, where one firm that barely passed emphasized the passage of say-on-pay as approval for their current compensation plans, and continued with their package. On the other hand, another firm saw its failure of say-on-pay as reason to massively shake up the executive team. Many such examples exist, and a large portion of firms around the vote cutoff of 50% cite approval or failure for specific actions that they take.

While it is possible that such citation may not be causal, and instead be rhetorical, the fact that it is emphasized so much does suggest that there is discrete impact at the cutoff, whatever the cause. And further evidence goes to show that it is likely not all rhetorical. Table 2.6 shows that besides discrete changes in citing approval or regret, there are large differences in structural pay changes or off cycle pay changes on the two sides of the cutoff. For both pay change types, the majority of firms right above the payoff did not make changes, and the majority of firms to the left did, with large differences in the two groups.

By examining SEC forms, the discrete effect on publicity from failing, and the voluntary emphasis of firms themselves on the passage or failure of say-on-pay, we have shown that there is likely reason to believe that a discrete effect would occur at 50%. The data on CEO pay and other outcomes supports this hypothesis strongly.

Table 2.6: Pay Policy around Dodd-Frank Say-on-Pay Vote Cutoff.

Comparison of payout policy of the 10 companies immediately above and below the 2011 vote cutoff. The Table examine whether the DEF 14A 1) mentions the word "approved" or "favored" for last year's vote 2) mentions the word "regret" for last year's pay policy 3) explicitly mentions a change in pay structure, and 4) explicitly mentions a pay change off-cycle.

	Mentions Approved or Favored	Express Regret	Structural Pay Change	Off Cycle-Pay Change
Below Cutoff	0 of 10	9 of 10	9 of 10	6 of 10
Above Cutoff	7 of 10	3 of 10	4 of 10	2 of 10

2.7 Main Results: Impact of Votes

2.7.1 Graphical Evidence of Impact on Total Pay

Figure 2.1 shows the percentage change in total pay from 2010 to 2011 as the Dodd-Frank Say-on-Pay vote fraction varies. To be non-parametric about the dependency of pay change on vote fraction, the firms are assigned to 5% bins at round cutoffs. Observe that generally the dependency of pay change on vote fraction is continuous. However, at the 50% cutoff point, there is a discrete jump in pay. In fact, the bins below and above the 50% mark are the only ones where the 95% confidence intervals do not overlap.

Also note that the relationship between vote fraction and pay change does not seem monotonic. While we do not run any formal tests, non-monotonicity is not a surprise. An immediate explanation is that firms already have some idea of their say-on-pay vote fraction before it occurs. They can predict using surveys of investors, news articles, as well as econometric regressions. Balsam and Yin (2012) provide strong evidence of vote predictability. Thus 2010 pay is set conditional the company's prior expectations for the vote fraction. A 48% vote then can be worse news than a 30% vote; in the 48% vote case the company may have thought they would have passed, whereas a 30% vote firm likely knew they would fail. Analogously, a 52% vote could be better news than a 70% vote. This then would explain the non-monotonicity in pay change, especially the non-monotonicity observed around the 50% cutoff.

115

Simple bins like Figure 2.1 provide graphical evidence of the discontinuity

and give a sense of the size of discontinuity. However, it is not rigorous – the

boundary of the bins do not account for optimal bandwidth. Bins can be too large

and suffer from bias if the effect of the score variable is no longer constant across

the bin. Bins can be too small and lose power. Thus this paper proceeds run

optimal-bandwidth non-parametric estimates.

2.7.2   RDD Regression of Total CEO Pay and Robustness

The graphical output of say-on-pay vote fraction on change in Total CEO

Pay is shown in Figure 2.2. Here this paper again sees a large discontinuity before

and after 50%. There is again also non-monotonicity. The theorized reason of

non-monotonicity above, that it comes from outcome surprise, is again validated

here, as the non-monotonicity is confined to around 50%, and is more obvious for

smaller kernel bandwidths.

Table 2.7 shows that the estimate of the change in pay is quite significant:

corresponding to a 59.8% rise in pay at the 50% cutoff.[7] The t-stat is also strong

at above 4.0, providing evidence against data snooping. While this may seem like

a large change, note that CEO pay changes as a baseline are quite variable. A

---

[7] Note that to calculate the 59.8% figure, this paper uses as the denominator the average of the two years. By way of example, a drop from \$100,000 to \$50,000 would be calculated as $\frac{\$100,000-\$50,000}{\frac{\$100,000+\$50,000}{2}} = 67\%$. All changes are thus bounded between +200% and -200%. The paper calculates changes this way to avoid outlier effects on estimates, and changing the calculation base year to either the first year or the second year does not qualitatively change estimates.

majority fraction of all CEO pay comes from non-salary components, which allows for this variance. In fact, as the Table 2.1 summary statistics shows, the standard deviation of CEO pay change is itself 47%. Thus, the pay rise at the cutoff represents about 1.3 standard deviations of CEO pay change. Further, the 59.8% rise in pay matches the magnitude of the OLS estimate of CEO pay change as the vote fraction changes by a standard deviation (see Table 2.4), providing validation that the magnitude is correct.

This paper further verifies the regression through a few more robustness checks. First, this paper tests against the main misspecification possible with non-parametric estimates, bandwidth selection. Using both double and half the optimal bandwidth, Table 2.3 Panel B shows that the total pay change % is still large and significant (t>3). In fact, all estimates lie within the original estimate's 95% confidence interval, providing a measure of internal validation.

Figure 2.1 Panels B and C show the graphical plot of these alternate-bandwidth regressions. For the 0.5x optimal bandwidth estimate, the non-monotonicity around 50% is clearer, and the non-parametric fit to this non-linearity gives a higher estimate of the discontinuity. For the 2.0x optimal bandwidth, other than a linear fit, the two sides of the discontinuity are treated as approximately constants. As expected, a broader bandwidth goes more towards the naïve regression case as shown in equation (3).

The next robustness check that this paper runs is dummy tests at the 45% and 55% cutoff. Unlike the suggestive a priori evidence for the 50% cutoff, there

is no reason to believe that these two points should have any special significance. Indeed, we see that for both cases in Table 2.7 Panel B, the t-stats are well below 1. Thus, running cutoffs at these points provides a baseline giving suggestive evidence to the validity of the 50% result, much like a bootstrap. Additionally these baselines provide evidence against a mere round-number effect as well.

As a final robustness check, the paper tests to see whether there is a discontinuity at the 50% cutoff for pre-existing variables like market capitalization, levels of CEO pay, SIC code, and past year pay changes. No discontinuity are found in any such pre-existing variables, although given the small sample size, this robustness check does not have a large amount of power.

Table 2.7: Main Results: Changes in Company Characteristics at the Vote Threshold.

This Table presents univariate discontinuity regressions of changes of various independent variables (Y_i) on Dodd-Frank Say-on-Pay using discontinuity kernel regressions (function k) with optimal bandwidth. The sample is the same as Table 2.1. The coefficient α on the discontinuity indicator is shown below. The regression run is:

$$Y_i = \beta + \alpha \cdot I(VOTE > 50\%) + k(VOTE) + \omega_i$$

Base variables from Table 2.1 & 2 are differenced between 2010 and 2011 to obtain Y_i. Percent (%) differences for a base variable X is calculated as $(X_{t+1} - X_t)/((X_{t+1} + X_t)/2)$. Percentage point (pp) differences for a base percent variable X is $X_{t+1} - X_t$. Panel A regresses against CEO Pay Variables. Panel B checks the robustness of CEO pay results; 0.5x and 2.0x indicates a kernel regression with half and double the optimal bandwidth; and 45%,55% indicating redefining the cutoff from 50% to those points. Panel C extends the data to 2012 for the plus or minus 20% region around the ctuff. Panel D regresses against company accounting and other variables. Unless specified otherwise, the cutoff is defined at 50%, and the kernel bandwidth is 1x the optimal. t-statistics in parenthesis are calculated via Nichols (2007), with p<10% bold.

	N	Pass (α)
Panel A: CEO Pay		
Total CEO Pay Change 2010-2011 (%)	1,325	**59.8%**
		(4.05)
Salary as fraction of Total Pay (pp)	1,325	**-13.1%**
		(2.47)
Bonus as fraction of Total Pay (pp)	1,325	**-7.9%**
		(1.81)
Options as fraction of Total Pay (pp)	1,325	6.4%
		(0.67)
Panel B: Robustness of CEO Pay		
0.5x Optimal Bandwidth	1,325	**70.7%**
		(3.88)
2.0x Optimal Bandwidth	1,325	**47.5%**
		(3.35)
45% dummy cutoff	1,325	11.7%
		(0.44)
55% dummy cutoff	1,325	0.5%
		(0.03)
Panel C: Extension of Data to 2012		
Total Pay Change 2010 to 2012, -20%; +20% around cutoff	116	**48.0%**
		(1.98)
Total Pay Change 2011 to 2012, -20%; +20% around cutoff	116	**44.8%**
		(1.65)
Panel D: Company Financials		
Announcement Date Return (pp)	2,188	-0.0%
		(0.11)
EBIT (%)	2,175	**-35.9%**
		(1.82)
Dividend (%)	2,175	0.2%
		(0.31)
Paying Dividends (pp)	2,175	-4.3%
		(1.28)
Equity ratio (pp)	2,175	18.6%
		(1.11)

2.7.3 Extension to Further Years

Finally, one strong piece of validity comes from extending the data year to 2012, much of which is being reported during the writing of this paper (and such data releases will occur up until the end of calendar year 2013). I manually collect 2012 data, including the portion of accounting, compensation, and other data that is missing from databases to construct a 2012 data set. I collect two datasets. The first consist of firms who are within 20 percentage points the 2011 vote cutoff for which full 2012 data is available. The second dataset consists of firms who are within 20 percentage points of the 2012 vote cutoff for which full 2012 data is available.

On this dataset, I run two RDDs. The first RDD looks at the same 2011 50% cutoff as the rest of the paper, but with the predicted variable being total CEO pay change between 2010 and 2012 (versus 2010 to 2011). As Table 2.7 Panel B show, the 2010 to 2012 total CEO pay change as a result of passing say-on-pay is 48% and statistically significant.

This provides evidence for the robustness of the 2010-2011 result, and also goes on to show that the pay results do not revert away. One concern with the pay drop at the 50% cutoff is whether the change in pay is temporarily put in place just to appease investors. Even worse would be the case where the board implicitly delays pay to give the impression of next year pay drop, while the NPV of pay stays in same. The extension of the result from one to two years shows no

such inclinations. The pay drop seems to stay in place, with minimal statistical evidence of further pay drops or reversions.

The second RDD on the new data set looks at a different set of firms. Instead of the firms around the same 2011 vote cutoff as the rest of the paper, the regression examines firms around the 2012 vote cutoff, generally a different set of firms. The predicted variable is the 2011 to 2012 change in total CEO pay. This is the analogue of our baseline 2011 regression shifted up by one year.

As Table 2.7 Panel B shows, the regression coefficient is significant at the 10% level with an estimated increase in total CEO pay at the cutoff of 44.8%. The low power may be attributed to the much smaller data-set, less than $1/10^{th}$ the full sample size due to limitations in the number of companies who have released 2012 data, especially in redistribution databases like EXECUCOMP.

In terms of size, 44.8% is in line with the baseline data. The ability to reproduce the result in future years provides external validation of the baseline result – especially since the 2012 data was created after the first draft of this paper and thus being truly out of sample.

Further, that the effect works in 2012 data shows that the effect is not limited to the first year of implementation. This may be thought to be the case since Dodd-Frank was only signed into law in the middle of 2010, and so the surprise of such a reinforcement system might be limited to the first year.

That the magnitude is lower than in the first year is in line with the informal theory that previous-year pays are already adjusted for company

expectations of say-on-pay vote fraction. For 2011, with more data, companies might make more accurate predictions of next year's vote fraction. This would reduce future year's pay changes, if pay changes are partially seen as a measure of censure surprise from say-on-pay. Thus the lower magnitude from 2012 data may have a common explanation from the non-monotonicity around the 50% cutoff.

2.7.4   Composition of Salary

Ample evidence above shows that say-on-pay has causal impact on Total CEO Pay. The total CEO pay clearly changes at the cutoff. However, what about the composition of CEO pay? It seems unlikely that composition would stay exactly the same. If that were the case, then salary, bonus, and options would all see the exact same 59.8% increase. There is much reason a priori to believe some pay terms (bonuses) are more flexible than others (salary).

On one front, salary, bonus, and options have different incentive properties. One might imagine the cut that the board implements as the vote fraction falls from 50.1% to 49.9%. To maintain the same power of incentive for CEOs to work, the board might shift the composition of pay from low-powered components like salary to high-powered components like options and bonuses. This would be in line with later results this paper, which show that pay cut from failing say-on-pay does not cause a drop in firm value or accounting measures: CEO pay cuts reduce average pay while maintaining incentives. This would be

the type of pay change advocated by Murphy (1999) and Bebchuk and Fried (2007).

However, such a pay shift is not observed in the data. The fall from 50.1% to 49.9% does not cause a drop in salary and an increase in high power incentives, as observable in Table 2.7 Panel A[8]. Instead, the effect is exactly the opposite: salary is the most stable, getting cut the least. Bonuses are second most stable getting cut more than salary but still less than proportionately. Options and other compensation are the least stable, getting cut more than proportionately. This does not align with the incentives theory: the highest-power components of pay are cut the most. As an econometric note, I calculate components of total pay as a fraction instead of a level because the division controls for large common noise in total pay to improve the power of the tests.

However, this pattern of pay cuts lines up well with another consideration. Salaries are often the least flexible portion of pay. It is set out beforehand in many contracts, or otherwise expected to match some historical amount. Also, it is precisely because salaries don't depend on the vagaries of performance or circumstance that it is the most secure. From an agency-tradeoff point of view, it is the portion of pay that is most valuable to the CEO for a fixed expected value (Gabaix and Landier 2008). Thus, as pay is cut, in order to meet a minimum threshold to keep a CEO onboard, it makes sense to shift pay into more stable

---

[8] The Table shows the transition in the opposite direction, from 49.9% to 50.1%. To match the verbal analysis to the table then simply take the negative sign.

portions like salary (Hölmstrom and Kaplan 2001, 2003) . In this sense, the shift from high power to low powered incentives indicates that maintaining the IR constraint takes precedence over maintaining the IC constraint. Conversely, portions of pay like bonuses, and especially options, are incredibly volatile. Not only do they respond to the CEOs effort, but they also vary based on items that are outside of the CEO's control like the decline of the industry, and they are highly positively correlated with other parts of a CEO's implicit wealth, like job prospects.

Thus, the risk-based explanation highlight the fact that for the same expected dollar, a fixed salary is worth a lot more to the CEO than as part of an options package. When pay is cut, to keep the CEO onboard, it is natural in this framework to cut the options portion. Combine this with the institutional observation that boards often have a lot more year-to-year discretion over non-salary portions, and this paper arrives at an explanation for why pay cuts come mostly from non-salary portions.

2.7.5   Announcement Date Returns and Other Company Financials

The previous section shows that say-on-pay vote fraction has high impact on pay. What is the impact of the vote fraction on other characteristics of the firm, either through the pay channel or directly?

The first variable to examine is the announcement day return from say-on-pay passing. After the meeting in which the say-on-pay vote is taken, a few days

later the result is announced via SEC Form 8-K. The announcement day return is defined as the stock return from the last close-of-market stock price before the announcement to the first close-of-market stock price after the announcement. The announcement day return generally spans a single market day, and so has low unconditional variance and high power.

Table 2.7 Panel D shows that there is no announcement day return to passing or failing say-on-pay. The 95% confidence bounds are quite narrow at spanning [-1.4%, 1.6%]. This is significantly different from the 2.8% Figure that Cuñat Gine Guadalupe (2012b) finds for the passage of a generic vote. This implies that even though CEOs are paid significantly less on failing say-on-pay, the market did not predict that the lowered pay would have a significant adverse impact on firm value. This suggests that CEOs at the margin are indeed overpaid if their pay can be cut without harming firm value. This provides evidence against a Murphy (1999) story of politically induced underpayment of CEO salaries. At the same time, the zero announcement day result also suggests that investors are voting optimally. A consistent positive or negative effect would prove that investors are either consistently over voting or under voting. A zero result suggests investor efficiency in voting.

Such an interpretation is borne out in examining the impact of the say-on-pay vote cutoff on other company accounting variables. Table 2.7 Panel D shows that the cutoff has no impact on other company accounting variables for the year forward as well. The insignificant announcement day stock reaction is correctly

predicting that real accounting variable changes in the next year will be insignificant as well. As total CEO pay is drastically reduced under the vote cutoff, this paper finds no evidence that at the same time of any reduction in dividends, probability of paying dividends, finance policy, or earnings.

In these results, it is important to note that earnings does seem marginally significant. Being below the cutoff reduces CEO pay while at the same time raising next year's earnings. Both failing say-on-pay and the subsequent pay cut are signals to the CEO that his or her job is at risk. One interpretation is that the CEO as a result works harder to raise real earnings. That would suggest that failing say-on-pay improves the value of the firm. However, this contradicts the tight bound around zero of the announcement date stock return. Likely an alternate story is happening: the CEO responds to this increased pressure through earnings management. He or she delays depreciation, channel stuffs, and employs other standard accounting tricks. The CEO has discretion on when to recognize bad news and may simply be delaying that.

Another important note is that company accounting variables are particularly noisy because they are measured over the course of a year. The low power is exacerbated by the RDD identification only occurring at the margin, reducing effective sample size even more. The lack of accounting variable significance should not be taken as strong proof that say-on-pay has no effect, but rather lack of proof of any effect. Announcement day returns suffer much less from this power issue due to its low noise to start out with.

Another possible analysis with this data is in the instrumental setting. If it can be assumed that say-on-pay failure operates solely through reducing CEO pay, then this can be seen as an exclusion restriction. Say-on-pay can then serve as an instrument in estimating the elasticity of firm characteristics with respect to CEO pay. Under this interpretation, no firm characteristics have significant elasticity with respect to pay, except earnings which has negative elasticity. There is substantial reason to believe that exclusion restriction might not hold. Primarily, changes in CEO pay over a year cannot be seen as equivalent to having a lower overall level of pay – the first lowers pay but also sends a signal to the CEO that his or her job is at risk. This confound muddles elasticity estimates.

2.8 Extensions

As this paper relies on recent data, there are many obvious extensions, including both the data and the content being tested.

2.8.1 Data Extensions

The baseline data set here consists of two years of data: 2010 and 2011. Because the regressions require a lagged year for changes, there is only one year's worth of regression data. The paper makes an attempt at extension into 2012 for three years of data (and two years' worth of regression data). The results of this selective extension can be seen in Table 2.7 Panel C. However, since say-on-pay votes will happen annually, even in the next few years, it is expected that the relative amount of data available for analysis will increase tremendously. The

standard error on total CEO one-year pay change rests at about 14-20% right now, even with two years of data. In another decade, the data set will have increased to more than 12 years, giving us much smaller standard errors of less than 5%. While this paper already shows great t-stats of above 4 for some tests, new data will rapidly increase the power of the test and identify the exact size of the total CEO pay effect. This observation is not a trivial one that applies to all papers: it applies to this one especially because of the low number of years of data it relies on, and the guarantee of more annual data for the foreseeable future.

With more power, a few extra questions might be answered. Is the pay effect different between small firms and large firm? Historical studies of pay show large differences in composition and nature of pay between the firm types (Murphy 1999). Further, many theories of increasing CEO pay depend crucially on firm size (Gabaix and Landier 2008). This question can be answered by running an interaction of the treatment variable with log firm market cap. Similar studies can be done on industry, value, and other covariates.

One arena that extra power would especially help on is identifying company characteristic changes. As mentioned before EBIT, dividend, and financing policy are noisy measures. The magnitudes estimated in Table 2.7 Panel C are not small economically: a -4.3 percentage point drop in firms paying dividends would be economically interesting, and more data would help identify that. Similarly, EBIT which seems marginally significant can either be confirmed or denied for sure. The identification of other firm characteristics seems like one

128

of the most interesting uses of the extra power from future years. Under the instrumental interpretation of these firm characteristics, one can finally answer the question of the causal elasticity of firm behavior based on pay change. How do CEOs run their firms differently when pay changes?

Another arena to be verified with more data is the permanence of the total CEO pay change on the vote cutoff. With that high initial magnitude, an interesting question is how quickly it reverts away. Is the pay just shifted to a later date, or is failing a say-on-pay vote a reminder to the board that the general level of their pay is much too high, and so when they cut pay, the cut is permanent? Evidence from two years of data shows no reversion in pay, suggesting the later explanation, but it would be useful to still confirm the fact.

Extensions of the impact on failing say-on-pay to multiple years are especially important because generally pay contracts are set well ahead of time. Consider the 2011 Dodd Frank say-on-pay vote: the board is technically asking for investor approval of the past year's (2010) pay, but because the compensation committee has already done its work, rarely is the pay backwards-modified even on heavily negative vote. As can be logically induced and also gleaned from reading the SEC DEF 14A, most boards take the 2011 feedback on 2010 pay into account when setting 2011 pay. Even then some boards note that their hands are tied to changing pay due to past contracts. Thus, it may take many years for the full impact of Dodd-Frank say-on-pay to be felt. Hence, multiple-year predictions will be able to tease this distinction out.

A similar case exists for looking at company accounting variables and stock returns. If markets are not efficient, announcement-day returns are not sufficient for capturing the value change. CEOs may act with a lag, so to see the real impact of a pay cut on company performance, it may take more years of data.

Regardless of the case, the primary extension of this paper would involve collecting more data as it arises, and then using that extra data to 1) obtain more precise estimates 2) obtain estimates of interaction terms and 3) obtain longer duration change estimates.

2.8.2 Method Extensions

On the content front, one main result of this is that Dodd-Frank say-on-pay votes have real impact on CEO pay. Do other mandates have such impact? For example, Dodd-Frank also has a provision that requires firms to hold votes 1, 2, or 3 years depending on the desire of the shareholders. Does that provision impact pay or company valuation? If one imagines that CEO pay is generally too high, then being randomized into the 1-year group might immediately reduce pay and increase stock value. Future votes mandated by law can be used by in a RDD setting for these purposes as well.

Also, another area that this paper attempts to delve into is the impact of pay on firm performance. To that end, other instruments can be discovered that affect pay. One might believe for say-on-pay that perhaps there might be random noise in which block voters (institutions) votes in a given year. Since most say-

130

on-pay failures are related to institutional voters (see Summary Statistics Table 2.1), and institutional voters often follow ISS recommendations, instruments might be found in the ISS recommendation procedure. For example, if the ISS itself uses scoring, then a RDD can be run on the ISS internal score.

Finally, more power can be obtained if approval of Dodd Frank say-on-pay requires a higher number of votes. Since the median approval rate is much higher than 50%, having an approval cutoff of say even 70% would substantially increase the test power. Of course, the vote fractions might endogenously shift if the cutoff changes. Investors might vote more favorably knowing that the cutoff is higher. As long as the compensation effect is not one-for-one, the data can obtain higher power.

2.9 Conclusion

Few studies in executive compensation and corporate governance have strong identification. This is a particular issue with CEO pay, which is often set far in advance and whose impact manifests over a larger number of years. This paper obtains identification on both topics by exploiting a novel variation: that from voter noise around the 50% cutoff on the Dodd-Frank Say-on-Pay vote.

The paper began by looking at non-identified OLS studies of CEO pay, CEO pay change, and say-on-pay vote fraction. These regressions show that reverse causality and omitted variables seriously confound the OLS analysis. At

131

the same time, the regressions revealed descriptive insights into what other variables commove with these three.

When a firm receives just slightly below 50% of votes versus above, that firm statutorily fails say-on-pay and experiences discretely increased censure and thus discretely decrease next year CEO pay. This paper first verifies that such an effect does occur by examining media announcements, SEC DEF 14A narratives, vote forms, SEC 8-K vote results, and next-year SEC DEF 14A commentary on past years votes.

This paper then exposits the theory of optimal-bandwidth kernel regression discontinuity design, and uses this design to estimate the magnitude of the decrease at 59.8%. This estimate causal, economically substantial, and is robust to variation of the kernel estimation bandwidth. The paper further validates this number using next-year data (2012), and by showing that this decrease is sustained for two years (2010-2012), the entire data span available as of this writing.

Like many regression discontinuity studies, the causal identification has excellent internal validity. However, there is no guarantee of external validity. The causal effect is completely estimated on firms that receive around a 50% vote fraction. These firms may have unique properties that are specific to them – like have overpaid CEOs whose salary can be cut without hurting the firm. In taking all causal inferences from this study, it is worthy to keep in mind the local average treatment effect nature of the estimate (Imbens and Angrist 1994).

132

This paper then looks at the percentage point of total CEO pay that consists of salary, bonus, and options. This paper deduces that the cut comes less from salary more from bonuses, and most from options. This is in line with a risk-based compensation theory where the most risky dollars are cut first due to their relatively lower value to the CEO. The data also lines up with the fact that boards often have more discretion on non-salary portions of pay.

Finally, this paper uses the RDD to estimate the impact of the cutoff on other firm characteristics. The paper finds no statistically significant impact on announcement day returns, EBIT, dividends, or company financing policy. Taken together, this suggests CEOs at the say-on-pay 50% margin could be overpaid, and their pay may be reduced without hurting performance.

Chapter 3

3 Do Options Impact the Stock Market?

3.1 Introduction

In frictionless settings, standard options are redundant securities and can be synthesized from dynamically trading equities. The seminal works of Black-Scholes (1973) and Merton (1973) (BSM) show that any derivative that depends in a general fashion on continuous equity prices can be reconstituted through a dynamic hedging strategy on that equity. Thus, not only are options spanned by the space of equities, but each option is spanned by its single underlying equity. Many follow-on works to BSM provide specific implementations of their strategy. In the frictionless BSM theoretical framework, whether options exist as independent products should not matter.

134

However, the real world deviates from the frictionless ideal. Trading is not costless and losses arise from transaction costs. Thus rebalancing is not continuous, and often occurs at the daily to weekly frequency, according to current practitioners. There is heterogeneity in firm skill in reducing these frictions, leading to specialization. Empirically, end-users are generally net demanders of both puts and calls (Garleanu Pedersen Poteshman 2009). Market makers that take the other side do not desire the bet but synthesize it at a premium (Bates 2003). This is consistent with a story where end-users are unskilled at dynamic hedging and market makers are skilled, leading market makers to specialize in using dynamic hedging to synthesize options that are then purchased by end end-users.

To fix ideas, let us sketch the basic mechanism. Consider the risk exposures of market makers. End-user demand of both puts and calls respectively generate positive and negative equity exposure (the option delta). Because end-users purchase both puts and calls, most of the delta exposure cancels out in large samples. To the extent delta risk is not cancelled out, market makers can always hedge options delta risk by purchasing or shorting in the underlying stock market.

However, both puts and calls are also long concavity exposure (the options gamma). When end-users demand puts and calls, they are then consistently long gamma risk exposure, and market makers are short gamma risk exposure. Unlike delta risk, the gamma risk of puts and calls do not cancel out, since both derivatives have positive gamma exposure. Also unlike delta risk, gamma risk

cannot be simply hedged with a single trade in the underlying stock: the underlying stock has zero concavity and so hedging does not change gamma at all. Since options market makers are stuck being short gamma, dynamic hedging forces market makers to have upward sloping demand curves. That is, as a stock's price rises, market makers must mechanically buy that stock; similarly, when a stock's price falls, market makers must sell that stock. Thus, the issuance of both puts and calls causes a feedback effect where quantity demanded moves in the same direction as the previous price move. We term this theoretical phenomenon *hedging feedback demand* in this paper.

A natural and readily available empirical measure of hedging feedback demand is the total amount of gamma outstanding on a given equity on a given day. This can be calculated from databases like OptionMetrics by summing up the gamma over all options contracts. While not all options are between end-users and market makers, Garleanu Pedersen Poteshman (2009) give evidence that total gamma is a good proxy for the degree to which market makers are dynamically hedging to produce options for end-users.

We measure the resulting impact on hedging feedback demand through momentum. Since hedging occurs in practice on a daily to weekly frequency, our methodology examines a stock move on day zero, and then measures momentum one to ten days out. The underlying concept is as follows. On day zero, suppose the price of the underlying stock exogenously increases. Because market makers have short gamma exposure, they have upward sloping demand curves with

respect to the underlying stock. Hence, when the price of the underlying stock increases, hedging feedback demand forces market makers to buy more of the stock over the next few days (the exact time frame depends on often they hedge). The positive quantity demand then causes positive price effects through price impact. This price effect gives rise to measurable momentum. In other words, since hedging feedback demand is not directly observable, we measure its effect in price space through momentum.

However, total gamma and momentum should not be linked one-to-one. For example, a fixed amount of total gamma might cause minimal momentum effects in a large-cap company, which is highly liquid and has a tremendous amount of baseline volume. That same amount gamma might cause high momentum in a micro-cap that is thinly traded. Some controls on total gamma are needed. Hedging feedback demand is more accurately represented by total gamma controlled for baselines like size, liquidity, and other factors to be discussed in Section V. We term the resulting controlled measure of gamma *Residual Gamma*, which is intended to map directly to momentum.

We indeed find then that hedging feedback demand impacts the underlying equity; specifically we find that residual gamma does indeed causes positive momentum effects. After a day-zero uptick, momentum is experienced within one to four days and does not revert away within ten days (see Table 3.4), the furthest horizon that we test due to statistical power. This effect is robust to liquidity effects and time trends (see Table 3.5 and Table 3.6).

Causal identification of this effect is possible using instruments related to options expiration. The general principle behind all instruments considered is that options are issued with certain idiosyncrasies. For example, options expire at an arbitrary part of the month, or options tend to accumulate around past price paths of the underlying equity, or option strike prices tend to be round numbers. These idiosyncrasies can be seen as exogenous drivers of Residual Gamma, leading to identification.

The first instrument is that Residual Gamma decreases dramatically at the same point in a monthly cycle, generally options expiration around the 18th date of each month. On one day, an entire month's worth of options disappears and market participants do not immediately compensate. This leads to a sharp drop in Residual Gamma near the 18th of each month, and then a buildup again as time passes until the next drop. This causes a noticeable cycle in Residual Gamma (see Figure 3.1). The advantage of this instrument is that it is powerful and consistent across all stocks, although a disadvantage of this instrument is that it might inadvertently capture any monthly cyclic variation in momentum.

The second novel instrument this paper exploits is the fact that options are generally issued with strikes around current equity prices. As equity prices change, option purchasers either do not unwind past positions or do so with substantial lag. As a result, the option issuance structure is heavily affected by prices as far as months back (Figure 3.2). A stock that drifts away substantially from past prices will have relatively low exogenous Residual Gamma, and a stock

that maintains prices will have relatively high Residual Gamma (Figure 3.3). One may be concerned about the interaction between price moves and volatility. A stock that has moved a lot recently likely also has experienced higher volatility. We address this using volatility controls.

Finally, perhaps the most novel of instruments for Residual Gamma utilizes the fact that options are struck around round numbers (Figure 3.4 and Figure 3.5). A few days before options expiration, the Residual Gamma caused by an option is very closely clustered around the strike price. Therefore, if the underlying equity price happens to be round, Residual Gamma will be unusually high, and if the equity price is not round, Residual Gamma will be low. This instrument has little potential confound and provides a novel way of identifying our result.

In this paper, we utilize all three instruments to provide estimates of the causal effect of Residual Gamma on momentum. We find that the three instruments agree generally with each other and with the OLS estimate. This suggests options creation and issuances causes extra hedging feedback demand to destabilize the underlying stock through additional momentum. The momentum effect occurs within one to four days and does not revert away. The effect is visible in the baseline OLS regressions, as well as OLS regressions with controls. The magnitudes agree across three very different instruments with minimal confounds, and almost no common confounds. This cross verification provides

139

strong evidence for the existence and causality of this relationship between Residual Gamma and momentum.

This paper contributes 1) by expositing the feedback effect of options on equity prices, and 2) by presenting novel instruments for options that can be used in other settings as well.

Section II briefly reviews motivating literature and related research. Section III describes hedging feedback theory, aggregation amongst different players in the market, and the resulting price impact. Section IV describes the data collection procedure. Section V discusses the main OLS results. Section VI discusses the instrument regressions, a large contribution of our paper. Section VII discusses extensions and interpretations of the data. Section VIII concludes.
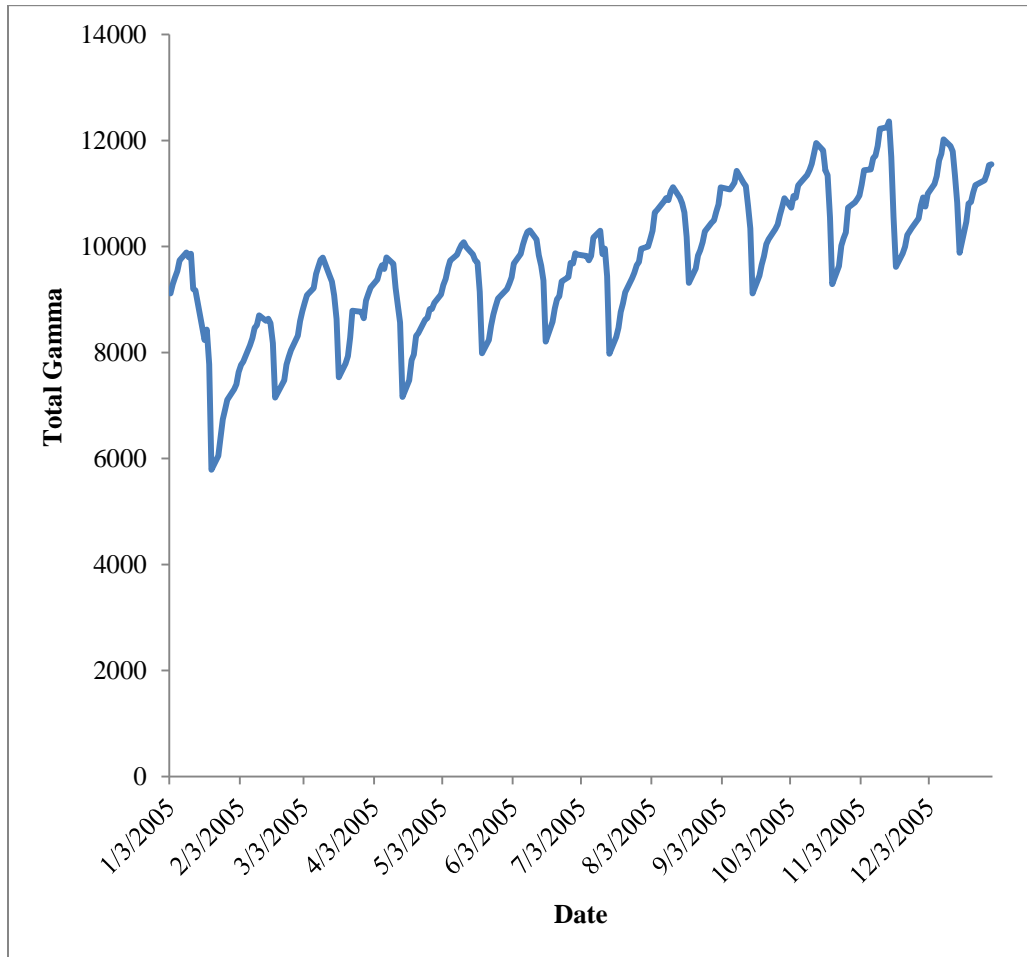
Figure 3.1: Instrument #1 (Time to Expiration): Cyclical Variation due to Options Expiration.

Gamma summed over all stocks and equities for a given date over an illustrative period in 2005. The figure shows monthly cyclicality of Gamma caused by options expiration. Gamma is aggregated by taking the sum of open interest for an option contract multiplied by the gamma specific to that contract calculated via OptionMetrics using its internal Cox Ross Rubinstein (1979) binomial tree model.
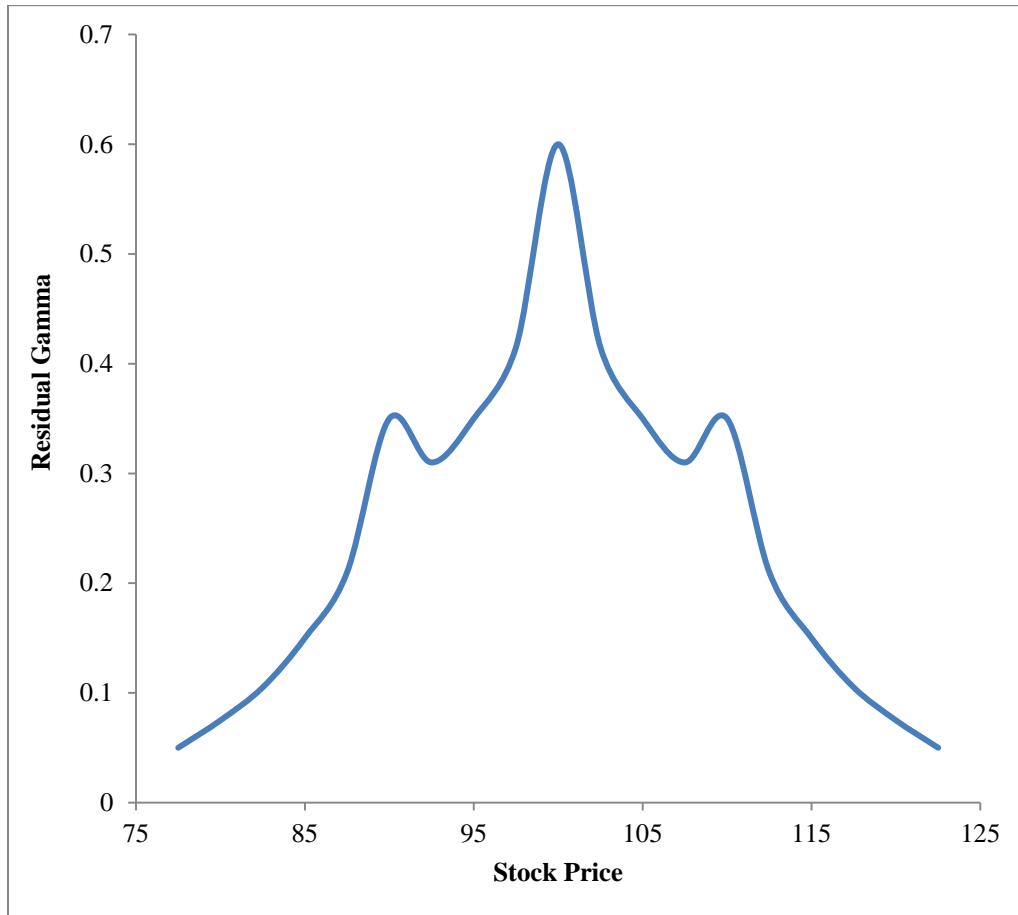
Figure 3.2: Instrument #2 (Past Price Path): Theoretical Illustration of Gamma For An Equity as a Function of Stock Price.

The following graph illustrates how historical prices affect Gamma. In the case below, the stock is assumed to have maintained a price of around 100 for the past many weeks. Thus options are generally issued around the historical stock price of 100, leading to a tent shape in Gamma as a function of stock price. If prices move significantly, to say 80, Gamma will drop sharply. Gamma is scaled to an arbitrary normalizing constant for illustrative purposes. Humps at 90, 100, and 110 simulate empirical observations of excess Gamma near round numbers (Instrument #3).

Figure 3.3: Instrument #2 (Past Price Path): Total Gamma as a Function of 20-Day Price Move on SPX.

Total Gamma as a function of the last 20 days price move on the SPX (S&P 500 index). A red optimal bandwidth (0.8) nonparameteric lowess regression line is fit through the data sample. The regression line is tent shaped, being linear on both sides of zero with similar magnitude of slope. Price move is expressed in dollars, with the S&P 500 index valued in the thousand dollar range.

143
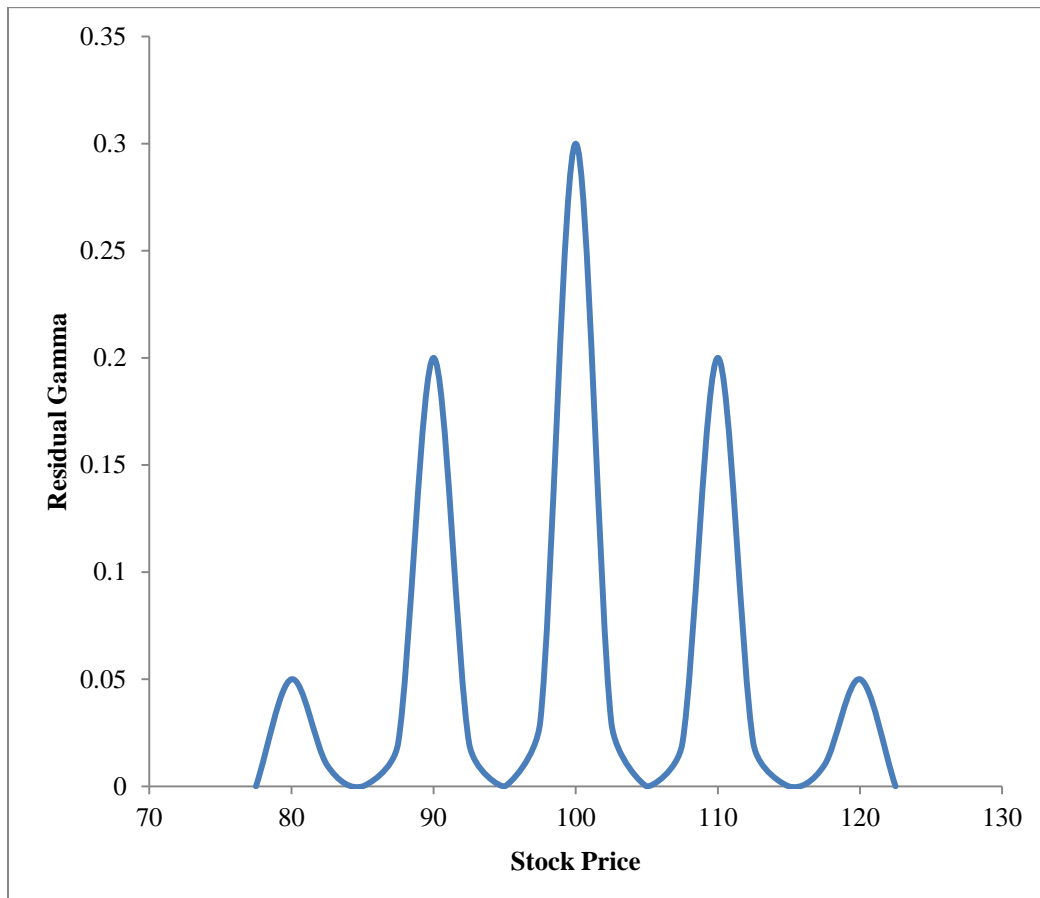
Figure 3.4: Instrument #3 (Distance to Round Numbers): Residual Gamma as a Function of Stock Price, for Options Close to Expiration.

The following graph plots theoretical gamma generated by near-expiry options for a given equity as a function of underlying stock price. The graph assumes recent prices near 100, and an exchange which mechanically only allows options to be struck near round prices. Observe that Gamma has a regular pattern that peaks at regular points. Price is expressed in dollars, and Gamma is normalized against an arbitrary constant for illustrative purposes.
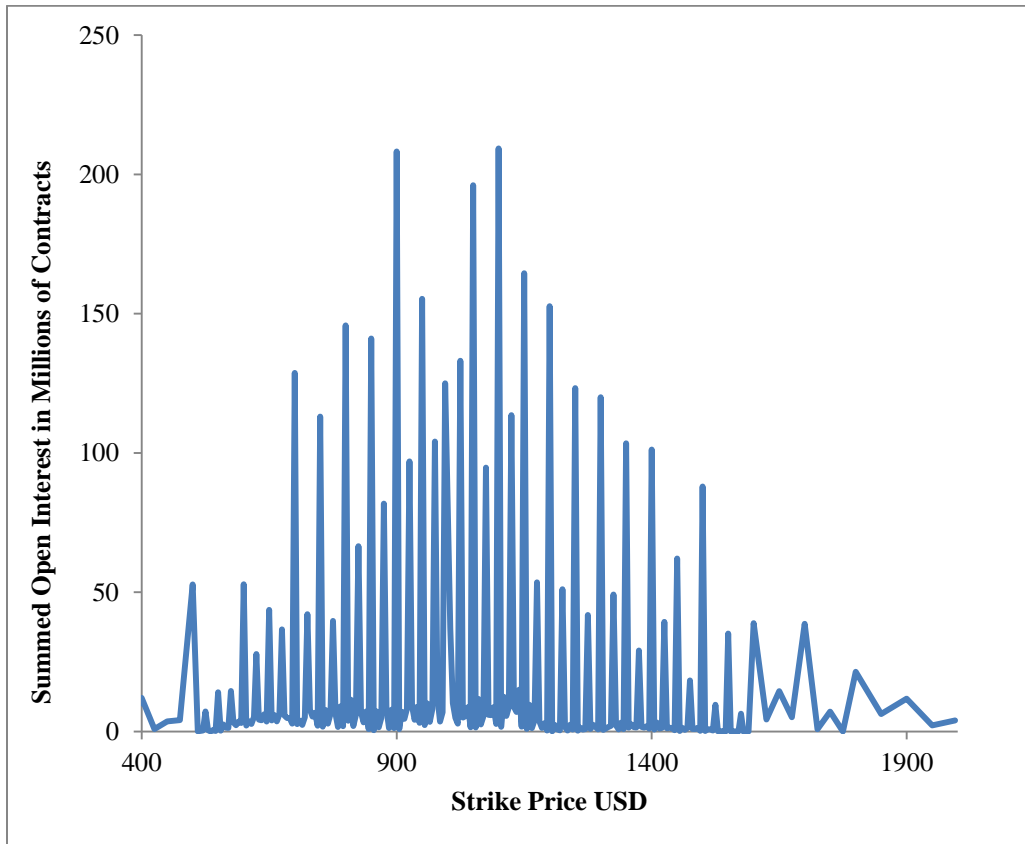
Figure 3.5: Instrument #3 (Distance to Round Numbers): Open Interest as Function of Strike Price on SPX.

Aggregate open interest in S&P 500 options for all days in our sample period between 1996 and 2013 as a function of that option's strike price. Prices like 1000, 900, 1100 have the most open interest, followed by less round numbers like 950. Non-round prices like 982 see very few options

3.2 Literature Review

The seminal work in options pricing is due to Black and Scholes (1973) and Merton (1973) (BSM). They show that any option on an underlying instrument is essentially made redundant by the underlying instrument as long as the underlying instrument can be traded continuously. Their redundancy theory holds in the first order: options traders today use the equivalency theories and formulas to value options and to replicate options.   Many works offer a complete treatment of the subject – for an example see the relevant section in Campbell, Lo Mackinlay (1997). Substantial work has gone into providing implementation details in various situations.

As much as Black Scholes Merton (BSM) is a standard, if the equivalency is taken to the extreme, the implication would be that options needed not exist at all. That options do exist and in fact are actively traded demonstrates that the second order effects of frictions, jumps, and non-equivalencies are important in options trading.

Bates (2003) shows that there are numerous options pricing anomalies with respect to BSM. In line with our paper, a large literature recognizes option pricing anomalies as the interaction between end-user demand for options and market makers. In particular, index options appear to be more expensive than predicted by BSM due to especially high end-user demand; low moneyness options also are more expensive than predicted due to high demand by end-users

146

who use these options for leverage (Rubinstein 1994; Longstaff 1995; Bates 2000; Jackwerth 2000; Coval and Shumway 2001; Bondarenko 2003; Amin, Coval and Seyhun 2004; Driessen and Maenhout 2008).

In the options literature, a paper that is close in topic to us is Garleanu, Pedersen, and Poteshman (2009). They give evidence that demand pressure from end-users affects options prices. In particular, they develop a model where options have risk factors that are unhedgeable in the underlying security, and these risk factors are consistently priced across the options space. Thus, high end user demand for a certain option will increase its price above the BSM standard, to compensate market-makers for bearing more risk.

The idea of demand pressure causing price effects is more general and includes stock index additions (Shleifer 1986, Wurgler and Zhuravskaya 2002, Greenwood 2005), mortgage-backed securities (Gabaix, Krishnamurthy, Vigneron 2007), option end-users (Stein 1989; Poteshman 2011), household risky asset holdings (Zhang 2013), and bonds (Greenwood and Vayanos 2009). Evidence of supply side catering is well supported in the literature (Baker Wurgler 2000, Baker, Greenwood Wurgler 2003, Greenwood Hanson Stein 2010). That sentiment plays a role in such demand has been studied both theoretically and empirically (Baker Wurgler 2006, Baker Stein 2004).

In terms of mechanism and style, this paper is closest to the strand of literature exposited by Cheng and Madhaven (2009) although they do not use instrumental variables. They show the analogue of our result in the leveraged

ETF industry. In particular, Cheng and Madhaven (2009) demonstrate levered or inverse ETFs must purchase the underlying when prices go up mechanically due to their dynamic hedging obligations, and vice versa. They argue that the dynamic hedging of ETFs causes destabilization by both increasing volatility of the underlying and increasing the amount of price impact and momentum experienced by trades in the market. This paper also argues a view of mechanical hedging feedback demand but applied to options instead of ETFs. In particular, we test momentum caused by the Residual Gamma of options.

To quantify the level of destabilization though momentum, this paper utilizes previous technical work in the momentum and price impact literature. Ferraris (2008) gives industry models of permanent price impact as stocks are traded. We account for baseline momentum as well: Jegadeesh and Titman (2001) show that there is short term reversion on the one month span. Asness (2008) shows that there is long run momentum on the one month to seven month span.

As far as we know, this paper is one of the first treatments of the feedback effects of BSM hedging of options back onto the price patterns of the underlying stock.

3.3 Theory

This paper examines the frictional effects of the implementation of BSM through dynamic trading. To start, it is natural to build a theory of BSM dynamic

trading. We first analyze the trading needed to hedge out a single options contract, then we aggregate amongst players in the market, and finally we motivate an empirical test of hedging through price effects.

3.3.1 Risk Properties of a Single Contract

First, for a fixed equity, consider the value of an option with strike K when the stock price is S, with C a binary indicator for whether the option is a call, $V_{C,K}(S)$. Assume for now that K is close to S and that the option is a call option. An option generally depends on many other parameters such as volatility, time to expiration, and risk-free rate. However, for simplicity of exposition, and because only a few parameters are key to the theory, we fix all other parameters as constant in the comparisons below.

BSM and related theories give us the following two properties for this call option:

$$\Delta := \frac{\partial V}{\partial S} > 0$$

$$\Gamma := \frac{\partial^2 V}{\partial S^2} > 0$$

Mechanically, this means that the value of the call option is increasing and concave in the stock price. Intuitively, this means 1) a call is worth more as the price goes up and 2) the degree of relationship between a call's price and the stock's price increases as the price goes up. These two properties hold under broad theoretical conditions and are almost never violated empirically.

Suppose a market maker sells such a contract while the underlying price is S_0. Then, the market maker's exposure can be approximated by the following second-order Taylor series:

$$-V_{C,K}(S) \approx -(S - S_0)\Delta(S_0) - \frac{(S - S_0)^2}{2}\Gamma(S_0)$$

Now let two periods occur. In the first period, the market maker may hedge by buying/shorting α contracts of the underlying. After the hedging, the price jumps from S_0 to S_1. The jump takes negligible time, thus not affecting option valuation through the time-to-expiration channel.[9] In the second period, the market maker may again hedge α', and then the price moves again.

The market maker incurs exposure in the two periods respectively of:

$$-V_{C,K}(S) + \alpha S \approx (S_1 - S_0)(-\Delta(S_0) + \alpha) - \frac{(S_1 - S_0)^2}{2}\Gamma(S_0)$$

$$-V_{C,K}(S) + \alpha'S \approx (S - S_1)(-\Delta(S_1) + \alpha') - \frac{(S - S_1)^2}{2}\Gamma(S_1)$$

For all utility functions that are risk averse everywhere (including CRRA and CARA utility) the optimal hedging in both periods is: $\alpha^* = \Delta(S_0)$, $\alpha'^* = \Delta(S_1)$. The quantity the market maker must purchase between the two periods to minimize risk of holding the option is $\alpha'^* - \alpha^* = \big(\Delta(S_1) - \Delta(S_0)\big) \sim \Gamma(S_0)(S_1 - S_0)$. Thus, the key quantity to recognize is that the quantity of hedging feedback demand is $\Gamma(S_0)(S_1 - S_0)$.

---

[9] For a realistic calibration of timescales, the period between S_0 and S_1 might be a day, while expiration might be three months out.

3.3.2 Aggregation Properties

Now let us consider the above for not just one single contract, but for q call contracts issued across strike prices K of both puts and calls C. A representative end-user demands $OI_{C,K}$ contracts of the option with strike price K, of call type C which is in zero net supply. A market maker fills the demand by selling $OI_{C,K}$ contracts. Now the market maker is short exposure by $-\sum_k OI_{C,K}\, \Delta_{C,K}(S)$. Since market makers must hedge out their linear exposure, the market maker must buy $\sum_k OI_{C,K}\, \Delta_{C,K}(S)$ shares of the underlying stock.

Now we arrive at the main feedback trading aggregation. As prices increase from S_0 to S_1, the market maker must purchase shares of underlying equity equal to:

$$\sum_k OI_{C,K}\, \Delta_{C,K}(S_1) - OI_{C,K}\, \Delta_{C,K}(S_0)$$

$$= (S_1 - S_0) \cdot \sum_k OI_{C,K}\Gamma_{C,K}$$

$$= (S_1 - S_0) \cdot G_0$$

Where $G_0 \coloneqq \sum_k OI_{C,K}\Gamma_{C,K}$ can be interpreted as total gamma outstanding on an equity. In other words, as more options are issued by market makers across all puts, strikes, and expirations, G_0 the total gamma of the option system increases. G_0 can be seen as a measure of magnitude of hedging feedback

demand, which is also the coefficient of the quantity of feedback trading as a

result of a price move.

3.3.3   Price Impact of Feedback Trading

As we see above, a dollar price change of $S_1 - S_0$ causes approximately a

mechanical quantity feedback demand of $(S_1 - S_0) \cdot G_0$. The impact of this

feedback quantity in how much it destabilizes prices should naturally be measured

in price space. Thus we need to translate quantities into prices. There are a

number of possible measures such as increased volatility, increased bid ask

spreads, decreased liquidity, and others as exposited in Madhaven and Cheng

(2009).

We choose here a very natural measure of feedback impact: price

momentum. We imagine the mechanism as follows: on day zero, price

exogenously increases. Feedback hedging causes positive quantity demand on

day one, which also pushes up the price on day one. Thus higher G_0 causes more

feedback hedging which causes more momentum.

Our method to map quantities to price changes follows the market

microstructure literature. G_0 reflects the absolute quantity of feedback trading in

the system. To translate this price into price impact, it is necessary to use a price

model. This paper follows the price model posited by Ferrais (2008) and divides

total gamma by trading volume to reach a normalized measure of gamma: $\frac{G_0}{Volume}$.

The division by volume is the result of the intuition that the same size trade will

152

cause much stronger movements in a thinly traded stock than a thickly traded stock. We take logs and define Residual Gamma = Log Gamma – Log Volume. This definition of Residual Gamma says that total gamma must be controlled by the liquidity of the trading volume to arrive at a measure of impact in price space.

Residual Gamma then can be directly mapped into price space:

$$S_2 - S_1 = \lambda \frac{\left(\sum_{k} OI_{C,K}\, \Delta_{C,K}(S_1) - OI_{C,K}\, \Delta_{C,K}(S_0)\right) \cdot S_0}{Volume\ Traded} \quad (1)$$

$$\approx \lambda R_1 \cdot Residual\ Gamma \quad (2)$$

λ here is a normalizing constant and R_1 is first period returns. While many price models treat λ as constant we explore the possibility that λ itself depends on other factors like size through third order interactions in Table 3.3. Residual Gamma can be calculated theoretically as well as empirically and the two match closely as seen in the Section V.

Table 3.1: Summary Statistics for Daily Level Equity Information.

The summary statistics below describe all dates and equities for which OptionMetrics price data exists as of construction of the dataset: from the start in 1996 to January 31, 2013. Each of the 11,266,998 observations is an equity-date pair and reflects the status of that equity at the end of that trading day. All options data is collected from OptionMetrics, which is originally in option-date pairs and has been summed into equity-date pairs: Open Interest of all options (OPEN_INTEREST), Gamma times Open Interest (GAMMA * OPEN INTEREST). All equity data is collected from CRSP: Return from holding the stock between two closes of trading days (RET), Close Price at end of trading day (PRC), Volume of shares traded that day (VOL), and Shares Outstanding at the end of that day (SHROUT). The remainder of the variables are computed as follows: Date Serial Number starts at one for the first date data exists, and increments by one every trading day; Dollar Volume is Close Price times Volume; Market Cap is Close Price times shares outstanding; Log Dollar Volume is the natural log of Dollar Volume; Log Market Cap is the natural log of Market Cap; Total Open Interest is the total number of open interests for all options across all equities for that day; Log Total Open Interest is the natural log of Total Open Interest; Day of Month is a integer from 1 to 31 of the day of that month; Last 20 Day Return is the sum of the 1st to 20th lags of Return; Last 20 Day Volatility is the square root of the sum of squares of the 1st to 20th lags of Return.

Variable	N	Mean	SD	Min	Max
Date Serial Number	11,266,998	2,349.739	1,247.693	1.000	4,278.000
Open Interest (thousands)	11,266,998	51.827	336.695	.000	30,800.000
Gamma * Open Interest (thousands)	11,266,998	2.801	25.988	.000	6,301.253
Return (%)	11,263,869	.044	3.712	-94.891	625.925
Close Price	11,264,218	29.166	29.317	.000	3540.000
Volume (millions)	11,264,215	1.586	8.1446	.000	1897.900
Shares Outstanding (millions)	11,266,778	157.239	494.596	.000	29,206.400
Dollar Volume ($millions)	11,264,214	50.988	384.227	.000	93,188.090
Log Dollar Volume	11,260,436	15.872	1.972	4.388	25.257
Market Cap ($billions)	11,264,218	5.340	18.613	.000	658.153
Log Market Cap	11,260,440	13.978	1.659	4.812	20.305
Total Open Interest (millions)	11,266,998	170.475	105.719	.000	372.427
Log Total Open Interest (millions)	11,259,870	18.647	.901	15.937	19.735
Day of Month	11,266,998	15.744	8.749	1.000	31.000
Last 20 Day Return (%)	11,013,012	.903	15.774	-306.419	701.251
Last 20 Day Volatility (%)	11,013,012	13.213	9.854	.000	627.684

Table 3.2: Baseline Determinants of Log Gamma.

OLS regression of Log Gamma, defined as the log of the sum of all gamma times open interest across all options for an equity-date pair, on Log Volume traded for that equity-date, Log Shares Outstanding for that equity on that day, the Price for the equity at the end of the trading day, the Log NASDAQ Volume for that equity traded on NASDAQ on that day, the Time Serial an integer that starts with one and is incremented by one for each trading day, Log Open Interest the sum across all options for that equity on that day, and Log Total Open Interest the total open interest across all equities and options on that day.

$$Log\ Gamma_{i,t} = \alpha + \beta_1 Log\ Volume + \beta_2 Log\ Shares\ Outstanding + \beta_3 Price$$
$$+ \beta_4 Log\ NASDAQ\ Volume + \beta_5 Time\ Serial + \beta_6 Log\ Open\ Interest$$
$$+ \beta_7 Log\ Total\ Open\ Interest + \epsilon_{i,t}$$

All logs are natural logs. t-stats are Eicker-Huber-White heteroskedasticity robust. The residual from regression (3) is defined as Residual Gamma for the remainder of the paper.

Panel A: Regressions (1) and (2)

Variable	(1)		(2)	
N	10,834,255		10,834,255	
R^2	.5053		.5282	
	b	[t]	b	[t]
Log Volume	.951158	(3086.7)	.743272	(1701.4)
Log Shares Outstanding			.363015	(679.9)
Price				
Log NASDAQ Volume				
Time Serial				
Log Total Open Interest				

Panel B: Regressions (3) and (4)

Variable	(3)		(4)	
N	10,834,211		10,834,211	
R^2	.5347		.5403	
	b	[t]	b	[t]
Log Volume	.755658	(1734.0)	.714132	(1504.5)
Log Shares Outstanding	.365671	(679.9)	.391373	(691.4)
Price	-.203234	(-353.2)	-.185148	(-315.0)
Log NASDAQ Volume			.006525	(49.3)
Time Serial			-.000244	(-191.5)
Log Total Open Interest			.468826	(273.7)

Table 3.3: The Impact of Residual Gamma on Momentum with Market Cap Interaction.

OLS regressions of equity-date Returns on 10 lags of: Returns (R) expressed as a net fraction; Residual Gamma defined in Table 3.2, shares outstanding, or volume; Demeaned Log Market Cap; and all higher order interactions of the previous variables. All variables besides Residual Gamma are defined as in summary statistic Table 3.1. Residual Gamma is defined as in Table 3.2. N = 10,531,923. R^2 =0.0016. T-stats are clustered by dates. Coefficients are multiplied by 100.

$$R_{i,t} = \alpha + \Sigma_{k=1}^K R_{i,t-k}\big(\beta_{1,k} + \lambda_{1,k}Residual\ Gamma_{i,t-k}$$
$$+ \lambda_{2,k}Residual\ Gamma_{i,t-k} \cdot Demeaned\ Log\ Market\ Cap_{i,t-k}$$
$$+ \beta_{2,k}Demeaned\ Log\ Market\ Cap_{i,t-k}\big)$$
$$+ \Sigma_{k=1}^K\big(\gamma_{1,k}Residual\ Gamma_{i,t-k}$$
$$+ \gamma_{2,k}Demeaned\ Log\ Market\ Cap_{i,t-k} + \gamma_{3,k}Residual\ Gamma_{i,t-k}$$
$$\cdot Demeaned\ Log\ Market\ Cap_{i,t-k}\big) + \epsilon_{i,t}$$

Panel A

Date	Return (R)		R * Residual Gamma		R * Residual Gamma * Demeaned Log Market Cap		R * Demeaned Log Market Cap	
	$100\beta_{1,k}$	$t(\beta_{1,k})$	$100\lambda_{1,k}$	$t(\lambda_{1,k})$	$100\lambda_{2,k}$	$t(\lambda_{2,k})$	$100\beta_{2,k}$	$t(\beta_{2,k})$
t - 1	0.606	(0.90)	0.787	(5.00)	0.042	(0.62)	-0.237	(-1.99)
t - 2	-1.592	(-1.81)	0.342	(2.07)	0.112	(1.65)	-0.514	(-3.71)
t - 3	0.196	(0.27)	-0.071	(-0.47)	-0.078	(-1.24)	-0.102	(-0.84)
t - 4	-0.087	(-0.11)	0.109	(0.71)	0.043	(0.68)	-0.157	(-1.27)
t - 5	-1.227	(-1.55)	0.161	(1.03)	0.037	(0.58)	-0.190	(-1.36)
t - 6	-0.651	(-0.84)	-0.128	(-0.8)	0.011	(0.17)	-0.281	(-2.07)
t - 7	-0.124	(-0.15)	-0.047	(-0.32)	-0.051	(-0.82)	-0.093	(-0.63)
t - 8	-0.061	(-0.08)	-0.184	(-1.22)	-0.114	(-1.85)	-0.075	(-0.56)
t - 9	0.097	(0.13)	0.087	(0.59)	0.002	(0.03)	0.047	(0.32)
t - 10	0.703	(0.90)	0.053	(0.37)	0.038	(0.63)	0.072	(0.48)

Panel B

Date	Residual Gamma		Demeaned Log Market Cap		Residual Gamma * Demeaned Log Market Cap	
	$100\gamma_1$	$t(\gamma_{1,k})$	$100\gamma_{2,k}$	$t(\gamma_{2,k})$	$100\gamma_{3,k}$	$t(\gamma_{2,k})$
t - 1	-0.070	(-3.57)	-0.574	(-1.86)	0.025	(5.34)
t - 2	0.030	(1.61)	0.374	(0.99)	-0.005	(-0.98)
t - 3	0.015	(0.78)	-0.155	(-0.49)	-0.003	(-0.56)
t - 4	0.002	(0.12)	0.090	(0.29)	-0.002	(-0.34)
t - 5	-0.015	(-0.84)	0.034	(0.11)	-0.006	(-1.35)
t - 6	0.031	(1.57)	0.023	(0.08)	0.005	(0.99)
t - 7	-0.016	(-0.89)	-0.053	(-0.18)	-0.009	(-1.86)
t - 8	0.008	(0.42)	0.164	(0.47)	0.002	(0.40)
t - 9	-0.028	(-1.6)	-0.093	(-0.26)	-0.003	(-0.6)
t - 10	0.024	(1.43)	0.185	(0.68)	0.004	(0.89)

3.4　Data

This paper uses panel data on two levels, the first of which is nested in the second. On the first level, an observation is an option-date pair. For example an observation might be the Exxon Mobile (XOM) call contract with strike of $42.50 and exercise date of 1/22/2005 as observed on 11/17/2004. First-level data is obtained from OptionMetrics IvyDB for all options in the database. This includes all observations from the start of the database in 1996 to the writing of this paper in 2013. The first-level data we process is 90 gigabytes in size and contains 898,578,403 records. We consider the comprehensiveness of our study, and the implicit computational non-trivialities and hurdles in processing such a dataset to be a second-order contribution of the paper.

Characteristic information of first-level data includes observation date, the underlying stock's CUSIP, the strike price, the exercise date, and a designator for whether the option is a call or a put. Substantive first-level data for each option-date includes open interest, delta, and gamma. The delta, gamma, and other greeks are calculated by OptionMetrics using a Cox Ross Rubinstein (1979) binomial tree method.

Again, processing and running regressions on such a large dataset on present-day commodity computational hardware requires substantial manual programming of many otherwise standard regression functions. For example, directly estimating fixed effects and per-equity coefficients on the 8108 equities

157

present in our dataset would normally require terabytes of memory, thousands of times more than readily available on machines today. Therefore, manual functions are often needed to improve computational tractability, and even then many subroutines have runtimes of days on a modern desktop computer. Despite these substantial difficulties, we believe having comprehensive data is critically important. It minimizes external validity concerns that often occur in studies that examine only a small fraction of the most liquid equities. It gives results that are true across the entire population of equities. It also maximizes the power of our tests. Therefore the comprehensiveness of our data processing can be seen as an auxiliary contribution of the paper.

On the second-level of data, an observation is an equity-day pair. Across the 17 years, our dataset of 898,578,403 option-level observations rolls up into 11,266,998 equity-level observations. These 11,266,998 equity-level observations are distributed across 8108 equities. A sample observation analogue to our example before would be Exxon Mobile (XOM) observed on 11/17/2004. This paper will mainly work on the equity-date level because both the main predictor variable (Residual Gamma) and the main predicted variable (equity level momentum) exist on this level. While first-level data might be useful for revalidating some instruments like distance to round numbers (our third instrument), we mainly work on the second-level even with these instruments because we believe the intuition, validation, and instrument first-stages are much clearer on this level.

We aggregate first-level data into second-level data through aggregating an empirical total gamma G_0, from now on known simply as the *Total Gamma* associated with an equity-date:

$$G_0 := \Sigma_{C,K} OI_{C,K} \Gamma_{C,K}$$

The right hand side variables of open interest and gamma are available from OptionMetrics IvyDB. This differs from the theoretical total gamma in Section III in that not every contract is necessarily between a market maker and end end-user, but Garleanu Petersen Poteshman (2009) give evidence that the two are highly correlated.

This paper obtains other equity level data directly from CRSP. From this database, the primary key is a CUSIP and date pair, and data includes open price, close price, volume, and holding period return – a value which includes splits, dividends and other adjustment factors. For summary statistics, please refer to Table 3.1.

3.5 Main Empirical Results.

We proceed in our empirical analysis in the following steps: (1) we calculate Residual Gamma from theoretical and empirical models. (2) we compare baseline momentum of the entire sample and subsamples conditional on Residual Gamma. This gives us straightforward results on the impact of Residual Gamma on momentum. (3) we formalize this by running an OLS interaction regression of the impact of Residual Gamma on momentum. We discuss

dynamics and magnitude of the price impact. (4) we discuss possible confounds, and show the impact is robust to specification and a battery of liquidity and trend controls.

3.5.1   Defining Residual Gamma: theoretically and empirically

As described in section III, absolute total gamma should not be used directly as a predictor of momentum: equities with larger size and liquidity will respond less in momentum space to the same amount of total gamma. To get at a sense of market impact on the scale of price returns, it is important to normalize by a factor of price impact, namely volume. Ferrais (2008) suggests using $\frac{Gamma}{Volume}$ as a measure, or in natural log form, $Log\ Gamma - Log\ Volume$. The Residual Gamma defined this way would be the theoretical approach answer to the problem.

A more empirical approach to the problem is to consider total gamma as factor that increases momentum but only in excess of a natural baseline. For example a firm that is ten times larger might mechanically have ten times more total gamma. To reduce such issues, comparisons ought to be made for gamma differences only on comparable firms – firms with similar size, volume, and other characteristics. The residual from such a regression could be defined as the Residual Gamma that would predict increased momentum.

It turns out that the theoretical and empirical approaches above give nearly identical answers, as Table 3.2 shows. Running a regression of Log Gamma on Log Volume (Table 3.2 Regression 1) gives a coefficient of near unity with an R^2

of .5053. The Residual Gamma from this definition is $Log\ Gamma - .951 \cdot$ $Log\ Volume$, almost identical to the theoretical value.

It is of course possible to add on further predictors. Price and size as measured by shares outstanding is a popular parameter in price impact models (Table 3.2, Regressions 2 and 3). Log Nasdaq Volume, Time Serial, and Log Total Open Interest can be added as well, as seen in Table 3.2 Regression 4. It turns out however that these additional covariate models have essentially the same R^2 as the baseline regression: .52-.54 for regressions (2), (3), and (4) vs .50 for the baseline regression (1). Thus the Residual Gamma calculated from all these regressions would be nearly identical as measured by correlation. We arrive at the remarkable result that the theoretical formula for Residual Gamma is nearly the same as the empirical formulae for Residual Gamma.

The standard price impact models generally include volume, shares outstanding, and price, thus matching up most with regression (3) of Table 3.2. While all models give nearly identical measure of Residual Gamma, we choose Table 3.2 regression (3) as the standard definition for the remainder of the paper. However, the results are not significantly changed by any of the definitions above.

3.5.2   Subsample Regressions

The fundamental empirical question of this paper is how Residual Gamma affects momentum. First, consider the simplest way to measure momentum, serial correlation. The estimation equation for this might looks like $R_{i,t} = \alpha +$

161

$\beta R_{i,t-1} + \epsilon_{i,t}$ with β being the baseline amount of momentum. We can generalize to k periods with $R_{i,t} = \alpha + \sum_{k=1,K} \beta_k R_{i,t-k} + \epsilon_{i,t}$ and observe the cumulative momentum through the cumulative coefficients.

As a baseline, equity returns are known to exhibit some serial correlation in returns. In fact, it is well known that on the order of a few days, there is generally reversion (Froot and Perold 1995). This reversion can be caused by bid-ask bounce (McInish and Wood 1992) amongst other factors. Figure 3.6 confirms this stylized fact, which shows the average momentum of the entire sample (solid line) to be slightly negative.

As a first look at the impact of Residual Gamma on momentum, we simply run the above analysis on two subsamples. The first subsample consists of the top ten percentile of Residual Gamma. The momentum on this sample is shown to be significantly positive in Figure 3.6 (green line). The second subsample consists of the bottom ten percentile of Residual Gamma. The momentum on this sample is shown to be more negative than average in Figure 3.6 (red line). Both subsamples have their means outside the 95% confidence interval of the sample average.

Thus, the rough glance provided by Figure 3.6 suggests a generally positive relationship between Residual Gamma and momentum.

162

Cumulative Momentum in Equities
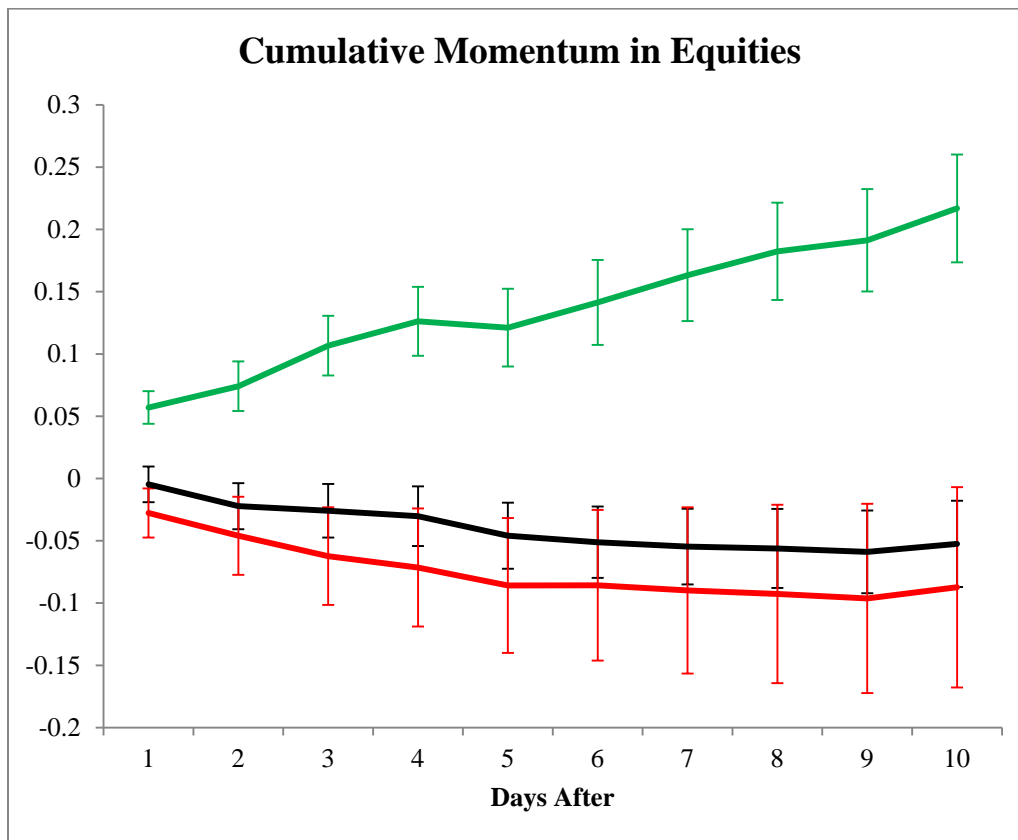
Figure 3.6: Cumulative Momentum in Equities and Percentiles of Gamma.

This figure shows the cumulative momentum from 1 to 10 days after a unit move in equity return. 95% confidence intervals are given by error bars. Momentum is calculated via the OLS regression clustered by date:

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} \beta_k R_{i,t-k} + \epsilon_{i,t}$$

The graph depicts three samples:

- All equity-date pairs in the sample (black line)
- Top ten percentile of Residual Gamma (green line)
- Bottom ten percentile of Residual Gamma (red line).

Residual Gamma is defined in Table 3.2 as the component of total gamma not predictable by price, shares outstanding, or volume. Cumulative Momentum is cumulative sum of the β_k with inferences accounting for covariances.

163

3.5.3   OLS Interaction Regression

Instead of just looking at momentum in subsamples, we can be more
rigorous and quantify the effect in a regression setting. There are multiple natural
ways to accomplish this. If identification comes completely from the cross
section, then one strategy is a Fama-MacBeth estimation in which β_i from
$R_{i,t} = \alpha + \beta_i R_{i,t-1} + \epsilon_{i,t}$ is estimated for each equity, and then a second cross
regression $\beta_i = a + \lambda G_i + e_i$ can be run. This paper does not use the Fama-
Macbeth strategy due to a desire to identify off of both time series and cross
sectional properties – after all there is no fundamental reason to believe that as
companies progress and change through multiple years that they should be seen as
one observation.

This paper uses the strategy of having one regression that decomposes the
momentum coefficient into a linear function of Residual Gamma and other
controls. This motivates the estimation equation using interactions:

$$R_{i,t} = \alpha + \left(\beta + \lambda Residual\ Gamma_{i,t-1}\right)R_{i,t-1} + \gamma Residual\ Gamma_{i,t-1}$$

$$+ \epsilon_{i,t} \quad (3)$$

β is the momentum constant, while λ is the variable of interest as it
represents the sensitivity of momentum to our predictor factor Residual Gamma.
Note that this unexpanded equation is identical to the expanded form:

$$R_{i,t} = \alpha + \beta r_{i,t-1} + \lambda Residual\ Gamma_{i,t-1} R_{i,t-1} + \gamma Residual\ Gamma_{i,t-1} +$$

$\epsilon_{i,t}$. The expanded form makes it clear that the coefficient of interest is on the interaction of Residual Gamma and past returns, and motivates the addition of a first-order term $\gamma Residual\ Gamma_{i,t-1}$ since $Residual\ Gamma_{i,t-1}$ shows up as an interaction. If Residual Gamma does cause extra momentum, we expect higher values of Residual Gamma to mean that when yesterday experiences a high positive return that tomorrow's return will be higher as well. The interaction with lagged return is necessary for covariates and other factors that affect *momentum*. Non-interaction terms should be added only to the extent they are believed to directly affect tomorrow's returns.

   We would like to consider for the baseline lags of more than a single day. After all there's no fundamental reason to believe that hedging must happen within a day. In fact conversations with practitioners suggest that rebalancing with options often takes a few days. This is in contrast with leveraged ETFs which Cheng and Madhaven (2009) explain must contractually rebalance at the end of the day. To consider multiple days of momentum we add additional lags to our fundamental specification:

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} R_{i,t-k}\left(\beta_{1,k} + \lambda_{1,k} Residual\ Gamma_{i,t-k}\right)$$

$$+ \Sigma_{k=1}^{K} \gamma_1 Residual\ Gamma_{i,t-k} + \epsilon_{i,t}\ (4)$$

It is this equation that we will actually estimate. Again, in the regression, the key estimated parameters of interest are the λ_k variables. These variables represent additional momentum caused by Residual Gamma.

The results of this regression are presented in Table 3.4 Regression 1, which show the coefficients of interest λ_k as k varies from 1 to 10 inclusive. Baseline values of momentum β_k are also shown. λ_k can be interpreted as the additional return in percent caused by a 1% move in prices as Residual Gamma increase by one unit. This is slightly less than a standard deviation of Residual Gamma, which is 1.45.

Note first that the return drift occurs mostly in the first few days. After the fourth day or so, there is minimal additional momentum. However, the momentum does not revert away significantly either. That the reaction mostly accumulates within the first day is not surprising. Conversations with practioners reveal that most of the hedging for a stock's price moves occurs within a couple of days. That the effect does not revert away in any timeframe for which our data has power for suggests that the feedback has long-term effects and is not a temporary microstructure quirk.

The magnitude of the momentum is also significant. The momentum prediction provided by Residual Gamma results in an annualized R^2 of about 10%. According to Campbell Thompson (2008) formula (14) which assumes log utility, this results in 30.6% extra returns. That is to say, a participant earning 7.6% a year previously in an optimal fashion can now earn 10% with this new information. This gain is quite significant.

166

Table 3.4: Momentum in Stocks versus Residual Gamma.

OLS regressions of equity-date Returns on 10 lags of: net Returns ($R_{i,t}$) expressed as a fraction; Residual Gamma defined in Table 3.2 as the component of Log Gamma not predictable by price, shares outstanding, or volume; Residual Gamma multiplied by Return. All variables besides Residual Gamma are defined as in summary statistic Table 3.1. The regressions are run on the full sample (Col 1 and Col 2), the top 10 percentile Residual Gamma subsample of the population (Col 3), and the bottom 10 percentile Residual Gamma subsample of the population (Col 4). t-statistics are clustered by date.

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K}\beta_k \cdot R_{i,t-k} + \Sigma_{k=1}^{K}\lambda_k Residual\ Gamma_{i,t-k} \cdot R_{i,t-k}$$
$$+ \Sigma_{k=1}^{K}\gamma_k Residual\ Gamma_{i,t-k} + \epsilon_{i,t}$$

Variable	(1)	(2)	(3)	(4)
N	10,545,064	11,135,122	1,039,210	1,477,949
R^2	.0011	.0007	.0049	.0018
Sample	Full	Full	Top 10%tile of Residual Gamma	Bottom 10%ile of Residual Gamma
100*Return λ_k:				
(t-1)	.15 (0.22)	-.47 (-0.72)	5.70 (5.78)	-2.77 (-3.87)
(t-2)	-.54 (-1.91)	-1.75 (-2.33)	1.71 (1.40)	-1.83 (-3.12)
(t-3)	-.09 (-0.12)	-.37 (-0.55)	3.26 (2.78)	-1.64 (-2.98)
(t-4)	-.27 (-0.36)	-.44 (-0.63)	1.95 (1.46)	-.91 (-1.73)
(t-5)	-1.36 (-1.78)	-1.56 (-2.17)	-.51 (-0.39)	-1.44 (-2.57)
(t-6)	-.72 (-0.97)	-.52 (-0.75)	2.03 (1.50)	.02 (0.03)
(t-7)	-.32 (-0.43)	-.35 (-0.50)	2.19 (1.56)	-.42 (-0.81)
(t-8)	-.09 (-0.14)	-.15 (-0.24)	1.92 (1.48)	-.28 (-0.62)
(t-9)	-.11 (-0.15)	-.28 (-0.43)	.89 (0.70)	-.36 (-0.74)
(t-10)	.66 (0.91)	.64 (0.96)	2.56 (1.93)	.90 (1.79)
100*Return* Residual Gamma λ_k				
(t-1)	.78 (6.02)			
(t-2)	.26 (1.97)			
(t-3)	.01 (0.06)			
(t-4)	.08 (0.69)			
(t-5)	.14 (1.07)			
(t-6)	-.13 (-1.01)			
(t-7)	-.01 (-0.05)			
(t-8)	-.09 (-0.73)			
(t-9)	.09 (0.77)			
(t-10)	.03 (0.26)			
Residual Gamma lags	Yes	No	No	No

Table 3.5: The Impact of Residual Gamma on Momentum with Liquidity Controls.

OLS regressions of equity-date Returns on 10 lags of: Returns (R) expressed as a net fraction; Residual Gamma defined in Table 3.2 as the component of total gamma not predictable by price, shares outstanding, or volume; Log Dollar Volume, and Log Market Cap, Returns multiplied by Residual Gamma, Returns multiplied by Log Dollar Volume, Returns multiplied by Log Market Cap. All variables besides Residual Gamma are defined as in summary statistic Table 3.1. Residual Gamma is defined as in Table 3.2. N = 10,545,064. R^2 =0.0014. t-stats are clustered by dates. Coefficients are multiplied by 100.

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} R_{i,t-k}\left(\beta_{1,k} + \lambda_{1,k} Residual\ Gamma_{i,t-k} + \beta_{2,k} Log\ Dollar\ Volume_{i,t-k}\right.$$
$$\left. + \beta_{3,k} Log\ Market\ Cap_{i,t-k}\right)$$
$$+ \Sigma_{k=1}^{K}\left(\gamma_{1,k} Residual\ Gamma_{i,t-k} + \gamma_{2,k} Log\ Dollar\ Volume_{i,t-k}\right.$$
$$\left. + \gamma_{3,k} Log\ Market\ Cap_{i,t-k}\right) + \epsilon_{i,t}$$

Panel A

Date	Return (R) $100\beta_{1,k}$	$(\beta_{1,k})$	R * Residual Gamma $100\lambda_{1,k}$	$t(\lambda_{1,k})$	R * Log Dollar Volume $100\beta_{2,k}$	$t(\beta_{2,k})$	R * Log Market Cap $100\beta_{3,k}$	$t(\beta_{3,k})$
t - 1	2.182	(1.19)	.923	(7.05)	.474	(3.17)	.695	(-3.86)
t - 2	6.115	(3.33)	.223	(1.66)	-.186	(-1.28)	.348	(-1.91)
t - 3	1.407	(0.84)	-.060	(-0.48)	-.258	(-1.96)	.218	(1.29)
t - 4	1.357	(0.83)	.155	(1.24)	.208	(1.70)	.354	(-2.10)
t - 5	1.252	(0.70)	.164	(1.28)	.081	(0.64)	.278	(-1.58)
t - 6	2.726	(1.56)	-.074	(-0.57)	.171	(1.42)	.447	(-2.64)
t - 7	.416	(0.22)	.020	(0.16)	.094	(0.68)	.151	(-0.84)
t - 8	-.081	(-0.05)	-.074	(-0.61)	.058	(0.49)	.064	(-0.4)
t - 9	-1.109	(-0.58)	.123	(1.05)	.126	(1.08)	.066	(-0.39)
t - 10	-.230	(-0.12)	.047	(0.41)	.070	(0.58)	.018	(-0.11)

Panel B

Date	Residual Gamma $100\gamma_{1,k}$	$t(\gamma_{1,k})$	Log Dollar Volume $100\gamma_{2,k}$	$t(\gamma_{2,k})$	Log Market Cap $100\gamma_{3,k}$	$t(\gamma_{3,k})$
t - 1	-.003	(-0.12)	.096	(4.26)	-.510	(-1.64)
t - 2	.048	(1.91)	.021	(0.94)	.384	(1.02)
t - 3	-.026	(-1.06)	-.041	(-1.68)	-.211	(-0.66)
t - 4	-.004	(-0.16)	-.008	(-0.37)	.093	(0.30)
t - 5	-.042	(-1.81)	-.033	(-1.46)	.013	(0.04)
t - 6	.029	(1.11)	-.003	(-0.13)	.013	(0.04)
t - 7	-.017	(-0.76)	-.008	(-0.34)	-.055	(-0.19)
t - 8	-.007	(-0.28)	-.015	(-0.69)	.154	(0.43)
t - 9	-.005	(-0.22)	.020	(0.88)	-.081	(-0.22)
t - 10	.007	(0.35)	-.017	(-0.81)	.185	(0.69)

3.5.4 Liquidity and Time Controls

Recall in our conceptual estimation equation (3) that we are estimating the level of momentum as the factor $\left(\beta + \lambda Residual\ Gamma_{i,t-1}\right)$. One may be concerned that it is not the Residual Gamma that is directly causing momentum, but rather a covariate of Residual Gamma. This is dealt with by identifying the most likely covariates first and then then adding them as interaction covariates to remove omitted variable bias.

One likely omitted variable is company size. The numerator of Residual Gamma is aggregated gamma over open interest of options, which should scale up with the size of the company. We have already normalized out by volume (see subsection A) in our price impact model. However the relationship between volume and market cap may not exactly cancel out. There may be non-linearities in the interaction. Momentum may simply be expressed differently in larger companies than smaller companies. For that reason, company size represented as log market cap would be a key control.

Similarly, momentum likely would be affected by liquidity. More liquid stocks can be arbitraged more easily, which may reduce momentum effects. Likewise a more liquid stock should have less price impact in the first place. One industry standard for measuring liquidity is daily dollar volume, so we use log daily dollar volume as a control as well. This results in the estimation equation:

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} R_{i,t-k}\big(\beta_{1,k} + \lambda_{1,k} Residual\ Gamma_{i,t-k}$$

$$+ \beta_{2,k} Log\ Dollar\ Volume_{i,t-k} + \beta_{3,k} Log\ Market\ Cap_{i,t-k}\big)$$

$$+ \Sigma_{k=1}^{K}\big(\gamma_{1,k} Residual\ Gamma_{i,t-k} + \gamma_{2,k} Log\ Dollar\ Volume_{i,t-k}$$

$$+ \gamma_{3,k} Log\ Market\ Cap_{i,t-k}\big) + \epsilon_{i,t} \quad (5)$$

Comparing Table 3.5 and Table 3.4 shows the results of this regression in cumulative λ_k similar to before. This demonstrates that liquidity and market cap effects are not large confounds in our results.

We might also be interested in whether the effect might be stronger in larger companies. For one, since Residual Gamma is in logs, a 1% abnormal increase is a lot higher for a larger firm than a smaller one. Also, many market anomalies grow substantially stronger with smaller size. Is this the case for the current effect as well? For that purpose, we run a triple interaction regression similar to (5):

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} R_{i,t-k}\big(\beta_{1,k} + \lambda_{1,k} Residual\ Gamma_{i,t-k}$$

$$+ \lambda_{2,k} Residual\ Gamma_{i,t-k} \cdot Demeaned\ Log\ Market\ Cap_{i,t-k}$$

$$+ \beta_{2,k} Demeaned\ Log\ Market\ Cap_{i,t-k}\big)$$

$$+ \Sigma_{k=1}^{K}\big(\gamma_{1,k} Residual\ Gamma_{i,t-k}$$

$$+ \gamma_{2,k} Demeaned\ Log\ Market\ Cap_{i,t-k} + \gamma_{3,k} Residual\ Gamma_{i,t-k}$$

$$\cdot Demeaned\ Log\ Market\ Cap_{i,t-k}\big) + \epsilon_{i,t} \quad (5.5)$$

The standard interaction between Residual Gamma and return is given by $\lambda_{1,k}$, but now we have a triple interaction between this term and Demeaned Log Market Cap given by $\lambda_{2,k}$. The results of this regression are shown in Table 3.3. We see that there is no appreciable interaction between Log Market Cap and the R * Residual Gamma term (the third column). No coefficient out of the ten lags is significant at the 5% level. Thus, the impact of Residual Gamma on momentum seems equally strong for large and small equity-dates.

Another concern we may have is that options behavior may have shifted over time. For one, the number of options has dramatically increased exponentially over time (Figure 3.7). Further, dynamic hedging techniques have developed to be more sophisticated and have lower price impact. The baseline momentum in the stock market may have time trends itself. In fact, it is known that for individual stocks at least, there has been an upward trend in volatility until 1997 (Campbell et al, 2002), which has since reversed, and volatility has impact on other market parameters (Campbell and Hentschel 1992).
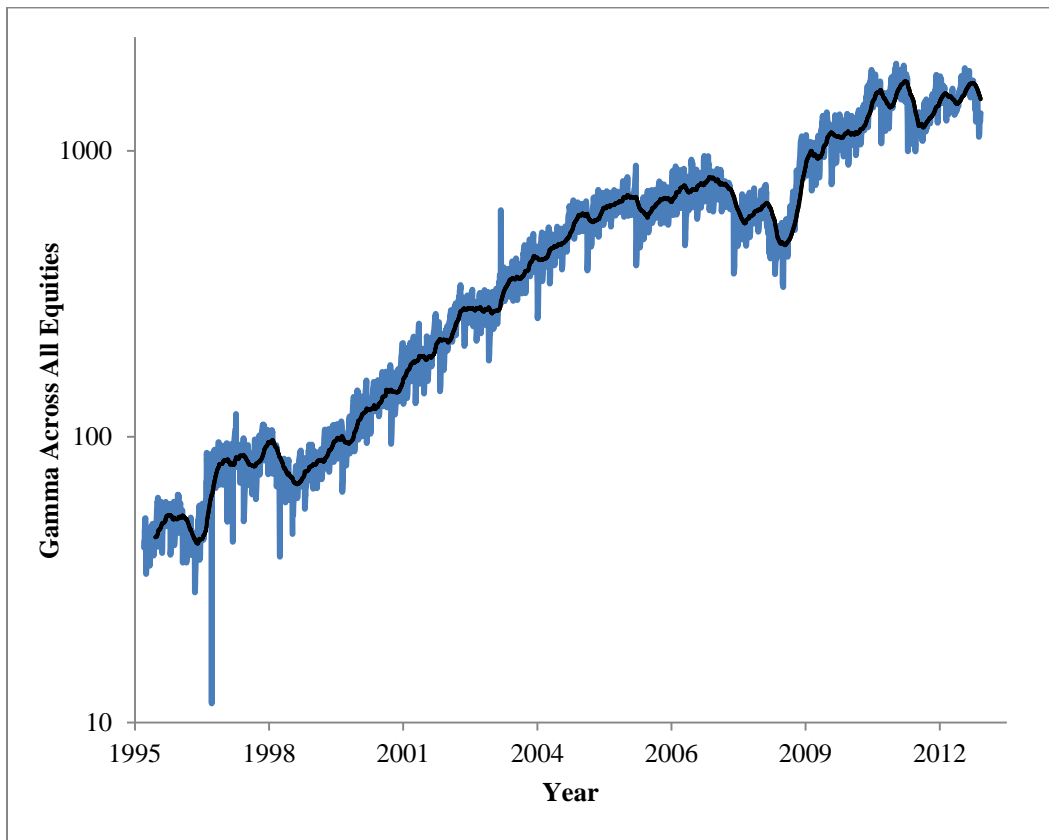
Figure 3.7: Total Gamma Across All Equities over Time.

Gamma across all options and equities summed for a given date through time. The sample consists of all options in OptionMetrics from the database start date in 1996 until 2013. Gamma is aggregated by taking the sum of open interest for an option contract multiplied by the gamma specific to that contract calculated via OptionMetrics using its internal Cox Ross Rubinstein (1979) binomial tree model. A 100 trading-day moving average is given by the dashed black line.

172

To address trend issues, we add two time controls. The first control, Time Serial, is simply an integer number representing the number of trading days from the start of the data series in 1996. To more specifically address the issue of open interest increasing over time, we add in a time variable of the log of the total amount of open interest over time as well. This is open interest summed over all equities and so it only has a time component to address time effects. We then arrive at the estimation equation:

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} R_{i,t-k}\big(\beta_{1,k} + \lambda_{1,k} Residual\ Gamma_{i,t-k}$$

$$+ \beta_{2,k} Log\ Total\ Open\ Interest_{t-k} + \beta_{3,k} Time\ Serial_{t-k}\big)$$

$$+ \Sigma_{k=1}^{K}\big(\gamma_{1,k} G_{i,t-k} + \gamma_{2,k} Log\ Total\ Open\ Interest_{t-k}\big)$$

$$+ \gamma_{3-1} Time\ Serial_{t-1} + \epsilon_{i,t} \quad (6)$$

Table 3.6 presents results from this regression, which is again very similar to the baseline Table 3.4 regression (1) both qualitatively and quantitatively. Further, combining controls from (5) and (6) and adding other transformations (e.g. square log daily volume) produce similar results, the output of which is not shown here.

Overall the baseline regression equation (4) has shown remarkable robustness to factors known to potentially affect momentum and be covariates of Residual Gamma. The controls added in (5) and (6) give strong evidence that it is not liquidity or time trend effects driving the main results.

Table 3.6: The Impact of Residual Gamma on Momentum with Time Controls.

OLS regressions of equity-date Returns on: Time Serial is an integer that starts with 1 and is incremented by 1 per trading day, and then 10 lags of: Returns (R) expressed as a net fraction; Residual Gamma defined in Table 3.2 as the component of total gamma not predictable by price, shares outstanding, or volume; Log Total Open Interest is the natural log of the total open interest of all options across all stocks on that day; Returns multiplied by Residual Gamma; Returns multiplied by Log Total Open Interest, Returns multiplied by Time Serial. 10 lags of Time Serial are not taken due to collinearity. All variables besides Residual Gamma are defined as in summary statistic Table 3.1. Residual Gamma is defined as in Table 3.2. N = 10,641,326. R^2 =0.0011. T-stats are clustered by dates. Coefficients are multiplied by 100.

$$R_{i,t} = \alpha + \Sigma_{k=1}^{K} R_{i,t-k}\left(\beta_{1,k} + \lambda_{1,k} Residual\ Gamma_{i,t-k} + \beta_{2,k} Log\ Total\ Open\ Interest_{t-k}\right.$$
$$\left. + \beta_{3,k} Time\ Serial_{t-k}\right) + \Sigma_{k=1}^{K}\left(\gamma_{1,k} G_{i,t-k} + \gamma_{2,k} Log\ Total\ Open\ Interest_{t-k}\right)$$
$$+ \gamma_{3-1} Time\ Serial_{t-1} + \epsilon_{i,t}$$

Panel A

	Return (R)		R * Residual Gamma		R * Log Total Open Interest		R * Time Serial	
	$100\beta_{1,k}$	$t(\beta_{1,k})$	$100\lambda_{1,k}$	$t(\lambda_{1,k})$	$100\beta_{2,k}$	$t(\beta_{2,k})$	$100\beta_{3,k}$	$t(\beta_{3,k})$
Date								
t - 1	-16.337	(-0.36)	0.848	(6.32)	1.156	(0.43)	-0.002	(-1.18)
t - 2	50.462	(0.95)	0.241	(1.77)	-3.165	(-1.02)	0.003	(1.43)
t - 3	-47.809	(-1.08)	0.029	(0.24)	2.864	(1.10)	-0.002	(-1.32)
t - 4	-3.601	(-0.08)	0.109	(0.86)	0.218	(0.08)	0.000	(-0.18)
t - 5	12.972	(0.26)	0.169	(1.34)	-0.761	(-0.26)	0.000	(-0.04)
t - 6	5.531	(0.11)	-0.134	(-1.04)	-0.395	(-0.14)	0.000	(0.25)
t - 7	9.478	(0.19)	0.003	(0.03)	-0.560	(-0.19)	0.000	(0.15)
t - 8	7.886	(0.18)	-0.086	(-0.73)	-0.498	(-0.19)	0.001	(0.30)
t - 9	8.044	(0.19)	0.084	(0.74)	-0.471	(-0.19)	0.000	(0.16)
t - 10	54.994	(1.24)	0.002	(0.02)	-3.304	(-1.27)	0.003	(1.77)

Panel B

	Residual Gamma		Log Total Open Interest		Time Serial	
	$100\gamma_{1,k}$	$t(\gamma_{1,k})$	$100\gamma_{2,k}$	$t(\gamma_{2,k})$	$100\gamma_{3,k}$	$t(\gamma_{3,k})$
Date						
t - 1	-0.077	(-4.48)	-0.091	(-0.22)	0.000	(1.03)
t - 2	0.031	(1.92)	0.052	(0.1)		
t - 3	0.014	(0.84)	-0.351	(-0.61)		
t - 4	0.005	(0.31)	0.216	(0.41)		
t - 5	-0.011	(-0.67)	0.232	(0.44)		
t - 6	0.028	(1.64)	-0.817	(-1.59)		
t - 7	-0.005	(-0.31)	-0.078	(-0.15)		
t - 8	0.008	(0.48)	0.214	(0.43)		
t - 9	-0.026	(-1.68)	-0.029	(-0.06)		
t - 10	0.015	(1.05)	0.554	(1.55)		

Table 3.7: Using Instrument #1 (Time to Expiration): The Impact of Residual Gamma on Momentum.

IV regressions of equity-date Returns on separate constants for each equity and 10 lags of: Returns (R) expressed as a net fraction, with coefficients generated separately for each equity; Residual Gamma defined in Table 3.2 as the component of total gamma not predictable by price, shares outstanding, or volume; Returns multiplied by Gamma Residual. Residual Gamma is instrumented by the following exogenous variables: indicator variables for each day of the month, generated separately for each equity. Residual Gamma is defined as in Table 3.2. N = 10,132,923. R^2 =0.0010. T-stats are clustered by dates. The IV estimator is two-stage least-squares. Coefficients are multiplied by 100.

$$\text{Instruments: } \Sigma_{k=1}^{31} b_{i,k} I(Day\ of\ Month = k)$$

$$R_{i,t} = \alpha_i + \Sigma_{k=1}^{K} \beta_{i,k} r_{i,t-k} + \Sigma_{k=1}^{K} \lambda_k Residual\ Gamma_{IV,i,t-k} R_{i,t-k}$$
$$+ \Sigma_{k=1}^{K} \gamma_k Residual\ Gamma_{IV,i,t-k} + \epsilon_{i,t}$$

	R * Residual Gamma (IV)		Residual Gamma (IV)		Returns (Per Equity)	Constants (Per Equity)
	$100\lambda_k$	$t(100\lambda_k)$	$100\gamma_k$	$t(100\gamma)$	$\beta_{i,k}$	α_i
Date					YES	YES
t - 1	1.299	(6.23)	-0.078	(-4.27)		
t - 2	-0.043	(-0.23)	0.035	(2.02)		
t - 3	-0.072	(-0.39)	0.014	(0.82)		
t - 4	0.122	(0.65)	0.003	(0.19)		
t - 5	0.363	(1.84)	-0.012	(-0.72)		
t - 6	0.184	(0.97)	0.029	(1.57)		
t - 7	-0.213	(-1.07)	-0.013	(-0.82)		
t - 8	-0.052	(-0.27)	0.006	(0.35)		
t - 9	0.008	(0.04)	-0.027	(-1.66)		
t - 10	-0.192	(-1.00)	0.022	(1.47)		

3.6   Instrumental Regressions

We move on to identify the causal effects of Residual Gamma on momentum in an instrument setting. We explain three instruments, illustrate their first stages, describe the empirical specifications, and finally discuss the results.

3.6.1   Instrumental Regressions: Time to Expiry

Residual Gamma may vary endogenously across equities and dates with respect to momentum. Reverse causality could occur in many possible stories. One such story could be that expected momentum for a stock affects the cost of issuing options. Thus, stocks with exogenously more momentum might have end users who most want to hedge volatility risk, resulting in more options issued and more Residual Gamma. Exogenous factors could affect both momentum and Residual Gamma as well. For example, during earnings announcement season end-users might purchase unusually high amounts of options to lever up their opinions or insider information. At the same time, news regarding earnings and its impact on the company's future could be slowly integrating into the market, causing momentum (Hong Stein 1999).

We can bypass these endogeneity issues if can find exogenous variation of Residual Gamma. The exogenous sources of variation that we will present will be a large portion of this paper's contribution. All three instruments will exploit the various idiosyncratic institutional structure of options issuances.

176

The first instrument we use is the monthly expiration and roll-off of options. This expiration causes a monthly cycle in options outstanding as end-users put on orders throughout the month, and then at one single instant, a significant fraction of options expire. Generally the roll-off date is around the 18th of each month. Figure 3.1 shows the cyclical nature of Residual Gamma through the month. The options expiration dates are not lined up with the start of months, earnings releases, or other natural cycles. Therefore, options expiration provides a potential instrument for Residual Gamma.

The empirical exogenous instruments we are using for the first stage regression is:

$$Residual\ Gamma_{i,t} = \ \Sigma_{k=1}^{31}b_{i,k}I(Day\ of\ Month = k) \quad (7)$$

In other words, for each equity we run dummies for each day of month. This creates 31 dummies (no constant) that fully capture any variation in that equity's Residual Gamma as part of the monthly cycle. As can be observed in Figure 3.1, the growth in Residual Gamma seems linear over the month, with a strong drop within the last two days of the expiration cycle.

The second stage is the direct instrumental analogue of (4):

$$R_{i,t} = \alpha_i + \Sigma_{k=1}^{K}\beta_{i,k}r_{i,t-k} + \ \Sigma_{k=1}^{K}\lambda_k Residual\ Gamma_{IV,i,t-k}R_{i,t-k}$$

$$+ \ \Sigma_{k=1}^{K}\gamma_k Residual\ Gamma_{IV,i,t-k} + \epsilon_{i,t} \quad (8)$$

Note that the estimated $Residual\ Gamma_{IV,i,t-k}$ has cross sectional variation, but within a given stock only cyclic variation over time. There is no quarterly variation or time trends. Identification is purely off of the monthly expiration cycle.

The output of the regression is given in Table 3.7. We see good agreement with the general price path λ_k in this regression and the baseline OLS regression in Table 3.4. Namely, the cumulative coefficient magnitudes are nearly the same. This validates our baseline regression.

3.6.2 Instrumental Regressions: Options Issuance and Past Price

The second instrument we exploit is the fact that options tend to be struck around current stock prices, and options purchasers have lag in rebalancing the strikes of options when the stock price moves. As an illustration, Figure 3.2 shows a hypothetical Residual Gamma as a function of the underlying stock price if the underlying stock price has been more or less around 100 in the past month. The underlying principle behind the illustration is that the price of the stock being around 100 results in most options being struck around this price.

Figure 3.3 demonstrates this phenomenon empirically, plotting the Residual Gamma against the price move the last 20 trading days. When large movements are seen, the shift away from past prices reduces Residual Gamma.

Figure 3.5 shows suggestive evidence of tapering of options away from current prices as well.

One possible objection is that using price movements has confounds with volatility. If a stock experiences a sharp drop, volatility will increase, market makers may sell due to risk limits, mechanically reducing Residual Gamma. This can be ruled out via two methods. First, we can control for volatility, which we do in our empirical specification. The shape of Figure 3.3 is maintained even with volatility controls. Second, if there is a volatility story, one would expect that downward movement of prices increases volatility more than upward movement of prices, leading to asymmetry (Chacko Viceira 2005). Note however that Figure 3.3 is symmetrical suggesting against the volatility story, but in line with the theory of options clustering around old prices.

By using a simple volatility control, we ensure that we are comparing between stocks with the same day-to-day volatility. One stock by chance just happens to have a long string of negative moves, and another has a 50%-50% balance of negative and positive moves.

The fact that end-users do not immediate rebalance strikes has foundations in much of financial economics and behavioral economics of defaults (Beshears et al 2009, Carroll et al 2009, Choi et al 2003; 2004, Kamenica 2012, Thaler and Sunstein 2008).

Since options are struck around past prices, if current equity prices are around past prices, Residual Gamma will be higher than otherwise. For example,

if over the course of two months, the price of Exxon Mobile (XOM) stock goes from 100 to 90 and then to 80, the Residual Gamma will be lower than if XOM goes from 100 to 90 and then to 100. This past-price-level dependency of residual gamma provides our second instrument to identify the effect of residual gamma on momentum.

The empirical specification we will use in our first stage IV regression for this instrument will be:

$$Residual\ Gamma_{i,t} = b_i \left| Last\ 20\ Day\ Return_{i,t} \right| + covariates + \epsilon_{i,t} \qquad (9)$$

$\left| Last\ 20\ Day\ Return_{i,t} \right|$ specifies the absolute value of the stocks return for the past 20 calendar days. The absolute value is used instead of a squared value due to descriptive data suggesting that the relationship is not quadratic, such as seen in Figure 3.3.

The second stage is similar to (8) but with $Last\ 20\ Day\ Volatility_{i,t}$ placed in as a volatility control – since we expect the instrument $\left| Last\ 20\ Day\ Return_{i,t} \right|$ to have omitted variable effects through volatility. $Last\ 20\ Day\ Volatility_{i,t}$ refers to the volatility of the stock in the last 20 day calendar days expressed in the square root of variance. The second stage is:

$$R_{i,t} = \alpha_i + \Sigma_{k=1}^{K} R_{i,t-k} \left(\beta_{i,k} + \lambda_k Residual\ Gamma_{IV,i,t-k} \right.$$

$$+ \delta_k Last\ 20\ Day\ Volatility_{20,i,t} \Big)$$

$$+ \Sigma_{k=1}^{K} \left(\gamma_{1,k} Residual\ Gamma_{IV,i,t-k} \right.$$

$$+ \gamma_{2,k} Last\ 20\ Day\ Volatility_{20,i,t} \Big) + \epsilon_{i,t} \quad (10)$$

The results are presented in Table 3.8, which again is qualitatively and quantitatively similar to our baseline Table 3.4.

Table 3.8: Using Instrument #2 (Past Price Path): The Impact of Residual Gamma on Momentum.

IV regressions of equity-date Returns on separate constants for each equity and 10 lags of: Returns (R) express as a net fraction, with separate coefficients for each equity; Residual Gamma defined in Table 3.2 as the component of total gamma not predictable by price, shares outstanding, or volume; Last 20 Day Volatility calculated as the square root of the sum of the last 20 days of returns; Returns multiplied by Residual Gamma; Returns multiplied by Last 20 Day Volatility. Residual Gamma is instrumented by the following exogenous variables: the sum of the last 20 days of returns, generated separately for each equity. All variables besides Residual Gamma are defined as in summary statistic Table 3.1. Residual Gamma is defined as in Table 3.2. N = 10,351,259. R^2 =0.0012. T-stats are clustered by dates. The IV estimator is two-stage least-squares. Coefficients are multiplied by 100. Controls for returns per equity and constants per equity are included in the results shown below even though they are not explicitly shown.

$$Instruments: b_i | Last\ 20\ Day\ Return_{i,t} |$$

$$R_{i,t} = \alpha_i + \Sigma_{k=1}^{K} R_{i,t-k}(\beta_{i,k} + \lambda_k Residual\ Gamma_{IV,i,t-k} + \delta_k Last\ 20\ Day\ Volatility_{20,i,t})$$
$$+ \Sigma_{k=1}^{K}(\gamma_{1,k} Residual\ Gamma_{IV,i,t-k} + \gamma_{2,k} Last\ 20\ Day\ Volatility_{20,i,t})$$
$$+ \epsilon_{i,t}$$

	R * Residual Gamma		R * Last 20 Day Volatility		Residual Gamma		Last 20 Day Volatility	
	$100\lambda_k$	$t(\lambda_k)$	$100\delta_k$	$t(\delta_{2,k})$	$100\gamma_{1,k}$	$t(\gamma_{1,k})$	$100\gamma_{2,k}$	$t(\gamma_{2,k})$
Date								
t - 1	1.630	(5.90)	1.049	(0.28)	-0.077	(-4.13)	-0.021	(-0.02)
t - 2	0.175	(0.56)	-10.881	(-2.62)	0.036	(2.10)	0.026	(0.02)
t - 3	-0.296	(-1.14)	-1.568	(-0.43)	0.011	(0.65)	0.499	(0.49)
t - 4	0.177	(0.58)	-5.003	(-1.24)	0.008	(0.50)	1.128	(1.17)
t - 5	0.531	(1.79)	-6.155	(-1.56)	-0.003	(-0.20)	-1.746	(-1.93)
t - 6	0.174	(0.60)	-0.444	(-0.12)	0.023	(1.32)	-0.206	(-0.23)
t - 7	0.025	(0.08)	-0.965	(-0.27)	-0.017	(-1.00)	-0.254	(-0.26)
t - 8	-0.003	(-0.01)	-2.154	(-0.65)	0.004	(0.22)	1.072	(1.06)
t - 9	-0.010	(-0.03)	0.088	(0.03)	-0.024	(-1.49)	-1.147	(-1.17)
t - 10	-0.246	(-0.88)	0.415	(0.13)	0.018	(1.20)	0.809	(1.05)

3.6.3 Instrumental Regressions: Options Issuance and Round Numbers

Finally, we use a third instrument, the round number clustering of option strikes. Option exchanges only issue options at prices that are round (e.g. 40, 50, 60, etc). Further, the more round a price is, the more market makers and end users will coordinate to issue at that strike and increase liquidity. Even though strikes exist at both 50 and 45, the strike at 50 will be a stronger coordination point and see more options issued. This instrument is particularly clean. There are relatively few a priori reasons to believe that stocks should otherwise behave differently around round numbers. There are however some previous research of round number effects: Harris (1991) models round numbers as simplifying negotiations; Donaldson and Kim (1993) show some market actors attach special meaning to the DJIA exceeding multiples of 100 inducing non-continuity; Christie and Schultz (1994) famously uncovered that market makers collude by coordinating on round prices. None of these effects seem like they would add an obvious confound to our IV analysis. As an institutional artifact there are few reasons to believe in serious fundamental reasons round prices should otherwise be special.

The issuance of options at round numbers causes essentially exogenous variation in Residual Gamma. In the few days before options expiration, if the stock's price is in line with a round number (e.g. 50), then Residual Gamma will be much higher than if the stock's price is not round (e.g. 53). For an illustration, Figure 3.4 shows a hypothetical example one day before all options expiration in

which all options are issued at the round numbers ending in 0. Residual Gamma would cluster around these values. Figure 3.5 shows an empirical example of option strikes clustering at round numbers, with more clustering near numbers that are more round.

Whether the stock price hits a round number is essentially exogenous. Only a very weak version of the random walk model needs to hold for this variation to be both random and unpredictable.

To develop our first stage for the instrument, we need to have a measure for how round a price is. A few issues come to play here. First is institutional convention for what counts as a round price. We would like to identify prices that 1) the exchange has issued strikes at and 2) actually trade with high volume and open interest. We could dive into first-level options level data to see for each equity exactly what the strike grid looks like every day. This level of analysis would be complex if it is to be complete. We would find that closeness to the nearest \$10 matters, but for some stocks, so does closeness to the nearest \$5 or \$2.5, albeit less. Sometimes this can depend on time of month too or the year of the observation. The complexity would inhibit an easy interpretation and cleanness of the analysis.

Instead we opt for a simple and robust measure of roundness. We first normalize prices to have two digits before the decimal point: so for example \$4.33, \$43.30 and \$433.02 all map to \$43.30. Then we look at the distance of this normalized price to the nearest \$10, \$5, \$2.5, \$1, \$.50, and \$.25. These divisions

of tenths and halves tend to be empirically the round prices around which options are struck. We run an empirical regression for each equity on how much roundness to each of the six increments matter. In particular, we regress Residual Gamma on instruments below:

$\text{Residual Gamma}_{i,t}$

$$= b_{i,1} Distance\ from\ Normalized\ Price\ to\ Nearest\ 10$$

$$+ b_{i,2} Distance\ from\ Normalized\ Price\ to\ Nearest\ 5$$

$$+ b_{i,3} Distance\ from\ Normalize\ Price\ to\ Nearest\ 2.5$$

$$+ b_{i,4} Distance\ from\ Normalize\ Price\ to\ Nearest\ 1$$

$$+ b_{i,5} Distance\ from\ Normalize\ Price\ to\ Nearest\ 0.5$$

$$+ b_{i,6} Distance\ from\ Normalize\ Price\ to\ Nearest\ .25 + covariates$$

$$+ \epsilon_{i,t}$$

The advantage of this empirical method is that we do not need to take a stance on how much each level of roundness matters – this is estimated by the model.

By way of example, if AAPL has a price around \$500, we would normalize its price by dividing by 10 to get its prices fluctuating around \$50. Then, we observe how much Residual Gamma goes up when AAPL price is near round \$10s (\$40,\$50,\$60), and similarly for \$5 (\$45, \$50, \$55), and \$2.5 (\$47.5,\$50,\$52.5), and soforth, and estimate an empirically determined model for that as our instrument for exogenous Residual Gamma.

This measure of roundness, while not perfect, is easy to understand. Further, it is less likely that this instrument is weak due to its empirical estimation.

The second stage is again the standard equation below. Results are presented in

Table 3.9

$$R_{i,t} = \alpha_i + \Sigma_{k=1}^{K} \beta_{i,k} r_{i,t-k} + \Sigma_{k=1}^{K} \lambda_k Residual\ Gamma_{IV,i,t-k} R_{i,t-k}$$

$$+ \Sigma_{k=1}^{K} \gamma_k Residual\ Gamma_{IV,i,t-k} + \epsilon_{i,t}$$

Table 3.9: Using Instrument #3 (Distance to Round Numbers): The Impact of Residual Gamma on Momentum.

IV regressions of equity-date Returns on separate constants for each equity and 10 lags of: Returns (R) express as a net fraction, with separate coefficients for each equity; Residual Gamma defined in Table 3.2 as the component of total gamma not predictable by price, shares outstanding, or volume; Returns multiplied by Residual Gamma; Returns multiplied by Last 20 Day Volatility. Residual Gamma is instrumented by the following exogenous variable: first the Price of an equity-date is multiplied or divided by 10 to form a Normalized Price that has exactly two digits before the decimal point; then 6 variables are generated by taking the distance of this Normalized Price to the nearest multiple of 10, 5, 2.5, 1, .5, and .25 respectively. All variables besides Residual Gamma are defined as in summary statistic Table 3.1. Residual Gamma is defined as in Table 3.2. N = 10,545,064. R^2 =0.0009. T-stats are clustered by dates. The IV estimator is two-stage least-squares. Coefficients are multiplied by 100.

$$\text{Instruments: } b_{i,1}\text{Distance from Normalized Price to Nearest } 10$$
$$+ b_{i,2}\text{Distance from Normalized Price to Nearest } 5$$
$$+ b_{i,3}\text{Distance from Normalize Price to Nearest } 2.5$$
$$+ b_{i,4}\text{Distance from Normalize Price to Nearest } 1$$
$$+ b_{i,5}\text{Distance from Normalize Price to Nearest } 0.5$$
$$+ b_{i,6}\text{Distance from Normalize Price to Nearest } .25$$

$$R_{i,t} = \alpha_i + \Sigma_{k=1}^{K}\beta_{i,k}r_{i,t-k} + \Sigma_{k=1}^{K}\lambda_k Residual\ Gamma_{IV,i,t-k}R_{i,t-k}$$
$$+ \Sigma_{k=1}^{K}\gamma_k Residual\ Gamma_{IV,i,t-k} + \epsilon_{i,t}$$

	R * Residual Gamma (IV)		Residual Gamma (IV)		Returns (R Per Equity)	Constants (Per Equity)
	$100\lambda_k$	$t(100\lambda_k)$	$100\gamma_k$	$t(100\gamma)$	$\beta_{i,k}$	α_i
Date					YES	YES
t - 1	1.342	(6.61)	-0.078	(-4.27)		
t - 2	-0.154	(-0.83)	0.034	(2.00)		
t - 3	-0.134	(-0.70)	0.014	(0.82)		
t - 4	-0.036	(-0.19)	0.003	(0.18)		
t - 5	0.312	(1.66)	-0.012	(-0.71)		
t - 6	0.197	(1.05)	0.029	(1.58)		
t - 7	-0.190	(-1.00)	-0.013	(-0.81)		
t - 8	-0.127	(-0.68)	0.006	(0.35)		
t - 9	0.005	(0.03)	-0.027	(-1.65)		
t - 10	-0.182	(-0.95)	0.022	(1.48)		

3.6.4 Interpretation

In our analysis above, we first ran a baseline (OLS) regression of the impact of Residual Gamma on momentum. We then controlled for possible confounds by adding covariates. Then we moved on to identify the causal effects of Residual Gamma on momentum in an instrument setting using three instruments: time to expiry, past price paths, and round numbers.

All the different variations of analysis above resulted in the quantitatively and qualitatively similar impacts of Residual Gamma on momentum. This is particularly substantial because the driving forces behind the regressions above vary widely. For example, time and liquidity effects that make the OLS regression problematic should not be an issue for the time-to-expiration instrument. Similarly, monthly cycle variations that may pose problems for time-to-expiration should not affect whether prices are round or not. But regardless of how the effect is generated or cut, the effect still remains the same. This provides strong evidence that the Residual Gamma and hence hedging feedback demand of options does indeed cause extra momentum.

3.7 Extensions

3.7.1 In what other ways are options destabilizing?

We have measured destabilization in this paper through momentum, but under a purely rational expectations world, one might expect that prices would be impounded not after a few days, but on the very day of the stock move. Cheng

and Madhaven (2009) capture this intuition by looking at price movements until 3:00PM, and then designate the period between 3:00PM and 4:00PM as the momentum period. Taken to the extreme, the price should actually move contemporaneously. In a purely rational Kyle (1985) model, market makers should anticipate the hedging and simply reduce liquidity to compensate.

These already provide two alternate measures of destabilization: volatility as measured by variance of future returns, and liquidity as measured by price impact or bid-ask spread. Other measures of stock destabilizing should exist as well.

3.7.2   Why do end-users purchase options?

In general equilibrium it would be useful to have an explanation of why end users purchase options in the first place, which gives non-linear exposure to the underlying equity. One theory perhaps could be that options mechanically bundle leverage and concavity together. End-users purchase options for the leverage factor, and purchase along with that concavity, an unintended add on. End-users do not pay attention every period, which both explains why end-users do not simply create their own options through BSM, and also why they do not rebalance away their concavity when prices change.

3.7.3   What relationship does this have with time-varying returns?

189

Another theory for the purchase of options can reflect the inherent preference by end-users or end users to hold more equities when the price is higher, and less when the price is lower. Numerous studies show that equity returns are time-varying, and in particular households hold less risky assets in bad times driving up returns (Zhang 2013). Could it be possible that investor heterogeneity creates a class of agents that prefer to take risk in good states of the world only? If so, then options seem like the optimal instrument with which to take on the risk. In that case, options can be seen as more than a simple dynamic hedging instrument, but an indicator of who and to what extent different agents hold risk in different states of the world.

Also, to the extent that bad fundamental news can cause feedback selling, then perhaps this can be one channel through which bad cashflow news induces bad discount rate news, in the Campbell Shiller (1988) decomposition sense.

3.8   Conclusion

This paper presents empirical evidence to show that the options create hedging feedback demand. This hedging feedback demand, measured empirically as Residual Gamma in our paper, causes market makers to purchase stocks exactly after prices rise, and sell stocks exactly after prices fall. In a frictional world, this should cause destabilization through momentum.

The paper exhibits this momentum effect. The effect is large in magnitude, with a one standard deviation increase in Residual Gamma causing 1.5% extra

190

momentum in stocks. The momentum occurs within one to four days and does not dissipate for a significant period of time, at least 10 days.

This effect on stock prices is robust to liquidity controls, market cap controls, open interest controls, and time trend controls. Instruments including option expiration cycle, past price effects, and round price effects are also used to identify the destabilization. The remarkable robustness of the results gives strong evidence that Residual Gamma indeed increases stock momentum. This feedback of options shows that derivative structure does indeed have effect on stock prices.

Chapter 4

4   Appendices

4.1   Appendix to Chapter 1

Table 4.1: Summary Statistics.

The summary statistics below covers all 243 quarters from the beginning of the *Federal Reserve Flow of Funds* series in 1951 to 2012. All data on sectors and their holdings are collected from the Federal Reserve Flow of Funds: Household equities is the series on households and nonprofit organizations, corporate equities, asset (FL153064105.Q); Household Equity Mutual Funds is the series on households and nonprofit organizations, mutual fund shares, asset (FL153064205.Q); Household Credit Markets is the series on households and nonprofit organizations, total currency and deposits including money market fund shares, asset (FL154000025.Q); Household Deposits is the series on households and nonprofit organizations, credit market instruments, assets (FL154004005.Q).

Household risky ratio is calculated as the sum of the first two series above divided by the sum of the second two series. Total risky ratio is calculated as the analogously except for the series corresponding to all sectors. Change in household risky ratio is the current period household risky ratio minus a one quarter lag. Log household risky ratio is the natural log of household risky ratio. Log total risky ratio is the natural log of the total risky ratio.

In future tables, R denotes the excess return calculated as the difference between the value-weighted CRSP minus the one-month Treasury bill rate from Ibbotson's in percentage points. S denotes the equity share in new issues as defined in Baker Wurgler (2000). CAPE denotes the ten year cyclically adjusted price to earnings ratio, defined as in Campbell Shiller (1988); Term Spread denotes the yield premium of ten year (from Robert Shiller's website) over one-month treasury bill rates. CAY is the consumption wealth ratio proxy defined by Lettau Ludvigson (2001). Risky valuation is the projection of CAPE on the household risky ratio. Risky Issuance is the residual from this projection.

	1951-2012		1951-1981		1982-2012	
Variable	Mean	SD	Mean	SD	Mean	SD
Panel A: Federal Reserve Flow of Funds Data Series						
Household Equities ($trillions)	4.289	5.356	.567	.258	8.009	5.445
Household Equity Mutual Funds ($tn)	2.441	3.443	.104	.090	4.776	3.579
Household Credit Market ($trillions)	1.437	1.593	.214	.099	2.525	1.459
Household Deposits ($trillions)	1.552	1.592	.226	.117	2.646	1.461
Panel B: Other Primary Data Series						
Excess Return R (%)	1.829	8.399	1.652	8.138	2.001	8.676
Lettau-Ludvigson CAY	.000	0.017	-0.004	0.01	0.004	0.021
Term Spread	2.275	1.231	1.794	0.941	2.743	1.302
Baker-Wurgler Equity Share of Issuance	0.186	0.095	0.224	0.082	0.143	0.090
Campbell-Shiller CAPE	18.887	7.534	15.735	4.597	21.963	8.525
Panel C: Derived Data Series						
Household risky ratio	2.509	0.815	2.540	.761	2.470	.866
Change in Household risky ratio	0.007	0.240	0.001	0.239	0.012	0.243
Log Household risky ratio	0.869	0.321	0.888	0.305	0.850	0.337
Total risky ratio	0.501	0.162	0.492	0.141	0.509	0.181
Log Total risky ratio	-0.747	0.340	-0.754	0.311	-0.740	0.367
Risky Valuation	-0.747	0.301	-0.865	0.243	-0.632	0.310
Risky Issuance	.000	0.157	0.111	0.096	-0.108	0.126

Table 4.2: Values of Percentiles of the Household risky ratio.

The table displays values of select percentiles of the household risky ratio variable. The household risky ratio is the ratio of household high risk over household low risk assets as reported by the *Federal Reserve Flow of Funds* data. Percentiles are selected to be approximately 1/6 of the distribution. 95% confident intervals are given by the Thiel-Sen method.

Percentile	Value	95% Conf Interval
17%	1.685	(1.62 , 1.77)
33%	1.996	(1.87 , 2.09)
50%	2.304	(2.20 , 2.52)
67%	2.866	(2.73 , 3.03)
83%	3.352	(3.23 , 3.55)

References

Alissa, W. (2009). Boards' response to shareholders' dissatisfaction: the case of shareholders' say on pay in the UK. *Available at SSRN 1412880*.

Amin, K., J.D. Coval, and H. N. Seyhun. (2004): Index Option Prices and Stock Market Momentum. *Journal of Business*77:835-74.

Asness, Clifford, Tobias Moskowitz and Lasse Pedersen (2008), " Value and Momentum Everywhere," working paper.

Bainbridge, S. (2008). Remarks on Say on Pay: An unjustified incursion on director authority. Working Paper.

Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. *The journal of Finance*, *43*(3), 593-616.

Baker, M., Greenwood, R., & Wurgler, J. (2003). The maturity of debt issues and predictable variation in bond returns. *Journal of Financial Economics*, *70*(2), 261-291.

Baker, Malcolm, and J. Stein (2004): "Market Liquidity as a Sentiment Indicator," *Journal of Financial Markets* 7 (3): 271-299.

Baker, M., & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance*, *55*(5), 2219-2257.

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, *61*(4), 1645-1680.

Balsam, S., & Yin, J. (2012). The Impact of Say-on-Pay on Executive Compensation. *Available at SSRN 2026121*.

Bates, D.S. (2003): "Empirical Option Pricing: A Retrospection," *Journal of Econometrics,* 116, 387-404.

Baumol, W. J. (1959). Business Behaviour. *Value and Growth, New York*.

Bebchuk, L. A. (2007). The myth of the shareholder franchise. *Virginia Law Review*, 675-732.

Bebchuk, L. A., & Fried, J. M. (2004). *Pay without performance* (Vol. 278). Cambridge, MA: Harvard University Press.

Bebchuk, L. A., Fried, J. M., & Walker, D. I. (2002). *Managerial power and rent extraction in the design of executive compensation* (No. w9068). National Bureau of Economic Research.

Beshears, J., J. J. Choi, D. Laibson, B.C. Madrian (2009), "The Importance of Default Options for Retirement Savings Outcomes," *Social Security Policy in a Changing Environment*, 167-195.

Bertrand, M., & Mullainathan, S. (2001). Are CEOs rewarded for luck? The ones without principals are. *The Quarterly Journal of Economics*, *116*(3), 901-932.

Black F, M Scholes (1973) "The pricing of options and corporate liabilities" – *The Journal of Political Economy*, 81, 637-654.

Bondarenko, O. (2003) "Why Are Put Options So Expensive?" Working Paper, University of Illinois at Chicago.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (2008). The myth of long-horizon predictability. *Review of Financial Studies*, *21*(4), 1577-1605.

Brainard, W. C., & Tobin, J. (1968). Pitfalls in financial model building. *The American Economic Review*, *58*(2), 99-122.

Brealey, R. A., Myers, S. C., & Allen, F. (2006). *Corporate finance* (Vol. 8). McGraw-Hill/Irwin.

Cai, J., & Walkling, R. (2009). Shareholders' Say on Pay: Does it Create Value?. *Journal of Financial and Quantitative Analysis (JFQA), Forthcoming*, 2008-06.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2006). *Down or out: Assessing the welfare costs of household investment mistakes* (No. w12030). National Bureau of Economic Research.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of financial economics*, *18*(2), 373-399.

Campbell, J. Y., & Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *The Journal of Political Economy*, *107*(2), 205-251.

Campbell John Y. and L Hentschel (1992): "No News is Good News, An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics,* 31 (3): 281-318.

Campbell John Y., M. Lettau, B.G. Malkiel, Y. Xu, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 56 (1): 1-43.

Campbell John Y., AWC Lo, AC MacKinlay (1997): The Econometrics of Financial Markets : Princeton University Press.

Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial studies*, *1*(3), 195-228.

Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance*, *43*(3), 661-676.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. *Review of Financial Studies*, *21*(4), 1509-1531.

Campbell, J. Y., & Yogo, M. (2005). Implementing the econometric methods in "Efficient tests of stock return predictability". *Unpublished working paper. University of Pennsylvania*.

Campbell, J. Y., & Yogo, M. (2006). Efficient tests of stock return predictability.*Journal of financial economics*, *81*(1), 27-60.

Carroll, G.D., J.J. Choi, D. Laibson, B.C. Madrian, and A. Metrick (2009): *The Quarterly Journal of Economics* 124(4):1639-1674.

Chacko, G., & Viceira, L. M. (2005). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies, 18*(4), 1369-1402.

Cheng, Minder and Madhaven, Ananth, The Dynamics of Leveraged and Inverse-Exchange Traded Funds (November 4, 2009). Journal of Investment Management, Winter 2009.

Choi, J.J., D. Laibson, B.C. Madrian, and A. Metrick (2004): "For Better or for Worse: Default Effects and 401(k) Savings Behavior," *Perspective on the Economics of Aging*.

Choi, J.J., D. Laibson, B.C. Madrian, A. Metrick (2003): "Optimal Defaults," *American Economic Review*, Papers and Proceedings 93,2:180-5.

Christie, W. G., & Schultz, P. H. (1994). Why do NASDAQ Market Makers Avoid Odd‐Eighth Quotes?. *The Journal of Finance*, 49(5), 1813-1840.

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

Cochrane, J. H., & Piazzesi, M. (2005). Bond risk premia. *The American economic review*, *95*(1), 138-160.

Conyon, M., & Sadler, G. (2010). Shareholder voting and directors' remuneration report legislation: say on pay in the UK. *Corporate Governance: An International Review*, *18*(4), 296-312.

Core, J. E., Guay, W., & Larcker, D. F. (2008). The power of the pen and executive compensation. *Journal of Financial Economics*, *88*(1), 1-25.

Core, J. E., Holthausen, R. W., & Larcker, D. F. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of financial economics*, *51*(3), 371-406.

Coval, J. D., and T. Shumway. (2001): "Expected Option Returns," *Journal of Finance* 56:983-1009.

Cox, J.C., S.A. Ross, M. Rubinstein (1979): "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (3): 229-263.

Cuñat, V., Gine, M., & Guadalupe, M. (2012a). Say Pays! Shareholder Voice and Firm Performance. OECD ELSA Seminars Series.

Cuñat, V., Gine, M., & Guadalupe, M. (2012b). The vote is cast: the effect of corporate governance on shareholder value. *The Journal of Finance*, *67*(5), 1943-1977.

DeAngelo, H., & DeAngelo, L. (1991). Union negotiations and corporate policy: A study of labor concessions in the domestic steel industry during the 1980s.*Journal of Financial Economics*, *30*(1), 3-43.

Donaldson, R. G., & Kim, H. Y. (1993). Price barriers in the Dow Jones industrial average. *Journal of Financial and Quantitative Analysis*, 28(3).

Driessen, J., and P. Maenhout. (2007): "An Empirical Portfolio Persepctive on Option Pricing Anomalies," *Review of Finance* 11:561-603.

Ertimur, Y., Ferri, F., & Maber, D. A. (2012). Reputation penalties for poor monitoring of executive pay: Evidence from option backdating. *Journal of Financial Economics*, *104*(1), 118-144.

Fama, E. F. (1980). Agency Problems and the Theory of the Firm. *The Journal of Political Economy*, 288-307.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *The Journal of Political Economy*, 246-273.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy,* 607-636.

Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. *Journal of financial economics*, *5*(2), 115-146.

Ferraris, A. (2008) "Equity Market Impact Models," *Deutsche Bank Working Paper: Sifterverband fuer die Deutsche Wirtschaft.*

Ferreira, F., & Gyourko, J. (2009). Do political parties matter? Evidence from US cities. *The Quarterly Journal of Economics*, *124*(1), 399-422.

Ferri, F., & Maber, D. A. (2013). Say on pay votes and CEO compensation: Evidence from the UK. *Review of Finance*, *17*(2), 527-563.

Fuster, A., Laibson, D., & Mendel, B. (2010). Natural expectations and macroeconomic fluctuations. *The Journal of Economic Perspectives*, *24*(4), 67-84.

Froot, K. A., & Perold, A. F. (1995). New trading practices and short‑run market efficiency. *Journal of Futures Markets*, 15(7), 731-765.

Gabaix, X., A. Krishnamurthy, and O. Vigneron, (2007): "An Empirical Portfolio Perspective on Option Pricing Anomalies." *Review of Finance* 11:561-603.

Gabaix, X., & Landier, A. (2008). Why has CEO pay increased so much?. *The Quarterly Journal of Economics*, *123*(1), 49-100.

Garleanu, Nicolae, Lasse Pedersen and Allen Poteshman (2009), "Demand-Based Option Pricing," Review of Financial Studies, forthcoming.

Gordon, M. J. (1962). *The investment, financing, and valuation of the corporation*. Greenwood Press.

Greenwood, Robin. (2005): "Short and Long Term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage," *Journal of Financial Economics* 75:607-49.

Greenwood, Robin, Samuel Hanson, and Jeremy C. Stein. (2010): "A Gap-Filling Theory of Corporate Debt Maturity Choice," *Journal of Finance*, 65 (3): 953-1028.

Greenwood, R., & Shleifer, A. (2013). *Expectations of returns and expected returns* (No. w18686). National Bureau of Economic Research.

Greenwood, R., & Vayanos, D. (2010). Price pressure in the government bond market. *The American Economic Review*, *100*(2), 585-590.

Laibson, D. I., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., & Akerlof, G. A. (1998). Self-control and saving for retirement. *Brookings Papers on Economic Activity*, *1998*(1), 91-196.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*,*69*(1), 201-209.

Hall, B. J., & Murphy, K. J. (2003). *The trouble with stock options* (No. w9784). National Bureau of Economic Research.

Harri, A., & Brorsen, B.W. (2009). The Overlapping Data Problem. *Quantitative and Qualitative Analysis in Social Sciences,* 2009(3),78-115.

Harris, L. (1991). Stock price clustering and discreteness. *Review of Financial Studies*, 4(3), 389-415.

Hickok, Robert L. and Rainville, Gay P. (2013). Round Two of Shareholder Say-on-Pay Litigation. *The Legal Intelligencer.* June 4, 2013.

Hjalmarsson, E. (2006). On the predictability of global stock returns. *rapport nr.: Working Papers in Economics*, (161).

Hölmstrom, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 74-91.

Hölmstrom, B., & Kaplan, S. N. (2001). *Corporate Governance and Merger Activity in the US: Making Sense of the 1980s and 1990s* (No. w8220). National Bureau of Economic Research.

Hölmstrom, B., & Kaplan, S. N. (2003). The state of US corporate governance: what's right and what's wrong?. *Journal of Applied Corporate Finance*, *15*(3), 8-20.

Hölmstrom, B., & Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, 303-328.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance, 54*(6), 2143-2184.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467-475.

Imbens, G. W., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, *79*(3), 933-959.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615-635.

Jackwerth, J. (2000): "Recovering risk aversion from option prices and realized returns," *Review of Financial Studies*, 13, 433-451.

Jegadeesh Narasimhan and Titman, Sheridan , Momentum (October 23, 2001). University Of Illinois Working Paper.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, *3*(4), 305-360.

Jensen, M.C., Murphy, K., & Wruck, E. (2004). Remuneration: Where we've been, how we got to here, what are the problems, and how to fix them.

Joskow, P. L., Rose, N. L., & Wolfram, C. D. (1996). Political constraints on executive compensation: Evidence from the electric utility industry. *The RAND Journal of Economics*, 165-182.

Kamenica, E. "Behavioral Economics and the Psychology of Incentives," *Annual Review of Economics* (4): 427-52.

Kaplan, S. N. (2008). Are US CEOs Overpaid?. *The Academy of Management Perspectives*, *22*(2), 5-20.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of financial Economics*, *17*(2), 357-390.

Kendall, M. G. (1954). Note on bias in the estimation of autocorrelation.*Biometrika*, *41*(3-4), 403-404.

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

Kuhnen, C., & Zwiebel, J. (2008). Executive pay, hidden compensation and managerial entrenchment. *Rock Center for Corporate Governance Working Paper*, (16).

Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica: *Journal of the Econometric Society,* 1315-1335.

Lamont, O. A., & Thaler, R. H. (2003). Anomalies: The law of one price in financial markets. *The Journal of Economic Perspectives*, *17*(4), 191-202.

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*, 281-355.

Lee, D. S., Moretti, E., & Butler, M. J. (2004). Do voters affect or elect policies? Evidence from the US House. *The Quarterly Journal of Economics*, *119*(3), 807-859.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *the Journal of Finance*, *56*(3), 815-849.

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

Lewellen, W. G., & Huntsman, B. (1970). Managerial pay and corporate performance. *The American Economic Review*, *60*(4), 710-720.

Longstaff, F. A. (1995): "Option Pricing and the Martingale Restriction," *Review of Financial Studies* 8:1091-124.

McInish, T. H., & Wood, R. A. (1992). An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *The Journal of Finance*, 47(2), 753-764.

Merton, R. (1973): "The Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science* 4:141-83.

Murphy, K.J. (1997). Executive compensation and the Modern Industrial Revolution. *International Journal of Industrial Organization*, *15*(4), 417-425.

Murphy, K. J. (1999). Executive compensation. *Handbook of labor economics*,*3*, 2485-2563.

Nelson, C. R., & Kim, M. J. (1993). Predictable stock returns: The role of small sample bias. *The Journal of Finance*, *48*(2), 641-661.

Newey, W. K. (1993). 16 Efficient estimation of models with conditional moment restrictions. *Handbook of statistics*, *11*, 419-454.

Newey, W. K., & West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix.

Nichols, A. (2007). Causal inference with observational data. *Stata Journal*,*7*(4), 507.

Pástor, Ľ., & Stambaugh, R. F. (2009). Predictive systems: Living with imperfect predictors. *The Journal of Finance*, *64*(4), 1583-1628.

Porter, J. (2003). Estimation in the regression discontinuity model. *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison*, 5-19.

Poteshman, A. M. (2001): "Underreaction, Overreaction, and Increasing Misreaction to Information in the Options Market," *Journal of Finance* 56:851-76.

Rehavi, M. M. (2007). Sex and politics: Do female legislators affect state spending. *Unpublished manuscript, University of Michigan*.

Roberts, D. R. (1956). A general theory of executive compensation based on statistically tested propositions. *The Quarterly Journal of Economics*, *70*(2), 270-294.

Rubinstein, M. (1994). "Implied Binomial Trees," *Journal of Finance* 41:579-90.

Shleifer, A. (1986): "Do Demand Curves for Stocks Slope Down?" *Journal of Finance* 41:579-90.

Shleifer, A. (2003). *Inefficient markets: An introduction to behavioral finance*. OUP Oxford.

Solow, R.M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics,* 70(1), 65-94.

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

Stein, J. (1989): "Overreactions in the Options Market," *Journal of Finance* 44:1011-23.

Thaler R. and C. Sunstein (2008). Nudge: Improving Decisions About Health, Wealth, and Happiness .

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, *51*(6), 309.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838.

Wurgler, J., and E. V. Zhuravskaya. (2002): "Does Arbitrage Flatten Demand Curves for Stocks?" *Journal of Business* 75:583-608

Valkanov, R. (2003). Long-horizon regressions: Theoretical results and applications. *Journal of Financial Economics*, *68*(2), 201-232.

Van der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on CollegeEnrollment: A Regression–Discontinuity Approach*. *International Economic Review*, *43*(4), 1249-1287.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455-1508.

Yermack, D. (1997). Good timing: CEO stock option awards and company news announcements. *The Journal of Finance*, *52*(2), 449-476.

Zhang, Fan A., (2013): "Do Options Impact the Stock Market?" *Harvard University Working Paper.*

Zhang, Fan A., (2013): "Household risky ratio and the Equity Premium," *Harvard University Working Paper.*