# Essays on Financial Markets

by

Albert Sibo Hu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Christine Parlour, Chair
Professor Terrence Hendershott
Professor Yuriy Gorodnichenko

Fall 2017

**Essays on Financial Markets**

**Abstract**

Essays on Financial Markets

by

Albert Sibo Hu

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Christine Parlour, Chair

This dissertation consists of three chapters that concern financial markets. The first chapter analyzes how financial frictions impact local government infrastructure spending. The US is frequently criticized for its recent poor spending and upkeep in local infrastructure, such as schools, utilities, and roads. I provide evidence that financial frictions in municipal bond markets exacerbate infrastructure underfunding. To demonstrate this, I exploit the differential exposure to government-rescued monoline bond insurers during the financial crisis as a quasi-experiment that affected local government access to municipal bond markets. I show that local governments with more exposure to a government-rescued monoline insurer FSA and its purchaser Assured Guaranty had better borrowing outcomes and spending in capital investments, relative to others that had exposure to non-rescued insurers. Event-study coefficients and difference-in-difference regression estimates show that issuers in the treatment group issue more bonds in the years after 2008, and also spend more on capital investments. The effect is significant for categories of public goods for which federal resources are scarce – specifically in education, housing development, and some general utilities.

The second chapter analyzes the effect of passive indexing on informed trading. I develop a model of segmented trade that considers the effect of comovement on asset pricing, efficiency, and incentives to acquire information. The main contribution is to show that traders who participate via passive indexes have characteristics of both informed and uninformed traders. They are like uninformed traders as they diversify away all idiosyncratic risk, and so do not seek costly information, but they are alike informed traders as their presence causes any informed trading to be more quickly disseminated into prices. I provide some empirical justification for this model using a regression-discontinuity design around Russell 2000 index inclusions with institutional holding data.

The third chapter explores the asset pricing properties of cryptocurrencies, an emerging type of digital transaction that utilizes decentralized, cryptographic methods to verify ownership. This chapter is co-authored with my dissertation chair Professor Christine Parlour. We provide summary statistics on cryptocurrency return properties and measures of common variation for over 200 digital coins. Secondly, we provide investment characteristics of initial coin offerings (ICOs), a method of crowdfunding that utilizes cryptocurrencies as legal tender. We reconcile these statistics with traditional finance theories and develop a set of empirical facts for this new asset class.

# Contents

# List of Figures

# List of Tables

## Acknowledgments

I extend my gratitude towards all the professors and PhD students whom I had the good fortune to work with during my graduate studies. Their sharp minds have given me a constant source of motivation and inspiration.

I would to like to thank my family for their unending support. I thank my parents Wong Lai-Ming and Hu Jishan for their love and encouragement.

I am grateful to my dissertation committee members Christine Parlour, Terrence Hendershott, and Yuriy Gorodnichenko for their patience in guiding me to complete my dissertation. I would also like to thank Nicolae Gârleanu, Amir Kermani, Gustavo Manso, Hoai-Luu Nguyen, Christopher Palmer, Demian Pouzo, and David Sraer for their helpful research comments.

I would like to thank all my fellow students with whom I have worked closely at the Economics department and at the Haas School of Business. I thank Abhay Aneja, Carlos Avenancio, Francesco D'Acunto, David Echeverry, Farshad Haghpanah, Dmitri Koustas, Sheisha Kulkarni, Haoyang Liu, Ryan Liu, David Schöenholzer, Waldo Ojeda, Markus Pelger, Xinxin Wang, Calvin Zhang, and Dayin Zhang.

Finally, I want to thank the staff at the Haas School of Business for their help in administrative needs and funding. I especially thank Kim Guilfoyle, Melissa Hacker, Bradley Jong, and Lisa Sanders Villalba.

# Chapter 1

# Bond Insurer Collapses and Municipal Government Debt Issuance

## 1.1 Introduction

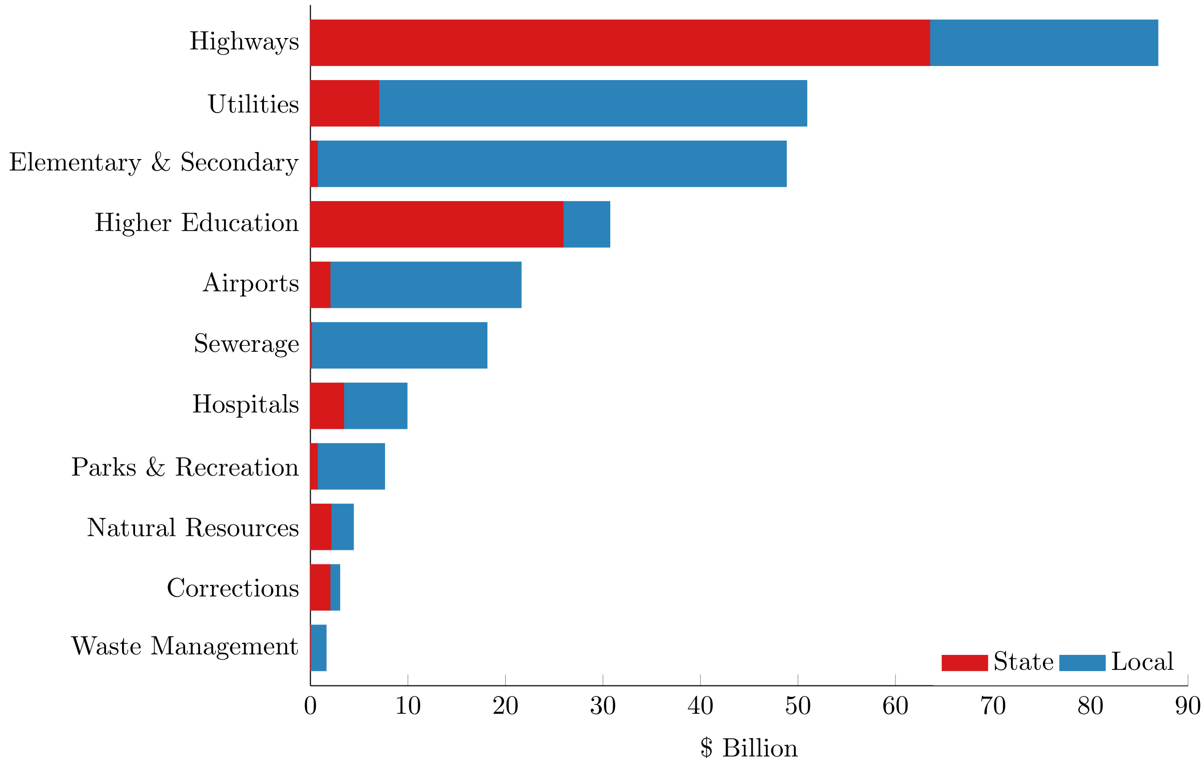
For more than two decades, the American Society of Civil Engineers has given grades ranging from "D" to "D+" to the country's overall infrastructure quality. Surveys and interviews with US mayors also say that the largest hurdle to economic prosperity in cities is better infrastructure, referencing the dilapidation and underfunding of roads, schools, water, transportation, sewerage, and waste systems. These deficiencies continue, in spite of longstanding economic scholarship on how public goods are to be adequately provided by the government (Samuelson (1954)).

However, the division of labor between federal, state, and local governments complicates optimal financing. Figure 1.1 provides a snapshot of capital outlays at the state and local level. Governments at these levels are responsible for high amounts of infrastructure spending. The most prominent source of federal funding is through direct grants or aid to these governments. US Census data show that these grants amounted to $630 billion in 2010 (Figure 1.2), most infrastructure spending delegated to state-funded highways. Since benefits from many local projects are not perceived to accrue nationally, and local governments cannot promise prudential use of grants, federal policymakers are politically inclined to withhold aid for needed local capital projects[1]. This concern came into focus, for example, after the devastation by hurricane Harvey in 2017, when the Harris County Flood Control District revealed that it had been lobbying the federal government for years to initiate flood mitigation projects.

In addition to grants, the federal government also offers subsidies in the form of tax-exemptions for interest payments on local government debt. These tax incentives allow government issuers to borrow more cheaply than possible otherwise. In 2014, the total sum

[1]See Sammartino, Frank. "Testimony: Federal Support for State and Local Governments Through the Tax Code", April 25, 2012.

Figure 1.1: US State and Local Capital Outlays, Fiscal Year 2014

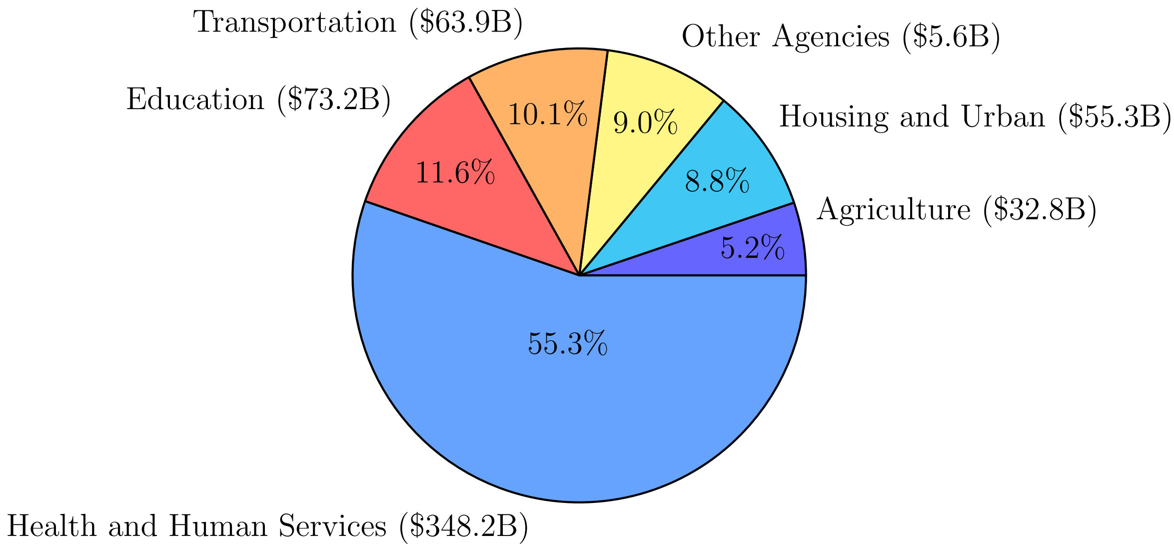



This figure shows local capital outlays by US state and local governments by category. Data is aggregated from the US Census Survey of Local Government Finances, for the 2014 fiscal year.

of debt, mostly in the form of long-term debt, neared $3 trillion. In fact, up to 90% of capital spending is funded by debt, the major source being municipal bonds (Marlowe (2015)). This alternative funding source relegates the responsibility of infrastructure funding to state and local governments. Figure 1.3 shows that, relative to other highly developed countries, the US has both low rankings of infrastructure quality – such as that measured by the World Economic Forum – and also a low ratio of *subnational* budget surplus relative to transfers from the central government. It is natural to ask whether these two facts are related.

As in any other market, financial frictions have a potentially important role for US infrastructure underfunding. The 2008-2009 financial crisis laid bare the possibility that fiscally decentralized funding potentially leaves local governments susceptible to nationwide debt market shocks. Figure 1.12 shows that new debt issued by municipalities fell significantly after the crisis, along with capital spending, and their recoveries have not kept up with the recovery in real GDP. Simultaneously, as shown in Figure 1.5, positive yield spreads over safe yields have increased and persisted across the rating spectrum for municipal securities. These changes from pre-crisis periods are strongly suggestive of structural changes to the municipal bond market.

Figure 1.2: Federal Aid to State and Local Governments, Fiscal Year 2010



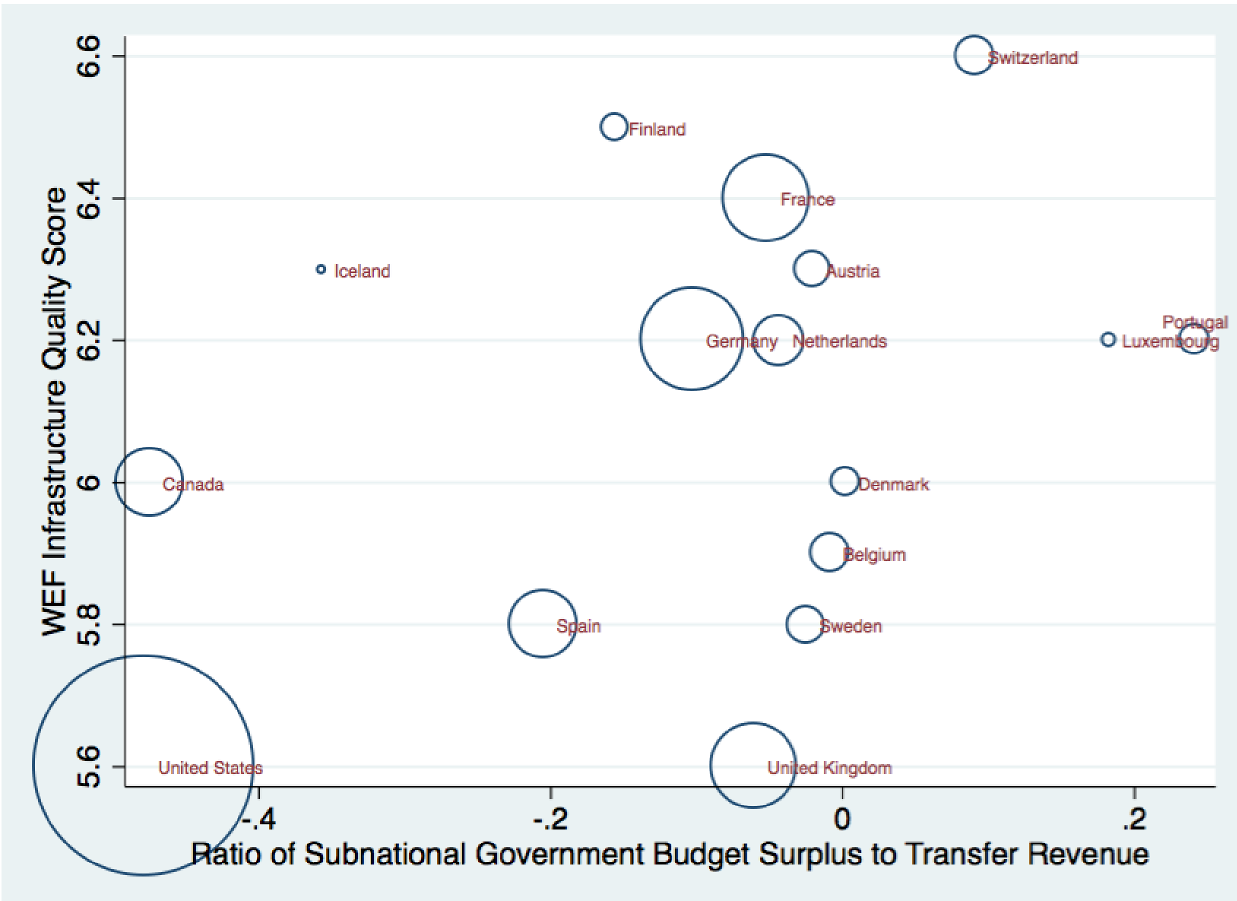

This figure shows the amount of federal aid to state and local governments in the 2010 fiscal year. Data only includes earmarked grants towards state and local governments for public investment purposes, and not other types of expenditure such as redistribution, welfare, or Medicare. The data source is the US Census Survey of Local Government Finances.

To study one such persistent distortion to the municipal market, I exploit the differential variation in exposure among municipal governments before the financial crisis to the collapse of US bond insurers (also known as *monolines*) during the crisis, which includes the highly publicized credit downgrades of AMBAC, MBIA, and FSA. These collapses were primarily attributable to adverse losses in subprime-related securitized products, and not due to municipal bond defaults. The variation in exposure, driven by the desire for bond insurers to diversify their portfolios, is also plausibly unrelated to municipalities' underlying economic conditions. These collapses resulted in increasing asymmetric information in municipal debt issuance (Cuny (2016)), increasing the cost of debt. Since issuance is endogenous to economic outcomes, I use this variation to establish a quasi-experimental variation for debt issuance to achieve a causal interpretation on investment outcomes.

Toward this end, I combine bond-level issuance data of US local municipal governments from SDC Platinum with detailed government spending data created by the US Census in order to link bond insurance, debt issuance, and local government finances. To achieve the statistical variation necessary for this study, I focus on debt issuance by county, city, and school districts.

I first show that local governments that are completely insured just prior to the financial crisis with FSA and Assured Guaranty issue up to 20% more bonds after 2008. They also are more likely to have new bonds insured. This new debt is additionally associated with a

Figure 1.3: Measure of decentralization vs. World Economic Forum Infrastructure Score (2012)

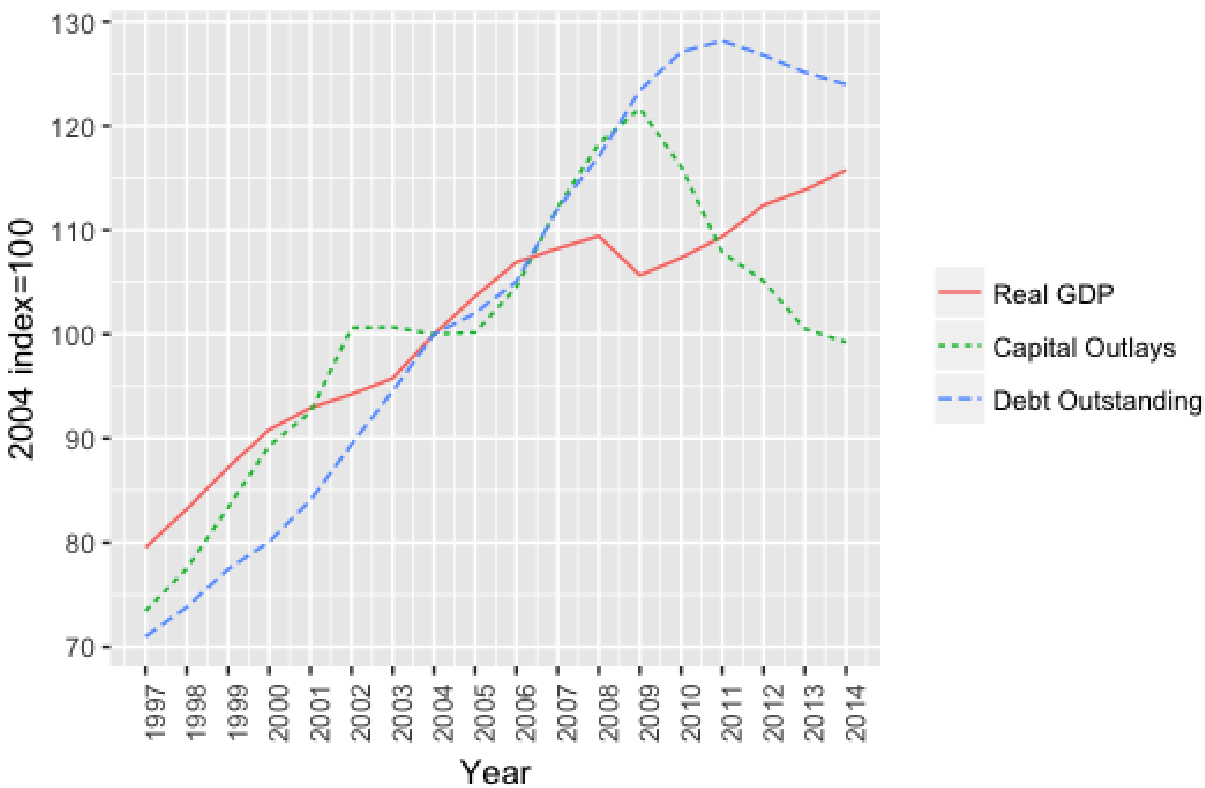



The x-axis variable is derived from measures published by the OECD, composed of ratios of budget surpluses of the subnational level, divided by the transfer revenue of those governments from the central government. The y-axis variable is the World Economic Forum infrastructure quality score for each of the countries shown, which ranges from 1 to 7. All variables are measured in 2012.

25% in higher levels of post-crisis capital outlays. These outlays are concentrated in school, airport, housing development, and certain classes of utilities – areas in which federal grants have been low. The evidence suggests that financial constraints matter for types of local infrastructure spending for which decentralized funding is most commonly used.

This paper is related to the optimal structure of government in the adequate provision of public goods. There is a large literature in public finance that attempts to address some of the benefits and costs of decentralized financing. The mechanisms that are related include inefficient levels of government due to tax competition between states (Wilson (1986)), the possibility of interregional destabilization (Von Hagen (2007)), and corruption due to com-

Figure 1.4: State and Local Government Debt vs. Capital Outlays (2009 dollars)



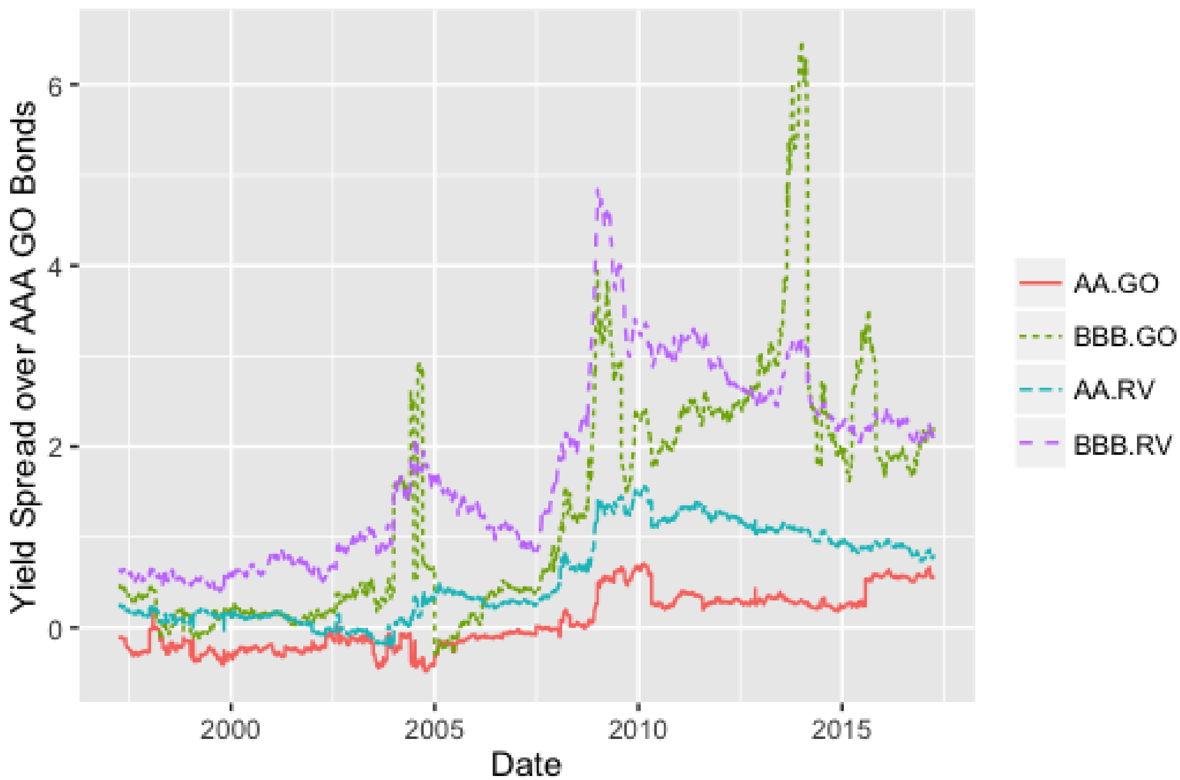

Capital outlay and debt outstanding data comprises of all capital investment, public construction and maintenance spending, collected from the US Census Survey of Local Government Finances. All series are deflated using 2009 dollars, and normalized to equal to 100 in 2004.

plexity of government (Fan et al. (2009)). To my knowledge, this is the first paper of its kind to measure the costs of financial frictions in the context of fiscal decentralization.

This paper is also related to the strand of literature on financing frictions and economic outcomes. The approach is in the spirit of Chodorow-Reich (2014) and Greenstone et al. (2014), both which use exposure to adverse banking conditions to study the effect on employment outcomes of firms. In this paper, I show that such frictions mattered also for the public sector during the 2008-2009 financial crisis, and not only for firms. The effect of debt issuance on capital outlays has also been examined by Cellini et al. (2010), which studies close bond elections in California school districts. In contrast, this paper examines debt effects for the universe of US municipal bonds and examine outlays for other major types of municipalities.

This paper contributes to a third literature concerning the effect of credit ratings on investment. Kisgen (2006) empirically examines credit ratings affect on capital structure

Figure 1.5: Yield Spreads of AA and BBB General Obligation and Revenue Municipal Bonds

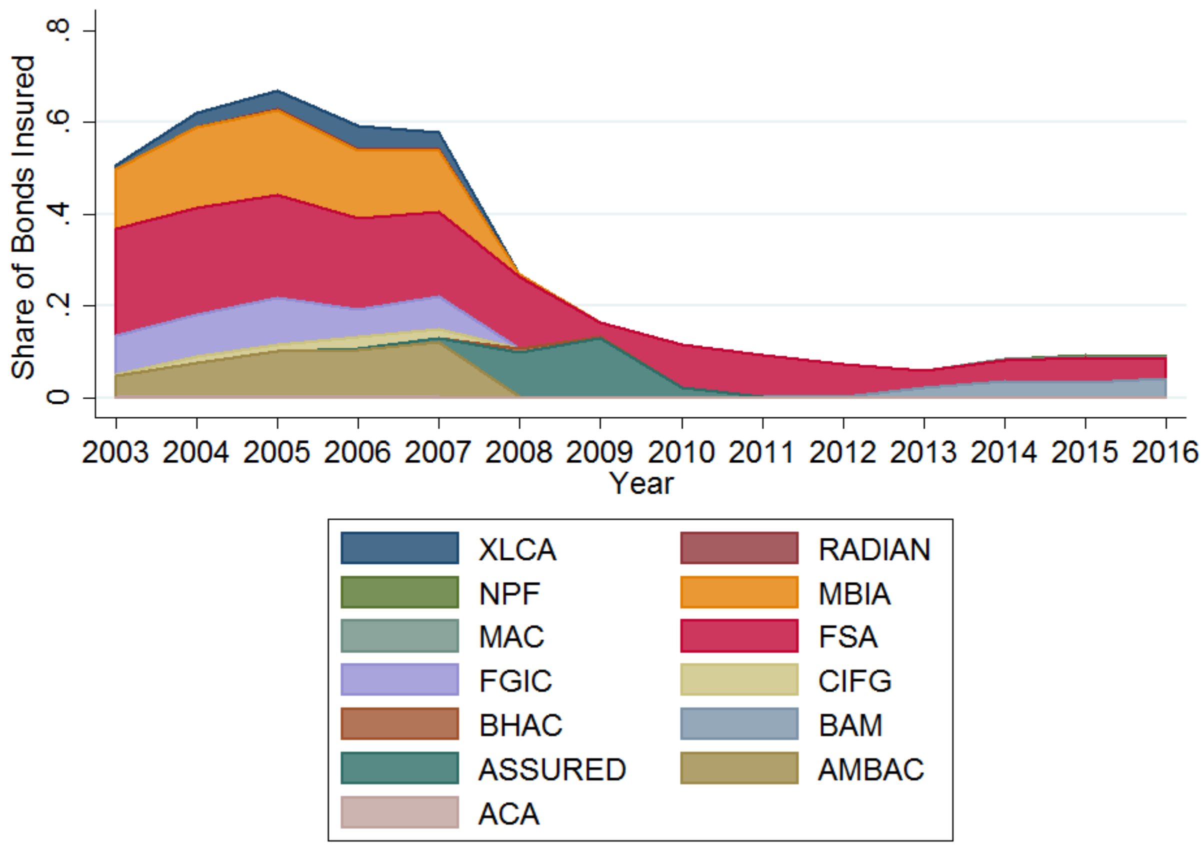



The four series are calculated from the the yields-to-maturities of the Bank of America Merrill Lynch US Municipal Securities Indices Data, collected from Thompson Reuters Datastream. Each series is calculated by taking the bond index with the relevant credit rating and bond type, subtracting the AAA General Obligation bond index.

due to informational benefits. Since all financial, spending, and revenue information for municipalities is public information, the empirical results here can be taken as a lower bound to the effect of credit rating on spending. A close paper on this subject is Adelino et al. (2017), where the authors use Moody's recalibration of their credit rating scale to find a positive effect of debt expansion on investment. The innovation in this present article is instead to focus on the role of monoline bond insurers during the financial crisis as the source of informational distortion.

The remainder of this article is organized as follows. Section 1.2 discusses the structure of the municipal bond market and the role of bond insurance. Section 1.3 gives an account of the data and methodology used, including defining the treatment and control groups and discussing econometric validity. Section 1.4 discusses the results. Section 1.5 concludes.

Figure 1.6: Fraction of Municipal Bonds Bonds Insured by Year




This figure plots the market share of each monoline bond insurer in the municipal bond market from 2003 to 2016. Each share is calculated as the notional of bonds that are insured in a given year divided by the total bonds issued that year. The data is from SDC Platinum database of global public finance bond issues. Each issue in the database lists up to three bond insurers.

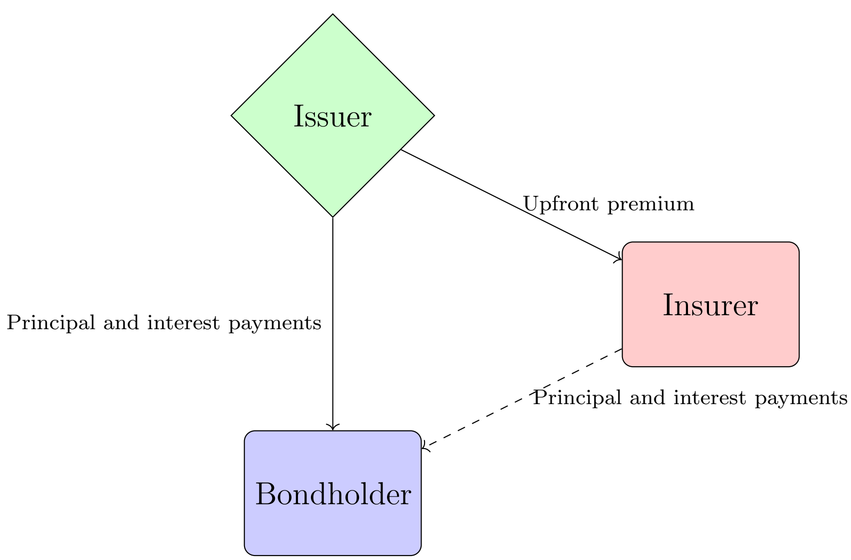
## 1.2 Background

Before the financial crisis, bond insurance was integral to the municipal bond market. Figure 1.6 shows by notional amount the fraction of newly issued municipal bonds insured by one of the dozen major bond insurers. Approximately 50% of new bond issues carried some form of insurance guarantee before the financial crisis, and in 2007 the total notional amount of municipal bonds insured exceeded $500 billion.

In a typical arrangement, the insurer charges the issuer a premium at time of issuance (Figure 1.7), in exchange for guaranteeing principal and interest payments on the insured bond in the event of a default. The insurers, in other words, take no role in potential bankruptcy proceedings or act to prevent any default scenarios from occurring. They merely agree to protect bondholders from any payment shortfalls (dashed arrow in Figure 1.7). This agreement is mediated by the underwriter of the bond issue, usually presenting a selection of monoline insurers. After a mutual selection process between the insurer and the issuer, this agreement then becomes part of the written prospectus agreement of each bond issue. The underlying rating of the bonds insured is subsequently derived from the credit rating of the bond insurer, rather than the underlying credit rating of the issuer.

Issuers benefit from this exchange by being able to afford lower interest rate payments. Figure 1.8 demonstrates the annual potential savings. In 2005, annual interest savings of insuring a BBB-rated $10 million 30-year fixed-rate bond is 39bp, or $39,000 annually in interest costs. According to Kriz and Joffe (2017), the median premium is 0.485 percent of

Figure 1.7: Diagram of Bond Insurance Process




This figure presents a diagram illustrating the relationship between issuers (local governments), bondholders, and the municipal bond insurer. Issuers pay an upfront premium to monoline insurers in return for protection on any principal or payment shortfalls to bondholders.

the total par value. This implies that large interest cost savings are possible to issuers who choose to insure.

Several reasons explain why this market exists. One major reason is the requirements of state-level conduit financing authorities that facilitate the debt issuance procedures of municipalities. Conduits such as the New Jersey Healthcare Facilities Financing Authority and the Illinois Financing Authority require debt issues they support to have investment grade ratings. Others, such as the California Educational Facilities Authority and the Delaware State Housing Authority require "A" ratings. Several of these conduits allow bond insurance in place to satisfy the required credit ratings. Even without the large interest savings benefit, the existence of such requirements ensures that bond issuers might undertake bond insurance, lest investors substitute away from their issues to higher rated ones. New difficulties of issuing debt in a manner that satisfied rating requirements by either regulatory authorities or market participants were revealed in the numerous lawsuits against the monoline insurers:

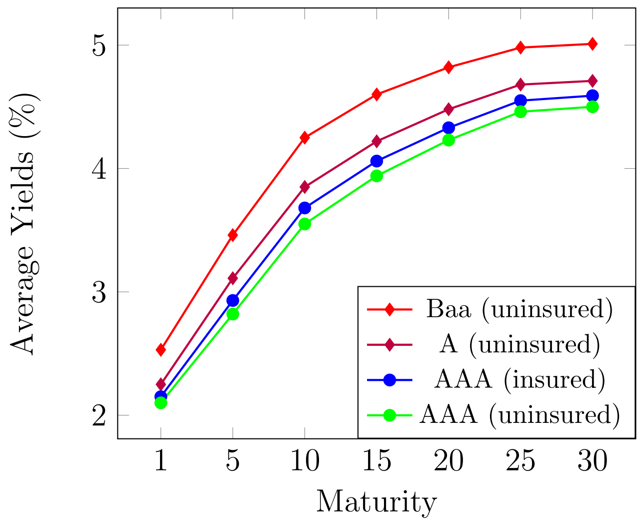
> 'Jersey City Medical Center has informed the Authority that it is undertaking a refunding project...made necessary by the fact that there is no longer a qualified bond insurer meeting the established rating standard to provide the wrap-around insurance for the bonds as required by the U.S. Department of Housing and Urban Development (HUD) for tax-exempt bonds." (New Jersey Health Care Facilities Financing Authority, August 23, 2012)

Other reasons relate directly to the benefit of insurance. Nanda and Singh (2004) argue that insurers, by preserving the tax-exempt status, function effectively as issuers of a sort of homogenized tax-exempt debt. Still other explanations draw upon either the ability of bond insurance to improve the liquidity of municipal bonds due to the credit improvement, or the ability for insurers to act as diversifiers of credit risk. More reasons for the existence of this market are provided in Wilkoff (2012) and Cirillo (2008).

According to the *California Debt and Investment Advisory Commission* (Angelides (2002)), issuers must evaluate bond insurers on a number of criteria, most notably whether or not they satisfy certain requirements as defined by the insurers. Several of these requirements include credit rating, net revenue coverage, additional bond tests, reserve requirements, additional hazard insurance, and capitalized interest. Other requirements are "soft" in nature, and include evaluations of economic, socio-political, structural, and historical factors, as well as demographics. Thus, there are presumably fixed costs for insurers in evaluating issuers.

Two notable features in Figure 1.6 are integral to the design of this study. The first is the dramatic fall in nationwide municipal insurance underwriting during the crisis, which has persisted. This episode includes the high profile failure of AMBAC Assured Corporation in November 2010 and the downgrade of MBIA from AAA to speculative grade. As a result of these failures, rating downgrades, and bankruptcies, the notional amount of municipal bonds enjoying bond insurance has fallen to one-tenth of pre-crisis levels. The source of this has primarily been due to bond insurer exposures in the subprime mortgage market, and not to the relatively healthy municipal market. In Ben Bernanke's memoir of the financial

Figure 1.8: Interest Cost Savings from Bond Insurance




This figure plots average yields in the municipal bond market for several credit ratings and insured status. AAA insured differs from AAA uninsured in that it summarizes the yield data for bonds where the rating may have been achieved through purchasing insurance, rather than achieved through an underlying credit rating. Data is collected from municipal bond yield indices for various credit rating categories from the Bond Buyer Municipal Market Data on January 4, 2005.

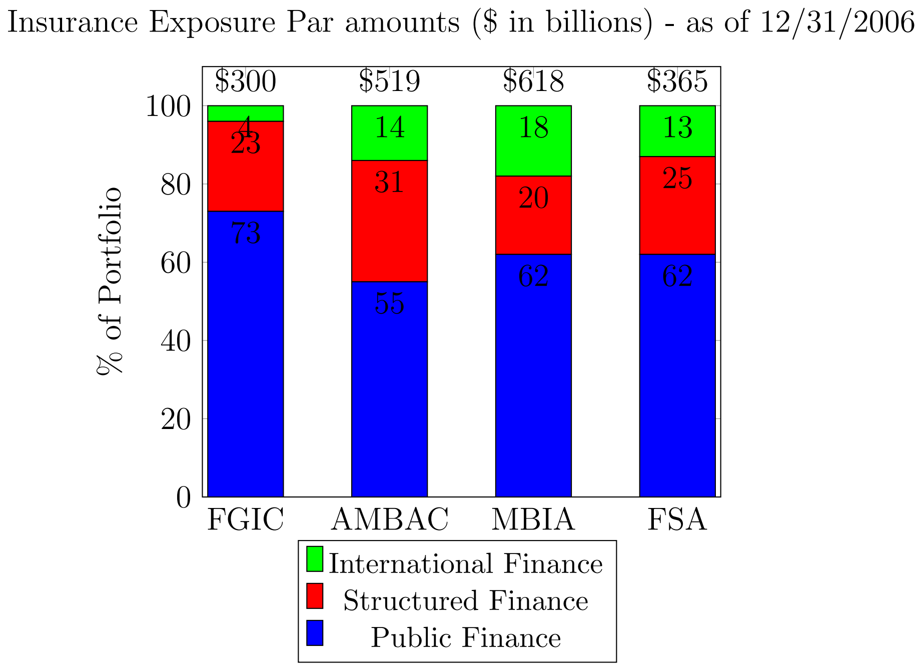
crisis, *The Courage to Act*, the former Federal Reserve Board Chairman recalls the fallout from insurer failures:

> "The monolines' guarantees of subprime securities were behind their ratings downgrades. Because the monolines also insured municipal bonds, however, investors grew wary of those bonds as well. We brainstormed about ways to help the muni market, which was largely an innocent bystander of the ongoing financial mayhem." (Bernanke (2015))

The second notable feature is that only two monoline insurers remained as sole survivors of this industry – Assured Guaranty Corporation and Financial Security Assurance, Incorporated (FSA). FSA was in fact sold to assured in July 2009, which led to Assured becoming essentially the sole insurance provider in the municipal bond market.

Indeed, the exposure of many of the insurers, while mostly exposed to public finance issues, had sizable exposure to asset-backed securities. Figure 1.9 demonstrates that right up to before the crisis, AMBAC, FGIC, MBIA, and FSA all had sizable exposures to structured finance assets. This demonstrates that FSA, like others, was not immune to fallout from insuring subprime securities.

Figure 1.9: Insurance Exposures for Four Major Monolines




Source: FGIC Investor Relations (2006). This data gives the percent of each monoline portfolio exposed to three major fixed income areas: international finance, structured finance, and public finance.

## 1.3 Data and Methodology

### Data

The municipal bond issuance data is from SDC Platinum, and consists of all municipal bond issues with 8 digit CUSIPs issued from 2003-2016. The data includes bonds from all US state, county, city, township, school district, and special authority issuers, and both general obligation and revenue bonds are included. Issues that do not contain face values, maturity dates, or issue dates were dropped, with the record treated as missing. This is a common occurrence as bond information seems to have been recorded on an issue-basis. As a feature peculiar to the municipal bond market, each bond issue can frequently contain up to 20 bonds, each maturing on successive dates. A shortcoming of this data is that not all bonds in a series will be insured, even if the data source notes that the issue carries bond insurance. Several issuers, albeit a small number, are also insured by multiple insurers. Thus an assumption made here is that the data provides a good enough proxy for the amount of insurance to which each local government is exposed.

The local government spending data comes from the US Census of Local Government Finances. I use data from 2002-2014. The US Census surveys finances for all local governments every 5 years, for years ending in 2 or 7. In between years, a balanced sample of

Table 1.1: Summary Statistics For Each US Census Survey Period

| Census Period | No. of Governments |
|---|---|
| 1997-2001 | 16,819 |
| 2002-2006 | 17,432 |
| 2007-2011 | 20,183 |
| 2012-2014 | 20,246 |

This table gives an account of the number of local government finances surveyed in each five-year US Census period. Source: US Census.

Table 1.2: Summary Statistics - Municipal Bonds, by Issuer Type

| Total | Face Value | ($ millions) | |
|---|---|---|---|
| Issuer Type | Mean | St. Dev. | N |
| County | 2.819 | 32.901 | 132,500 |
| City | 1.400 | 7.377 | 533,405 |
| School District | 1.638 | 12.690 | 513,624 |
| | 1.663 | | 1,179,529 |

| Pre-2008 | Face Value | | |
|---|---|---|---|
| Issuer Type | Mean | St. Dev. | N |
| County | 2.588 | 14.545 | 45,989 |
| City | 1.390 | 6.834 | 175,959 |
| School District | 1.780 | 17.786 | 175,271 |
| | 1.703 | | 397,219 |

This table gives an account of the amount of bonds issued by issuer type (county, city, or school district). The top table shows the entire time series from 2003-2016. Source: SDC Platinum.

roughly 30% of full population are surveyed. To keep the panel as balanced as possible while maintaining sample size, I only examine local governments that are repeatedly surveyed so that I maintain a balanced panel for all five-year periods. The number of governments per five-year period are provided in Table 1.1. The merge process between these two data sources is applied through a combination of exact and fuzzy string matching on district names.

Summary statistics for the municipal bond issuance data and government finances data are given in Tables 1.2 and 1.3. The tables are subdivided by the three major government issuer types – counties, cities, and school districts. Summary statistics show that these three types of issuers are comparable, the data not over-representing any single type.

Table 1.3: Summary Statistics - Local Government Finances, by Issuer Type

| (2007) | Expenditures | | |
|---|---|---|---|
| Issuer Type | Mean | St. Dev. | N |
| County | 294.891 | 859.859 | 987 |
| City | 185.504 | 1,933.789 | 2,049. |
| School District | 45.680 | 167.261 | 8,313 |
| | 92.598 | | 11,349 |

| (2007) | Total Debt Out. | | |
|---|---|---|---|
| Issuer Type | Mean | St. Dev. | N |
| County | 238.756 | 735.046 | 987 |
| City | 227.326 | 2215.830 | 2,049 |
| School District | 32.992 | 137.099 | 8,313 |
| | 85.973 | 976.989 | 11,349 |

This table gives an account of the amount of expenditures and total debt outstanding by issuer type (county, city, or school district) as of 2007. Source: US Census.

## Methodology and Causal Interpretation

## Variable Definitions

Dollar-value outcome variables from the US Census include all municipality and annual tax receipts, expenditure information, which includes capital outlays such as highway, school, sewage, and all construction expenditures. Log values are taken for all variables.

## Quasi-Experiment Approach

The approach I employ is a quasi-experiment that isolates the effect of being exposed to a government-rescued monoline insurer. First, I only use the population of issuers who have 100% of pre-crisis debt insured, so that there are no selection issues. Then, I form a control group and treatment group to allow for a difference-in-difference evaluation of relevant outcomes. The control group consists of all such issuers who had no exposure to any rescued monoline insurers, and the treatment group consists of issuers who had any exposure to rescued monoline insurers – specifically FSA or Assured Guaranty. The event study specification is, for issuer $i$ in month $t$, defining the year dummy $D_{ist} = \mathbf{1}(t = 2008 - s) \times \mathbf{1}(j \in \{\text{FSA}, \text{Assured}\})$,

$$y_{it} = \text{county}_{it} + \text{govtype}_{it} + \sum_{s=-5}^{s=4} D_{is}\beta_s + \mathbf{X}_i'\delta + \epsilon_{it},$$

where $y_{it}$ is any dependent outcome variable, $\text{county}_{it}$ are county-by-year fixed effects, $\text{govtype}_{it}$ are government-by-year fixed effects, and $\mathbf{X}_i$ a vector of controls.

The difference-in-difference specification is, given a Post x Treatment dummy, $D_{it} = \mathbf{1}(t \geq 2008) \times \mathbf{1}(j \in \{\text{FSA}, \text{Assured}\})$,

$$y_{it} = \text{county}_{it} + \text{govtype}_{it} + D_{it}\beta + \mathbf{X}_i'\delta + \epsilon_{it}.$$

The dependent variable will include measurement of debt issuance for local governments, as well as the associated capital investment outcome variables. Both specifications include county-by-year and government type-by-year fixed effects. This is to account for whether year-by-year variation in counties or in financing issues for specific government types can partially explain the regression outcomes. This fixed effect approach is also used in Cornaggia et al. (2017). For the second specification, I also include terms for Control $\times$ Post, which are the control variables each multiplied by the post dummy, in the case that levels of the controlled variables matter for post-2008 outcomes.

## Exogeneity

I argue for plausible exogeneity of this approach on several grounds. The first is the fact that municipal governments have historically achieved low rates of default before and through to the time of the Great Recession. The failure of bond insurers thus was a result of over-exposure to structured products related to subprime mortgages (Moldogaziev (2013)).

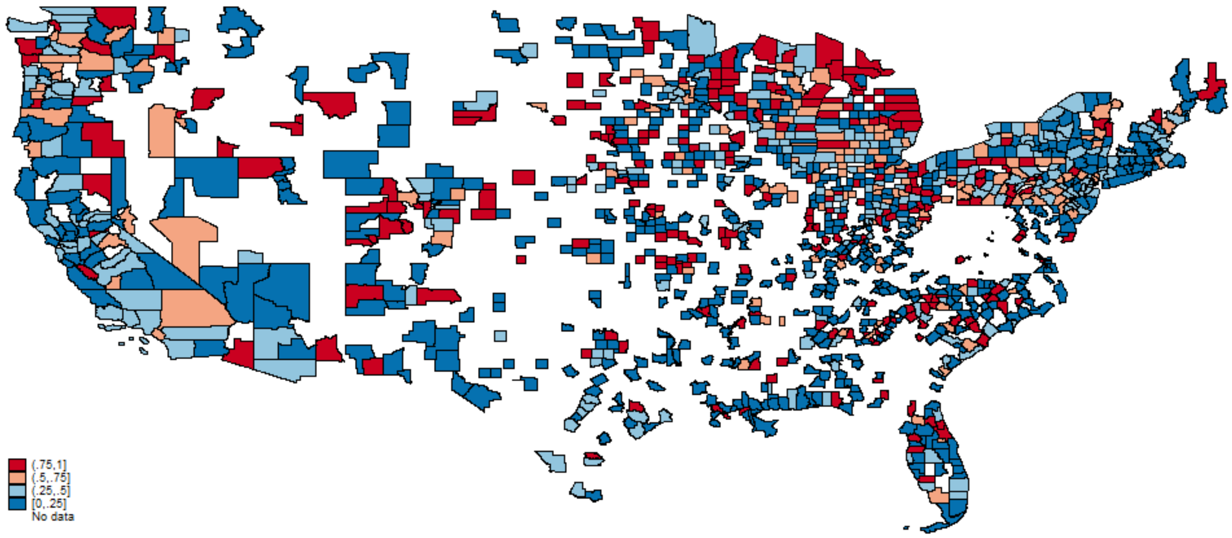
The second ground for exogeneity is that the eventual survival of a large insurer, FSA, was plausibly ex-ante unpredictable. In the US, almost all bond insurers were allowed to fail, including insurer FSA, which was owned by French bank Dexia. The subprime exposure of FSA spiraled Dexia into bankruptcy, before a deal was brokered by the French and Belgian governments to sell the FSA portfolio to Assured (Whitbeck (2013)). Other US insurers faced no such guarantee.

The third argument for exogeneity is the expectation that monoline bond insurers diversify their portfolios geographically. It is possible that insurers may possess some general focus or specialization on exposure to certain geographies or issuer types, but controlling for these, the issuers should still operate to diversify their holdings. If this is to be believed, then one should expect that the over-exposure to safe insurers was, in some sense, orthogonal to local economic conditions.

Figure 1.10 shows the fraction of pre-crisis bonds insured by FSA or assured, aggregated by county, red being the most exposed. The figure demonstrates that FSA and Assured have more exposure to some counties than to others, but otherwise are diversified by county within states that are exposed.

Balance tables for several observables for the control and treatment groups and provided in Table 1.4. These show that between treatment and control, there are a few significant differences between population outcomes, notably in important government financial figures. However, Table 1.5 provides the similar tables after also conditioning on the three government

Figure 1.10: Fraction of Pre-crisis Bonds Insured by FSA or Assured




This figure gives the percentage of exposure to FSA or Assured in each county in the United States prior to 2008. Red indicates 75% to 100% of all bonds issued in a county being insured by FSA or Assured.

types: county, city and school district. Once this is done, the within-type differences in financial figures – such as revenues, taxes, and pre-crisis debt issuance amounts – are no longer significant. This leaves differences in several characteristics that can be controlled for.

Table 1.4: Balance Table for Treatment and Control Groups

| | Control | Treatment | Difference | se |
|---|---|---|---|---|
| Mean Rating | 23.18 | 23.43 | -0.248** | 0.0843 |
| Fraction Unrated | 0.0252 | 0.0189 | 0.00630 | 0.00338 |
| Mean Duration | 4520.7 | 4452.2 | 68.52 | 51.06 |
| Pre-crisis Debt Issued ($mils) | 28.40 | 37.43 | -9.028** | 2.968 |
| Fraction of GO Bonds | 0.780 | 0.871 | -0.0903*** | 0.0118 |
| Fraction of Callable Bonds | 0.903 | 0.869 | 0.0340*** | 0.00965 |
| Population | 53320.7 | 27006.8 | 26313.9*** | 6377.2 |
| Revenue ($thous) | 107638.9 | 85464.5 | 22174.4* | 11043.2 |
| Tax Revenue ($thous) | 37998.7 | 29245.0 | 8753.7** | 3360.4 |
| Property Taxes ($thous) | 27723.7 | 24070.2 | 3653.5 | 2522.2 |
| Expenditures ($thous) | 105508.5 | 86386.2 | 19122.3 | 10674.8 |
| Capital Outlays ($thous) | 14860.9 | 11485.9 | 3375.0* | 1682.7 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table gives a balance table on several outcome variables on the differences between control and treatment groups.

Table 1.5: Balance Table for Treatment and Control Groups, by Issuer Type

| | (1)<br>Difference(county) | (2)<br>Difference(city) | (3)<br>Difference(school) |
|---|---|---|---|
| Mean Rating | 0.468* | -0.112 | -0.426*** |
| Fraction Unrated | -0.0245** | 0.000645 | 0.0152** |
| Mean Duration | 266.8 | 450.7*** | -85.85 |
| Pre-crisis Debt Issued ($mils) | -37.64* | -7.415* | -5.593 |
| Fraction of GO Bonds | -0.00742 | -0.00940 | -0.0927*** |
| Fraction of Callable Bonds | 0.0215 | 0.0669*** | -0.000923 |
| Population | -50045.4 | 2210.0 | 476.7 |
| Revenue ($thous) | -118492.0 | 4315.0 | -49.15 |
| Tax Revenue ($thous) | -32702.9 | 6657.1 | -755.0 |
| Property Taxes ($thous) | -27134.6 | 2720.8 | -622.3 |
| Expenditures ($thous) | -131289.7 | 2341.4 | -690.2 |
| Capital Outlays ($thous) | -17445.5 | 5575.0 | -828.8 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table gives a balance table on several outcome variables on the differences between control and treatment groups, after separating the population by government issuer type. These outcome variables are the same given in Table 1.4.

## 1.4 Results

### Debt Issuance

Event study coefficients for new debt issued are plotted in Figure 1.11. The plot depicts year-by-year coefficients for years 2005-2014, including 90% confidence intervals. The plot shows no clear pre-trend, and a positive jump at the time of the event, with an eventual recovery of differences around year 2011. The same event study specification, but for log total debt as an outcome, is depicted in Figure 1.12. For new issuances, the data is a flow variable from the bond issuance data. For log total debt, the data is a stock variable from the US Census. Both figures confirm the lack of a pre-trend, with an increase in debt levels during and after the financial crisis. This event study also implies a roughly 10% to 50% increase in debt levels for the treatment group.

Table 1.12 presents the difference-in-difference regression estimates. The coefficient on Treatment $\times$ Post implies that new bond issuances increased by 20.3% after the crisis, for the treatment group relative to the control group. Total debt increased by 25.4%, and interest payments by more. The difference is sizable and significant.

In order to justify that switching to a new insurer is costly for issuers in the treatment group, I show that the fraction of new debt insured for the treatment is persistent. Figure 1.13 shows that, by design, there is no difference between fraction of debt insured among issuers of both groups, as they are both selected to be 100%. After 2008, the treatment group is demonstrably more likely to issue insured debt. Table 1.13 reveals this amount

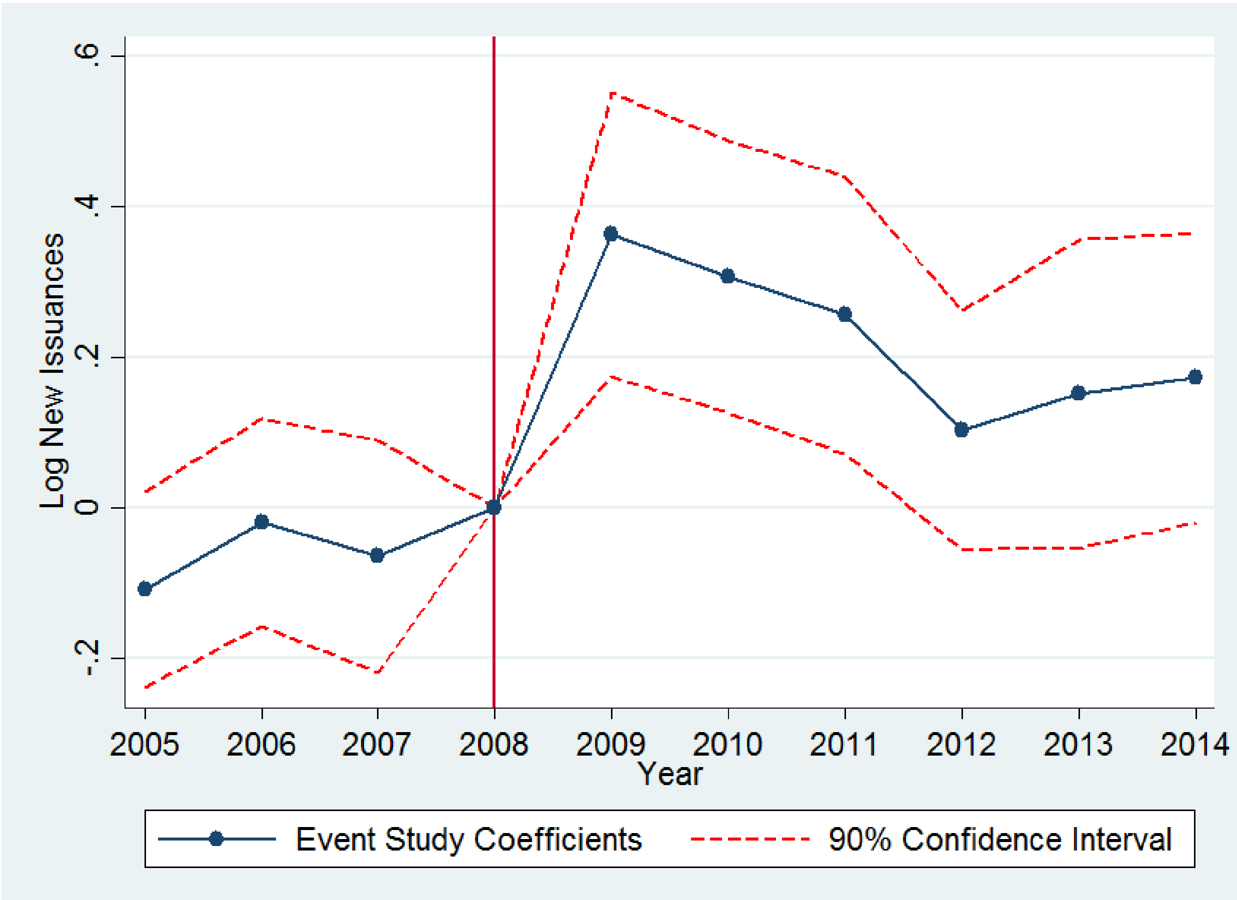
Table 1.6: Difference-in-Difference Regressions: Log Issuance/Log Debt

| | Log New Debt | Log Total Debt | Log Interest Payments |
|---|---|---|---|
| Treatment x Post | 0.203*** | 0.254*** | 0.282** |
| | (0.071) | (0.095) | (0.121) |
| County x Year FE | Yes | Yes | Yes |
| Local Government Type x Year FE | Yes | Yes | Yes |
| Issuer Controls | Yes | Yes | Yes |
| Issuer Controls x Post | Yes | Yes | Yes |
| R-squared | 0.33 | 0.34 | 0.34 |
| N | 8,748 | 5,974 | 5,974 |

Standard errors in parenthesis
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents difference-in-difference regression coefficients for log new bond issuances, log total debt, and log total interest payments. Log new bond issues are calculated by aggregating the face value of bond issues for each issuer by year. Log total debt is calculated by aggregating the total value of outstanding debt for all issuers by year. Log interest payments are calculated by adding all interest payments on outstanding debt for all issuers by year. Standard errors are clustered at the county level.

Figure 1.11: Event Study: New Municipal Bond Issuances



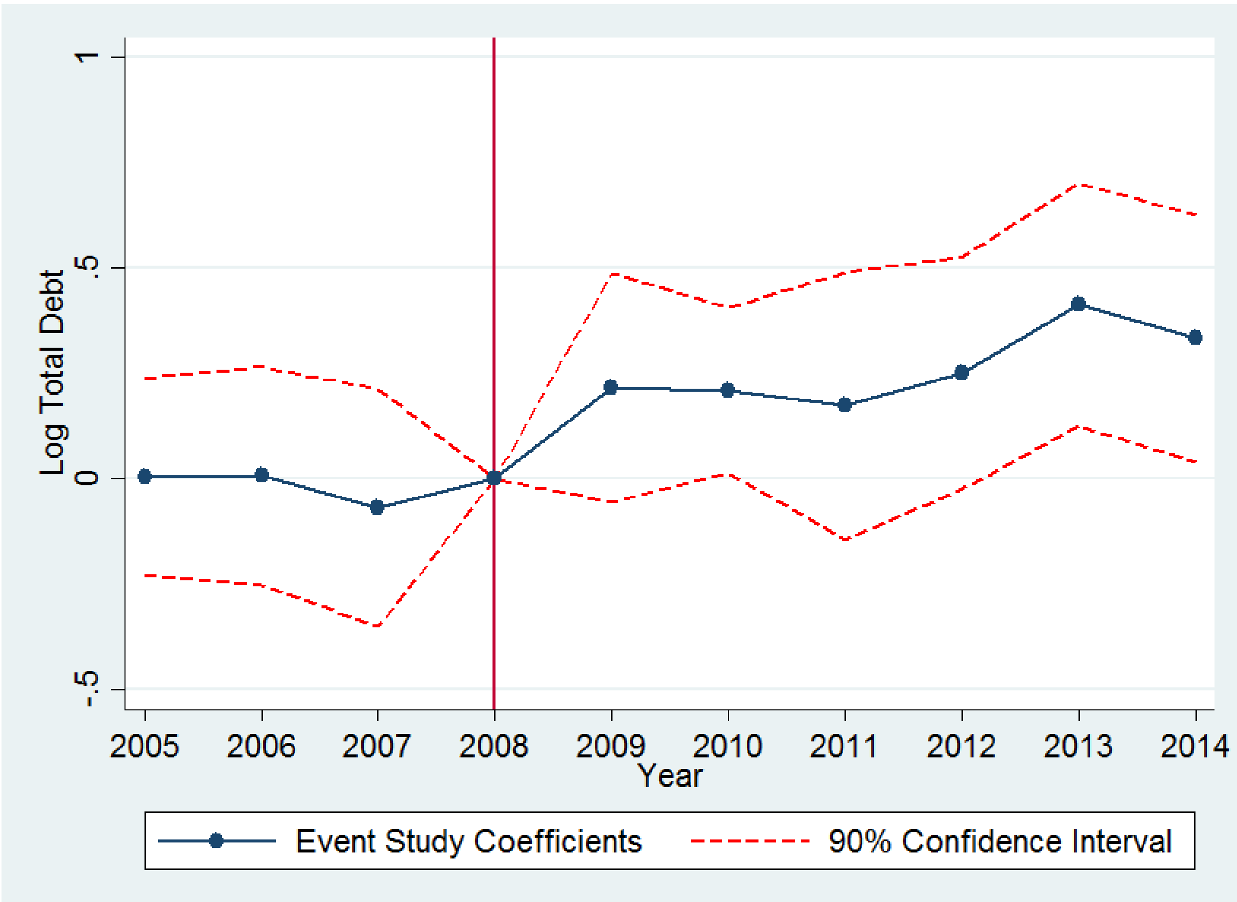

This figure presents event study coefficients for log new bond issuances. The measure is calculated by aggregating the face value of bond issues for each issuer by year.

to be 3.7% and statistically significant. This is a large amount, considering that the total fraction of post-crisis debt insured is around 5-10%.

## Yields

I present event study coefficient estimates for issuer bond yields in Figures 1.14 and 1.15. Given issuance results in the previous section, we should expect yields to fall for the treatment group, which is confirmed in both event studies. Again, no identifiable pre-trend exists, with a visible drop in yields after 2008. The former explains that new issuances are issued with lower yields on new issuances, and the latter should reveal that yields on existing debt in the treatment group are lower. The difference-in-difference estimates are -3.3% for offering yields and -3.5% for secondary market yields. These magnitudes are in line with post-crisis yield

Figure 1.12: Event Study: Log Total Debt



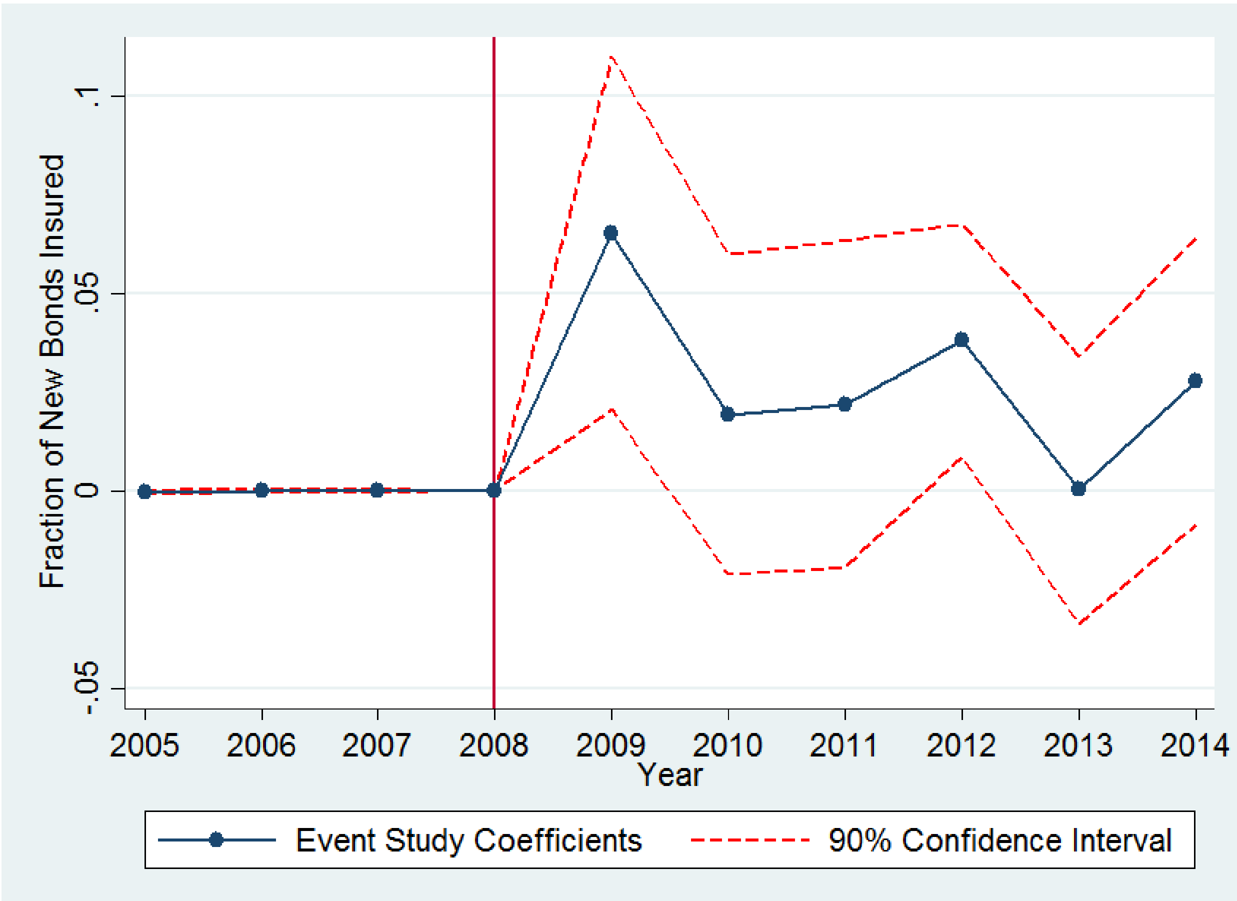

This figure presents event study coefficients for log total debt. The measure is calculated by aggregating the total value of outstanding debt for all issuers by year.

spreads presented in Figure 1.5, which are much larger in magnitude than those suggested by the benefits of insurance as estimated in Figure 1.8.

## Capital Investment Outcomes

Event studies for log total capital outlays are presented in Figure 1.16. This figure exhibits a slight negative pre-trend, but nonetheless an observable post-2008 shock, and eventual recovery, which mimics the pattern as seen with log bond issuance in Figure 1.11. Difference-in-difference regression estimates are given in Table 1.10. Issuers in the treatment group relative to the control group spend 25.3% more in capital outlays and 27.8% more on log construction. The former is significant at the 1% level and latter at the 10% level. Spending thus roughly increases one-to-one with debt issuance, highlighting the importance of debt

Figure 1.13: Event Study: Fraction of New Debt Insured




This figure presents event study coefficients for the fraction of new debt insured. The outcome variable is measured by dividing the amount of insured debt issued in a given year, divided by the total debt issued. The coefficients and confidence intervals are zero before the pre-period, as all issuers in the sample population have fraction of insured debt equal to 100% prior to 2008.

Table 1.7: Difference-in-Difference Regressions: Fraction of New Debt Insured

| | Fraction of new debt insured |
|---|---|
| Treatment x Post | 0.037*** |
| | (0.013) |
| County x Year FE | Yes |
| Local Government Type x Year FE | Yes |
| Issuer Controls | Yes |
| Issuer Controls x Post | Yes |
| R-squared | 0.59 |
| N | 13,644 |

Standard errors in parenthesis
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents difference-in-difference regression coefficients for the fraction of new debt insured. The outcome variable is measured by dividing the amount of insured debt issued in a given year, divided by the total debt issued. The coefficients and confidence intervals are zero before the pre-period, as all issuers in the sample population have fraction of insured debt equal to 100% prior to 2008. Standard errors are clustered at the county level.

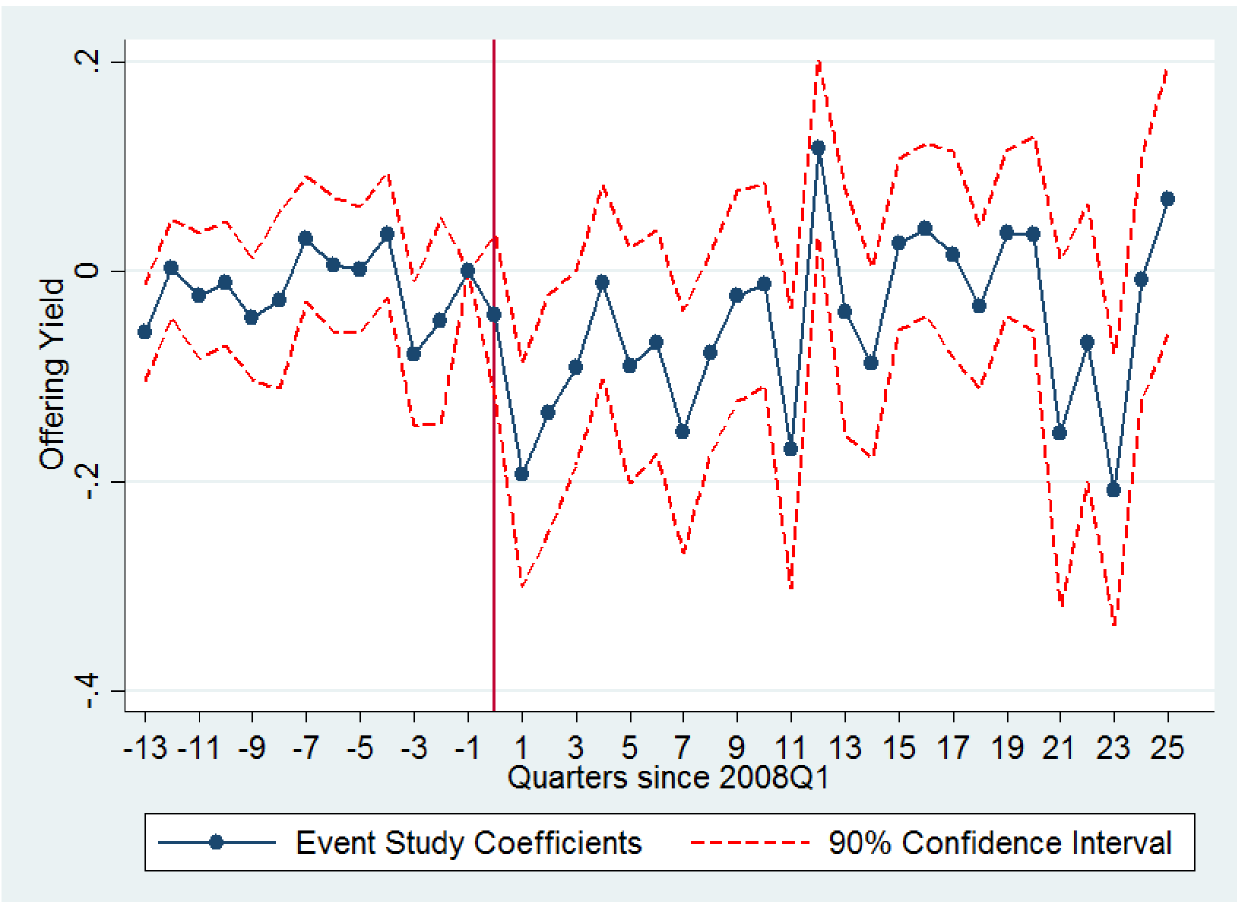
Table 1.8: Difference-in-Difference Regressions: Log Offering Yield

| | Offering Yield |
|---|---|
| Treatment x Post | -0.033*** |
| | (0.010) |
| County x Qtr FE | Yes |
| Local Government Type x Qtr FE | Yes |
| Issuer Controls | Yes |
| Issuer Controls x Post | Yes |
| R-squared | 0.55 |
| N | 20,353 |

Standard errors in parenthesis
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

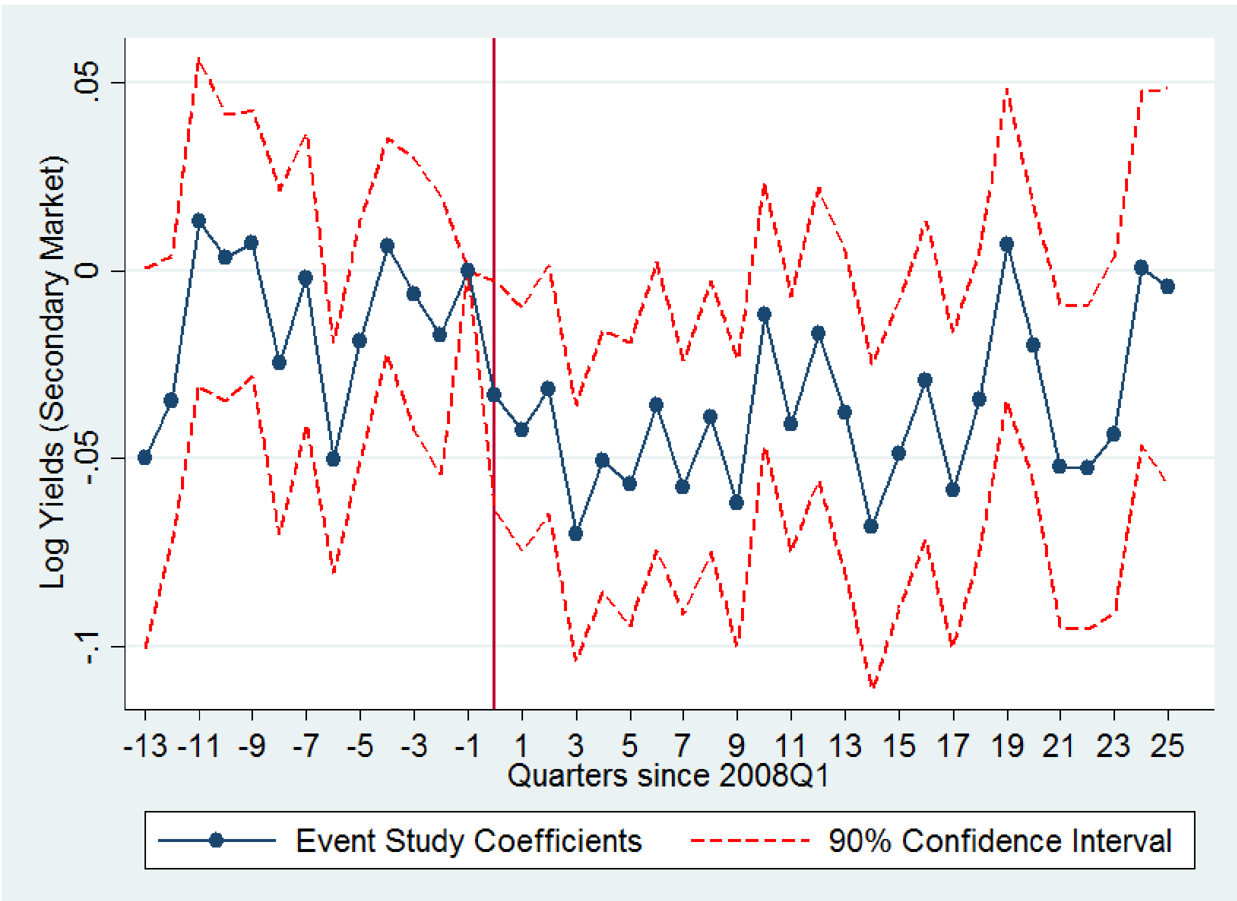
This table presents difference-in-difference regression coefficients for log offering yields. I compute quarterly average of log yields by aggregating all yields-to-maturity at time of offering for every bond issue, for each issuer. Event time is in quarters, with the event time equal to 2008Q1. Standard errors are clustered at the county level.

Figure 1.14: Event Study: Log Yields (Offering - Primary)




This figure presents event study coefficients for log offering yields. I compute quarterly average of log yields by aggregating all yields-to-maturity at time of offering for every bond issue, for each issuer. Event time is in quarters, with the event time equal to 2008Q1.

Figure 1.15: Event Study: Log Yields (Secondary Market)




This figure presents event study coefficients coefficients for log yields on bonds issued before the crisis. I compute quarterly average of log yields by aggregating all yields-to-maturity at time of offering for every bond issued prior to the event time, for each issuer. Event time is in quarters, with the event time equal to 2008Q1.

Table 1.9: Difference-in-Difference Regressions: Log Yield (Secondary Market)

| | Log Yield (Secondary Mkt) |
|---|---|
| Treatment x Post | -0.035*** |
| | (0.011) |
| County x Qtr FE | Yes |
| Local Government Type x Qtr FE | Yes |
| Issuer Controls | Yes |
| Issuer Controls x Post | Yes |
| R-squared | 0.23 |
| N | 98,459 |

Standard errors in parenthesis
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

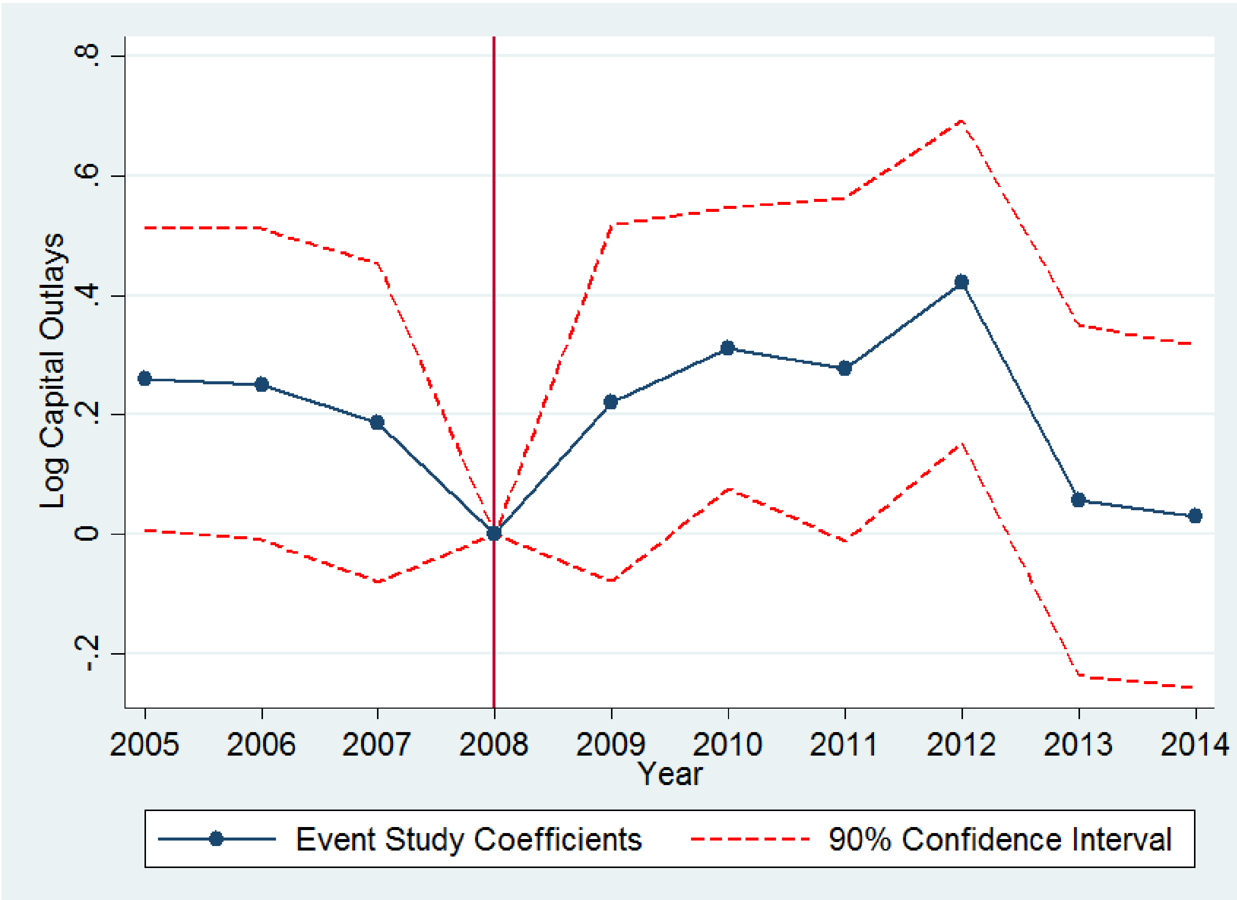
This figure presents difference-in-difference regression coefficients coefficients for log yields on bonds issued before the crisis. I compute quarterly average of log yields by aggregating all yields-to-maturity at time of offering for every bond issued prior to the event time, for each issuer. Event time is in quarters, with the event time equal to 2008Q1. Standard errors are clustered at the county level.

frictions.

Not all categories of spending have significant coefficients on the Treatment $\times$ Post variable in difference-in-difference regressions. Two representative categories of capital outlays are elementary and secondary education outlays, and highway outlays. The former shows a high statistical significance while the latter is not significant at the 10% level. A candidate explanation for this is that highway spending is highly subsidized by the federal government, whereas school spending is not. In order to analyze this in more depth, coefficients for all such univariate difference-in-difference estimates are given in Table 12 and listed in descending order by p-value. Categories significant at the 10% level are gas, electrical, water utilities, airport outlays, elementary and secondary education outlays, and housing development. By plotting this against a measure of federal subsidy strength in each of these categories, Figure 1.17 makes clear that having a low measure of federal subsidy is necessary for a category to observe a statistically significant coefficient. This measure is calculated for the year 2012 by dividing the total earmarked federal grants in each category by the total local capital outlays.

For higher levels of the ratio, four categories of estimates are not significant, namely those for waste, highway expenditures, transport utilities, and natural resources. At higher levels, the coefficients are also slightly more precisely estimated, but it is unclear whether this is a statistical artifact of the measurements or a meaningful relationship. Also, the remaining minor categories of expenditures, all with zero to no federal support, also exhibit estimates with low statistical significance, namely health spending, parks and recreation, welfare, corrections, and sewerage. This evidence is suggestive that low federal spending

Figure 1.16: Event Study: Log Capital Outlays




This figure presents event study coefficients for log capital outlays. The measure is calculated by aggregating the total value of capital outlays for all issuers by year.

for certain classes of capital outlays will exacerbate reactions to financial frictions in the municipal bond market.

Table 1.10: Difference-in Difference Regressions: Log Spending

| | Log Capital Outlays | Log Construction |
|---|---|---|
| Treatment x Post | 0.253*** | 0.278* |
| | (0.089) | (0.143) |
| County x Year FE | Yes | Yes |
| Local Government Type x Year FE | Yes | Yes |
| Issuer Controls | Yes | Yes |
| Issuer Controls x Post | Yes | Yes |
| R-squared | 0.35 | 0.27 |
| N | 5,974 | 5,974 |

Standard errors in parenthesis
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table presents difference-in-difference regression coefficients for log capital outlays and log construction. Log capital outlays is measured by aggregating the total value of capital outlays for all issuers by year. Log construction is measured by aggregating the total value of construction for all issuers by year. Note that part of the value of capital outlays consists of construction. Standard errors are clustered at the county level.

Table 1.11: Difference-in Difference Regressions: Log Education and Highway Spending

| | Log Education Outlay | Log Highway Outlay |
|---|---|---|
| Treatment x Post | 0.428** | 0.116 |
| | (0.195) | (0.103) |
| County x Year FE | Yes | Yes |
| Local Government Type x Year FE | Yes | Yes |
| Issuer Controls | Yes | Yes |
| Issuer Controls x Post | Yes | Yes |
| R-squared | 0.66 | 0.79 |
| N | 5,974 | 5,974 |

Standard errors in parenthesis
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

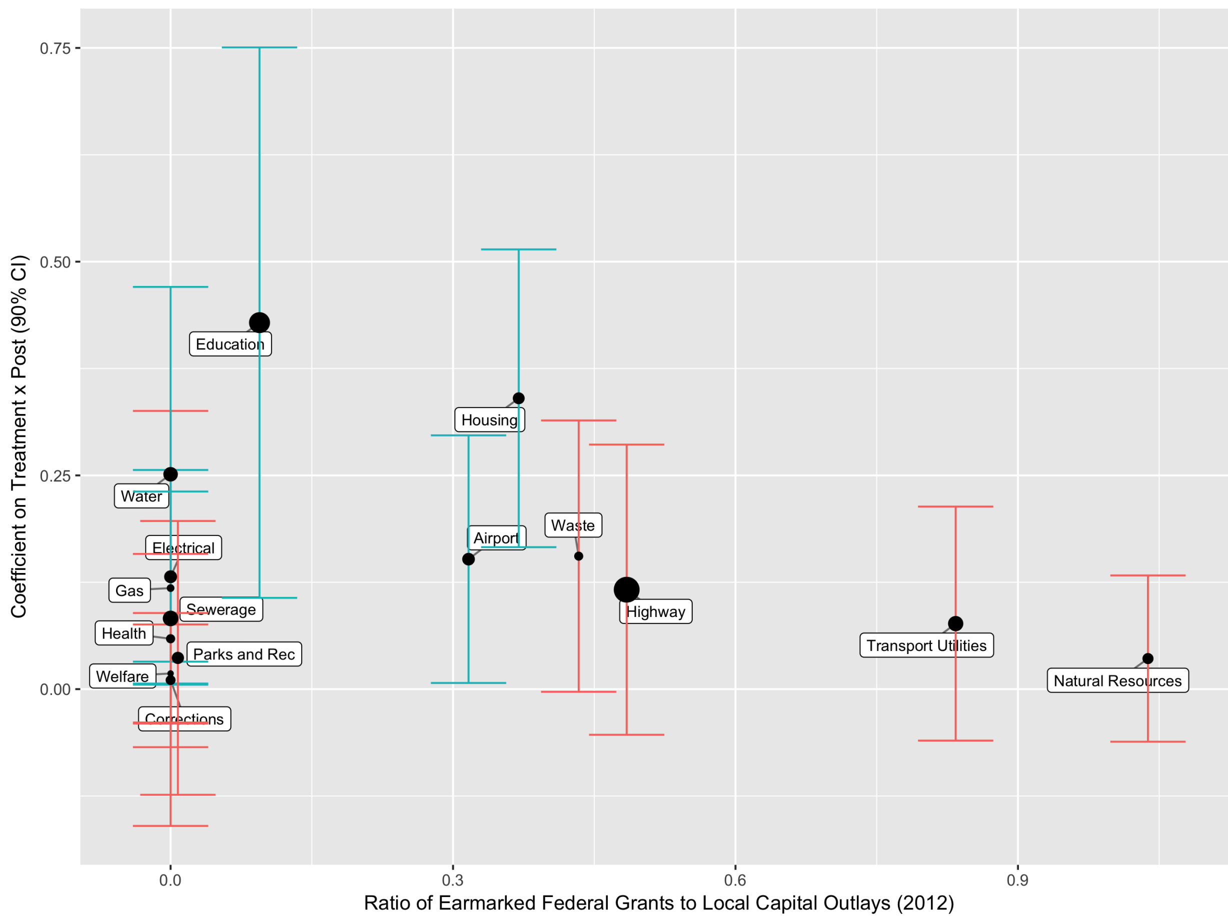
This table presents difference-in-difference regression coefficients for log elementary and secondary education spending and log highway spending. Each category of spending is aggregated at the issuer level by year. Standard errors are clustered at the county level.

Table 1.12: Log Regression Coefficients on Treatment × Post

| | b | se | t | p | 90% CI lb | 90% CI ub |
|---|---|---|---|---|---|---|
| Corrections | 0.0106 | 0.0476 | 0.2220 | 0.8245 | -0.0679 | 0.0890 |
| Parks Rec | 0.0365 | 0.0971 | 0.3761 | 0.7070 | -0.1236 | 0.1967 |
| Sewerage | 0.0827 | 0.1472 | 0.5617 | 0.5746 | -0.1600 | 0.3254 |
| Welfare | 0.0183 | 0.0348 | 0.5266 | 0.5988 | -0.0390 | 0.0756 |
| Nat Resources | 0.0357 | 0.0590 | 0.6061 | 0.5448 | -0.0615 | 0.1330 |
| All Utilities | 0.0766 | 0.0830 | 0.9233 | 0.3564 | -0.0602 | 0.2135 |
| Health | 0.0589 | 0.0602 | 0.9797 | 0.3279 | -0.0403 | 0.1581 |
| Highways | 0.1163 | 0.1029 | 1.1299 | 0.2593 | -0.0534 | 0.2860 |
| Waste | 0.1556 | 0.0963 | 1.6162 | 0.1069 | -0.0032 | 0.3143 |
| Gas | 0.1182 | 0.0685 | 1.7249 | 0.0854 | 0.0052 | 0.2312 |
| Air | 0.1520 | 0.0878 | 1.7309 | 0.0843 | 0.0072 | 0.2969 |
| Electrical | 0.1314 | 0.0757 | 1.7354 | 0.0835 | 0.0065 | 0.2563 |
| Water | 0.2513 | 0.1329 | 1.8906 | 0.0595 | 0.0321 | 0.4705 |
| Education | 0.4287 | 0.1952 | 2.1955 | 0.0287 | 0.1067 | 0.7506 |
| Housing | 0.3402 | 0.1056 | 3.2209 | 0.0014 | 0.1660 | 0.5144 |

This table presents difference-in-difference regression coefficients for 15 categories of capital outlay spending for local governments in the sample. The coefficient $b$ is the estimate on the Treatment × Post coefficient for the difference-in-difference specification, with the same regressors as those employed in Table 1.11. Standard errors, t-statistic, p-value and 90% confidence intervals of estimates are presented. Results are ordered decreasing in p-value. Standard errors are clustered at the county level.

Figure 1.17: Plot of 90% CIs of difference-in-difference coefficients versus Ratio of earmarked federal grants/local capital outlays, by outlay category (2012)




This figure plots the 90% confidence intervals of difference-in-difference coefficients given in Table 1.12. The x-axis measure is calculated for year 2012 by dividing federal reported earmarked grants towards each of the spending categories by the total capital outlays in the US Census. The confidence intervals lying above zero are marked in blue. These include gas utilities, airport outlays, electrical utilities, water utilities, educational outlays, and housing development outlays.

## 1.5 Conclusion

In this paper, I show that local governments with more exposure to a government-rescued monoline insurer FSA and its purchaser Assured Guaranty had better borrowing outcomes and spending in capital investments, relative to others than had exposure to non-rescued insurers. I show this using a quasi-experimental approach, forming a treatment and control group of local governments. Event-study coefficients and difference-in-difference regression estimates show that issuers in the treatment group issue more bonds in the years after 2008, and also spend more on capital investments. The effect is significant for categories of public goods for which federal resources are scarce – specifically in education, housing development, and some general utilities.

This suggests that patterns of federal support for local public goods may exacerbate financial frictions that affect local governments. This is a cost of decentralization that may be underestimated both in economic models of public good provision and in federal government policy deliberations. To the extent that federal policy leaves some local government projects more susceptible to macroeconomic shocks, this implies that it may be welfare improving for policy to account for unintended effects of unequal distribution of public good subsidies.

# Chapter 2

# Index Comovement and Informed Trading

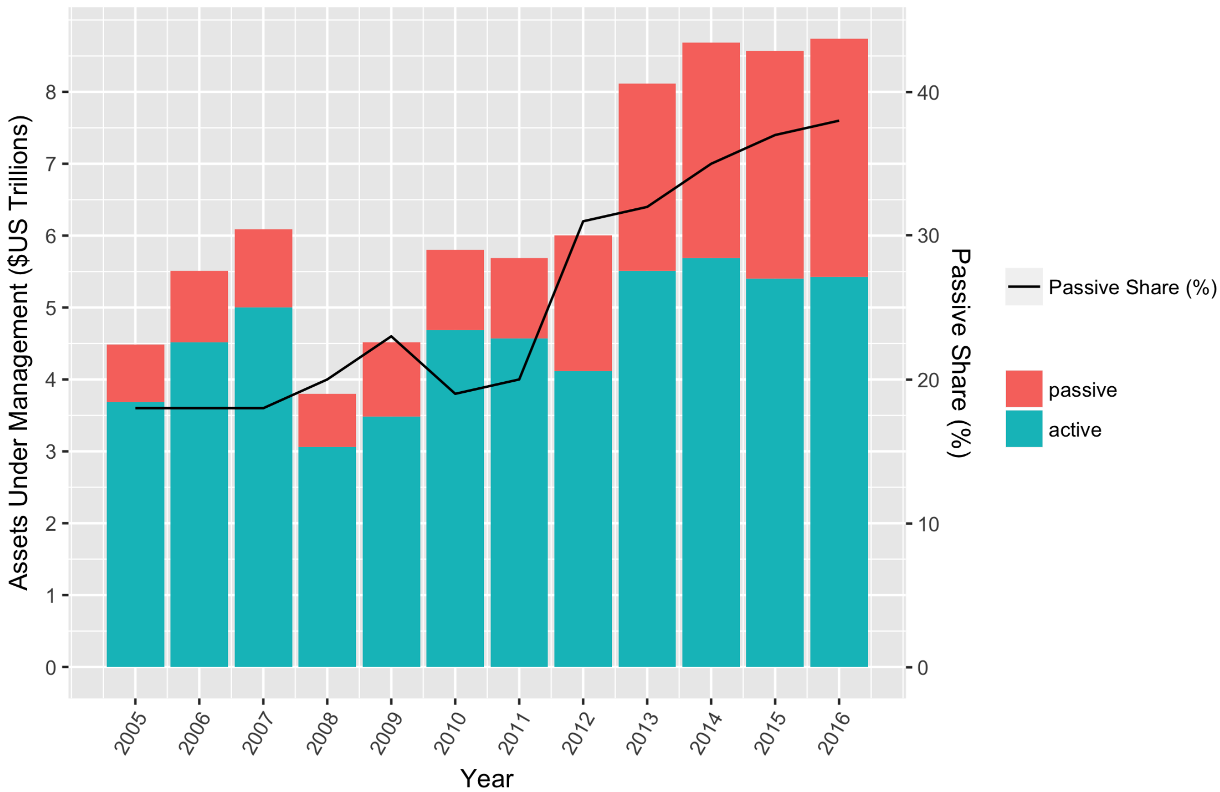
## 2.1 Introduction

Since the 1970s, standard finance theories, such as Malkiel and Fama (1970), suggest that markets are relatively efficient, which implies an optimal investment strategy of holding the market portfolio. Historical out-performance of passive benchmarks relative to open-ended mutual funds (Jensen (1968)) thus opens up an "active management puzzle", questioning why active managers are still highly compensated despite poor average performance. Rational arguments such as Berk and Green (2004) suggest there can be active rents, but that these are driven away competitively by capital flows into the most profitable funds.

A potentially large source of active rents are informational externalities arising from the presence of large passive investment capital itself, such as in exchange-traded funds (ETFs). As of 2016, Morningstar estimates that total assets under management in US equity-based passive index mutual funds and ETFs have exceeded \$3 trillion, which represents 38% of all US equity fund assets (Figure 2.1). Since most investors in these products would not need to remain informed about idiosyncratic risks, there can still exist active rents in the sense of Grossman and Stiglitz (1980). That is, passive investment capital can affect the incentives of active managers through directly changing the information environment itself.

I bring theory and evidence to bear on this. In this paper, I study how the mechanism of *comovement* – a significant source of price distortion caused by index investing – affects price efficiency and the incentives to acquire information. Using a model of informed and passive trading in a two asset Glosten and Milgrom (1985) setting, I describe conditions under which price efficiency is improved or exacerbated. The primary contribution of this paper is to show that passive investors have attributes of both uninformed and informed traders. They are uninformed in that they do not trade on privately acquired information; but they are informed in that their presence facilitates the diffusion of such information across asset prices. Secondly, I provide empirical evidence using a regression discontinuity design with

Figure 2.1: Assets of all US Equity Passive and Active Mutual Funds and ETFs




This figure shows the total assets under management for passive and active mutual funds and ETFs, for US equity assets. The right-hand axis is the share of passive assets as a percentage of the total. Source: Morningstar.

Russell 2000 and Thomson Reuters 13-F participation data. I show that when interacted with measures of short-run earnings uncertainty, index ownership predicts less trading by institutional firms. This suggests that passive ownership affects how active investors respond to short-run uncertainty in firms.

I contribute to the literature on price distortions in equity markets and the various hypotheses on how asset demand curves slope downward. The central mechanism of this paper is excess comovement, which is the primary price distortion mechanism of index investing (Barberis et al. (2005), Da and Shive (2013), Broman (2014)). The empirical portion of this paper borrows from Chang et al. (2015), which uses the Russell index inclusion regression discontinuity quasi-experiment to demonstrate price shocks on included stocks. This paper analyzes the associated informational consequences.

This paper also contributes to a growing literature on the unintended effects of passive ownership. Several papers show that large passive ownership shares affect the structure under which efficient corporate governance manifests (Appel et al. (2016), Mullins (2014), Azar et al. (2015)). This paper speaks to a major disciplining device that is not direct corporate governance, but selling pressure by active managers (Edmans (2014)), which has indirect implications for corporate governance.

Lastly, this paper also relates to theoretical treatments of the active management puzzle. Gârleanu and Pedersen (2015) links the relationship between price efficiency and the costs of seeking active management. This paper takes seriously the participation of passive investors to derive price implications of homogeneous allocations. My setting predicts the existence of equilibria where an increase in index ownership improves price efficiency and decreases the activities of informed investors – a potential exacerbation of the asset management puzzle.

The outline of the paper is as follows: Section 2.2 presents the model setup; I describe the asset structure, and incentives for different types of investors. Section 2.3 derives the prices implied from sequential trading and implications for price efficiency. Section 2.4 describes comparative statics that reveal the information consequences for informed trader behavior. Section 2.5 describes the empirical test. Section 2.6 concludes.

## 2.2 Model Setup

### Asset Payoffs

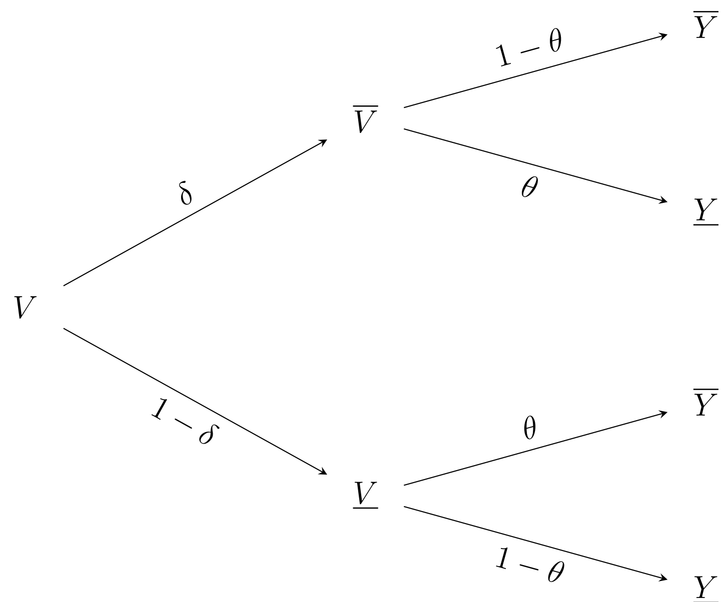
Consider the following economy, based on Glosten and Milgrom (1985). There are two assets, $V$ and $Y$, that can take on values $\underline{V}, \overline{V}$ and $\underline{Y}, \overline{Y}$, respectively.

For asset $V$, $\underline{V}$ occurs with probability $\delta$ and $\overline{V}$ with probability $1-\delta$. The payoff of asset $Y$ however depends on that of $V$: If $V = \underline{V}$ (or $V = \overline{V}$), then $Y = \underline{Y}$ (or $Y = \overline{Y}$) occurs with probability $1-\theta$. Conversely, if $V = \underline{V}$ (or $V = \overline{V}$), then the other options occur with probability $\theta$. In other words, if the return of $V$ is in the "high" state, then there is a probability $1-\theta$ that $Y$ is also in the high state. Therefore $\theta$ captures the correlation between assets $V$ and $Y$. We can think of $V$ as the market, and $Y$ of a payoff that has a large correlated component to $V$. A diagram of this payoff structure is given in Figure 2.2.

### Agents

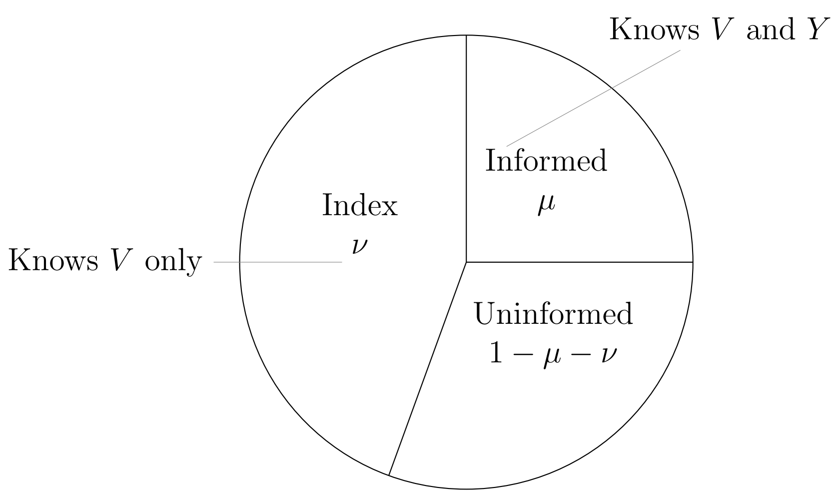
There are three types of traders in this economy: informed traders with mass $\mu \in [0,1]$, index traders with mass $\nu \in [0,1]$, and uninformed traders with mass $(1-\mu-\nu) \in [0,1]$. Informed traders have full information and know the realized values of both assets $V$ and $Y$. Index traders only know the realized value for asset $V$, and are limited to buying or selling at the same side in both assets. Uninformed traders know neither realized value, and don't have restrictions on buying and selling.

Figure 2.2: Probability diagram of contemporaneous asset payoffs for $V$ and $Y$ (per period)




This diagram gives a tree that shows the value realization probabilities of $V$ and $Y$. The probabilities for $Y = \overline{Y}$ and $Y = \underline{Y}$ depend on the value of $V$.

Figure 2.3: Agents and Information Structure




This diagram describes the composition of agents in the model and the information structure regarding which agents know the realizations for $V$ and $Y$.

### Sequential Trade

Trade occurs sequentially, and trades are for one unit only, occurring at each discrete time period $t$. Each type of trader submits either a buy or sell order based on her information set. Since uninformed traders have no information, they buy or sell each asset with independent probabilities of $\frac{1}{2}$. The market maker has infinite capital and sets bid and ask prices such that $B = \mathrm{E}[V|\mathrm{sell}]$ and $A = \mathrm{E}[V|\mathrm{buy}]$, or the expected value of $V$ conditional on the order being either a buy or sell. From the market maker's perspective, these prices can be solved using Bayes' rule. I assume that the market maker sets prices independently in both assets.

## 2.3 Trading

The trading restrictions for index traders simulates the existence of an "index" where traders will buy or sell a pool of assets. Even though informed traders have information that $Y$ might go in the opposite direction, they must trade against index traders who create correlated prices in the two assets. Thus, for example if $\theta = 1$, then index traders will never be correct about asset $Y$, and if $\theta = 0$, then the information set of informed and index traders are identical. Thus, the signal $Y$ available to informed traders is most "informative" at $\theta = 0.5$[1]

Considering these values for $\theta$, I will assume

**Condition 1** $\theta \leq 0.5$.

With Condition 1, this allows index traders to trade "correctly" most of the time.

### Conditional Probabilities and Prices

The bid and ask prices for asset $Y$ can be calculated using Bayes' rule. I omit an analysis for asset $V$, as index traders and informed traders have the same information set in asset $V$, so bid and ask prices for $V$ simplify to Glosten and Milgrom (1985).

**Proposition 1** *The conditional probabilities that any order is a buy or sell, given the realization of $Y$, are:*

$$\begin{aligned}
P(buy|\overline{Y}) &= \mu + \nu(1-\delta) + \frac{(1-\mu-\nu)}{2} \\
P(sell|\overline{Y}) &= \nu\delta + \frac{(1-\mu-\nu)}{2} \\
P(buy|\underline{Y}) &= \nu(1-\delta) + \frac{(1-\mu-\nu)}{2} \\
P(sell|\underline{Y}) &= \mu + \nu\delta + \frac{(1-\mu-\nu)}{2}.
\end{aligned}$$

[1]One can argue this, for example, by maximizing a Shannon entropy quantity of the form $H(\theta) = -\theta\log\theta - (1-\theta)\log(1-\theta)$.

These conditional probabilities are intuitive. The probability that an order originates from an index trader is $\nu(1-\delta)$ for buy orders and $\nu\delta$ for sell orders, as their order types completely depend on the realization of $V$, and not that of $Y$.

The total probabilities of the realizations for $Y$ are:

$$\begin{aligned}\overline{w} \equiv \mathrm{P}(Y = \overline{Y}) &= (1-\delta)(1-\theta) + \delta\theta \\ \underline{w} \equiv \mathrm{P}(Y = \underline{Y}) &= (1-\delta)\theta + \delta(1-\theta),\end{aligned}$$

which allows us to calculate prices.

**Proposition 2** *The bid price for asset $Y$ is*

$$\begin{aligned}B &= E[Y|sell] \\ &= \frac{\underline{Y}\mu\underline{w} + (\overline{w}\overline{Y} + \underline{w}\underline{Y})\nu\delta + (\overline{w}\overline{Y} + \underline{w}\underline{Y})\left(\frac{1-\mu-\nu}{2}\right)}{\mu\underline{w} + \nu\delta + \frac{(1-\mu-\nu)}{2}},\end{aligned}$$

*and the ask price is*

$$\begin{aligned}A &= E[Y|buy] \\ &= \frac{\overline{Y}\mu\overline{w} + (\overline{w}\overline{Y} + \underline{w}\underline{Y})\nu(1-\delta) + (\overline{w}\overline{Y} + \underline{w}\underline{Y})\left(\frac{1-\mu-\nu}{2}\right)}{\mu\overline{w} + \nu(1-\delta) + \frac{(1-\mu-\nu)}{2}}.\end{aligned}$$

The bid and ask prices expressions look comparable, but their comparative statics are quite different. The expected value characterization means that the price is a function that weighs the true asset value observed by the informed traders ($Y$), and a risk-neutral expectation of the price ($\overline{w}\overline{Y} + \underline{w}\underline{Y}$) observed by all others. Figure 2.4 illustrates the comparative statics with increasing the mass of index traders $\nu$. The bid and ask are decreasing in $\nu$, but the spread is increasing, where the bid is concave and the ask, convex.

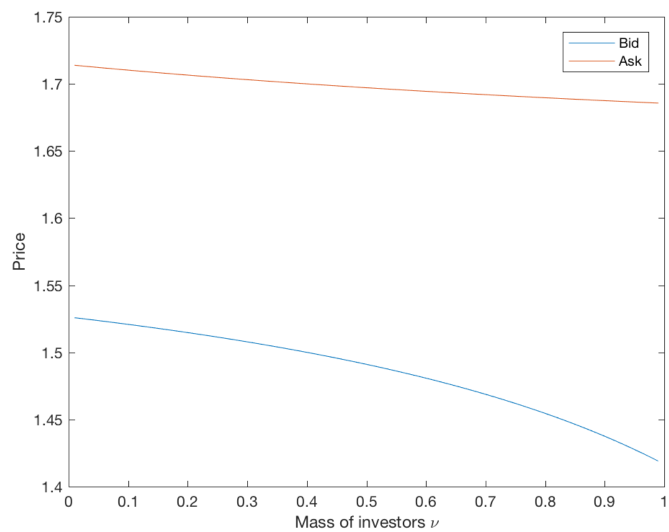
What gives rise to the asymmetry? There are two effects of increasing the mass of index traders, $\nu$. Index traders essentially behave like both informed and uninformed traders. In the way that they act like informed traders, they are more likely to increase the probability that an incoming order is a buy. This mimics the decision of informed traders, and in equilibrium an increase in $\nu$ causes the bid-ask spread to widen. Their behavior, however, mimics uninformed traders as well, so that the ask price also adjusts downward, away from the informed trader's valuation.

## Price Efficiency

The model has implications for price efficiency. Recall, the market maker sets bid and ask prices conditioning on the past history of orders so that:

$$\begin{aligned}B &= \mathrm{E}[Y|k+1 \text{ buys}, l \text{ sells}] \\ A &= \mathrm{E}[Y|k \text{ buys}, l+1 \text{ sells}],\end{aligned}$$

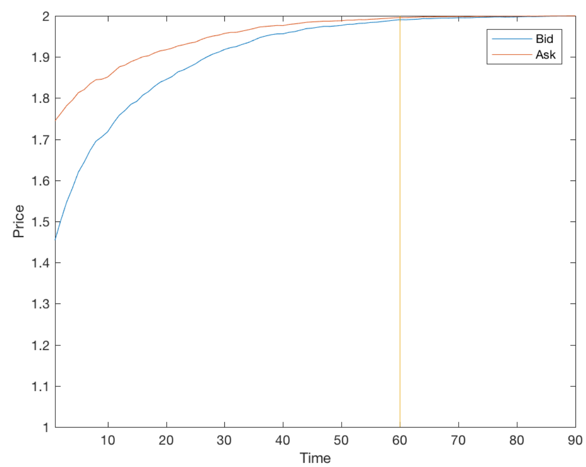
Figure 2.4: Bid-ask spread as a function of $\nu$




Parameters: $\mu = 0.2$, $\underline{Y} = 1$, $\overline{Y} = 2$, $\delta = \theta = 0.25$

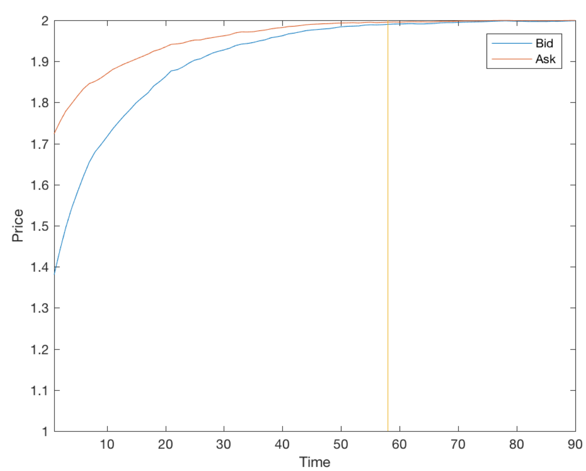
then one can parameterize the model and compare convergence times to the true realized asset value. One such comparison is captured in Figures 2.5 and 2.6. The average bid-ask prices converge to within 0.5% of the true realization of $Y$ sooner when $\nu$ is large. These dynamics are due to orders from informed and index traders being indistinguishable to the market maker. This result is counterintuitive – since the spread is larger with more index traders, and prices are more distorted, away from the true valuation. A sample of one simulation is given in Figure 2.7.

Figure 2.5: Average bid-ask for 500 simulations



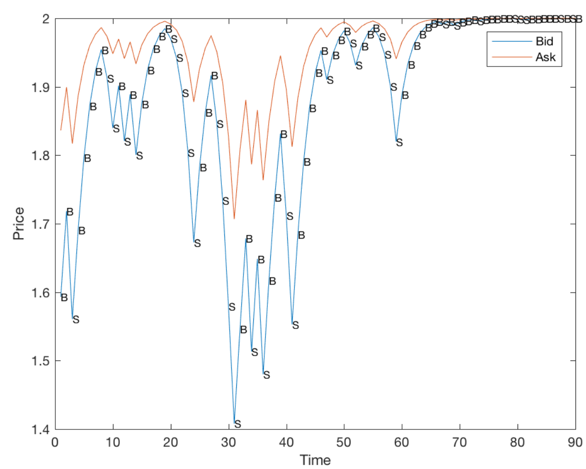

Parameters: $\nu = 0.2$, $\mu = 0.3$, $Y = \overline{Y} = 2$, $\delta = \theta = 0.25$

Figure 2.6: Average bid-ask for 500 simulations




Parameters: $\nu = 0.7$, $\mu = 0.3$, $Y = \overline{Y} = 2$, $\delta = \theta = 0.25$

Figure 2.7: Average bid-ask for 1 simulation




Parameters: $\nu = 0.2$, $\mu = 0.3$, $Y = \overline{Y} = 2$, $\delta = \theta = 0.25$

## 2.4 Comparative Statics

The analysis here focuses on the incentives of traders to become informed in the presence of index traders. Thus the following Proposition is helpful.

**Proposition 3** *The profit of informed traders $\pi_I$ is given by*

$$\pi_I = \overline{w}(\overline{Y} - A) + \underline{w}(B - \underline{Y}),$$

*and the profit of index traders $\pi_N$ is given by*

$$\pi_N = \overline{w}\Big((1-\delta)(\overline{Y} - A) + \delta(B - \overline{Y})\Big) + \underline{w}\Big((1-\delta)(\underline{Y} - A) + \delta(B - \underline{Y})\Big)$$

This characterization allows us to arrive at our main result.

**Proposition 4** *Holding $\mu$ constant,*

$$\frac{\partial(\pi_I - \pi_N)}{\partial \nu} < 0.$$

**Corollary 1** *Holding $\mu$ constant, there exists $\nu, \nu'$ where $\nu > \nu'$ and $\underline{\theta} < 0.5$ such that for all $\theta \in (\underline{\theta}, 0.5)$,*

$$\frac{\partial \pi_I}{\partial \theta} > 0,$$

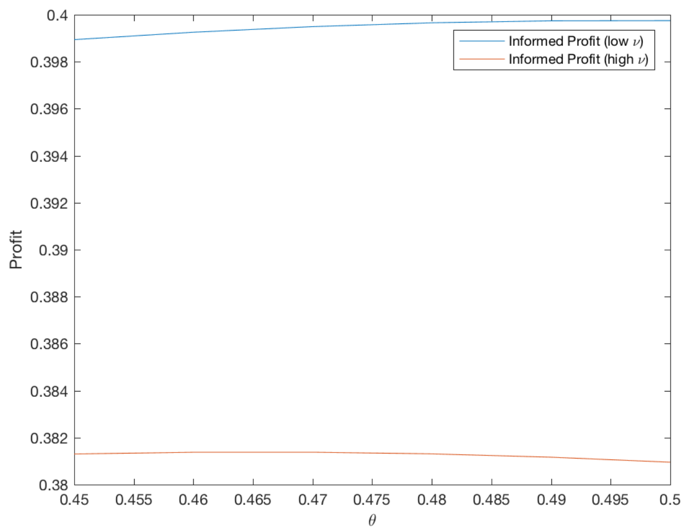
*while*

$$\frac{\partial \pi_I}{\partial \theta} < 0.$$

Proposition 4 states that when the fraction of index investor increases, holding constant the mass of informed traders, returns to being informed falls. When there are more index investors, prices are more likely to reflect the covariance between assets $Y$ and $V$; informed traders cannot hide their trading on having private information about $Y$.

Furthermore, an increase in $\nu$ shifts the point at where more precise signals become less useful. Thus in certain regions, an increase in $\nu$ allows informed profits to eventually be decreasing in the signal precision.

Figure 2.8: Trader profit function of $\nu$




Parameters: $\mu = 0.2$, $\underline{Y} = 1$, $\overline{Y} = 2$, $\delta = \theta = 0.25$

## 2.5 Empirical Test

### Index Data

I restrict the sample to stocks in the Russell 1000 and Russell 2000 index from 1998-2006 [2]. The Russell 2000 is the most commonly tracked small-capitalization market index in the United States, which comprises stocks ranked 2000-3000 by market capitalization. The Russell 1000 contains the top 1000 stocks by market capitalization. Due to the methodology of the Russell index, stocks at the top of the Russell 2000 have 5-10 times the index weights as those at the bottom of the Russell 1000. Thus, this lends itself to a regression discontinuity design where small shocks to market capitalization place stocks arbitrarily on either side of the cutoff.

### Institutional Holdings Data

The 13-F institutional holdings data is from Thompson Reuters. The SEC requires all firms managing in excess of $100 million dollars in notional value to report holdings quarterly.

Following Bushee (1998), I categorize each institution as one of three types: blockholders, quasi-indexers, transient investors. These classifications are based on investment horizon as measured by the frequency of portfolio turnover. This classification is useful as quasi-indexers

[2] The reason for the restriction is that after 2006, Russell made a change to its index methodology that obscured the discontinuity.

should be less information sensitive than transient institutions. Summary statistics for the holdings data are given in Table 2.1.

Table 2.1: Summary Statistics - Institutional Holdings

| | (1) | | (2) | | (3) | |
|---|---|---|---|---|---|---|
| | Blockholders | | Quasi-Indexers | | Transient | |
| | mean | sd | mean | sd | mean | sd |
| Total holdings (in billions) | 74.36 | 140.31 | 12.15 | 43.68 | 6.59 | 17.40 |
| # Russell 1000 stocks | 1.54 | 1.42 | 1.83 | 1.61 | 1.59 | 1.47 |
| # Russell 2000 stocks | 0.69 | 1.35 | 0.59 | 1.39 | 0.54 | 1.04 |
| # Non-Russell 1000/2000 stocks | 0.16 | 0.48 | 0.09 | 0.33 | 0.09 | 0.29 |
| Observations | 337 | | 12659 | | 7611 | |

This table gives summary statistics for the institutional holdings data, by the institution type as defined by Bushee (1998) and membership in the Russell indices.

## Earnings Forecast Data

I calculate the dispersion in earnings data using IBES earnings forecasts for the sample of stocks. Given a quarter $t$, I take the standard deviation of forecasts for that quarter, made prior to $t-1$, and normalize that number by the absolute number of the mean of the forecasts. The reason for the lagged measure is that institutional trading will have already responded to forecasts released in concurrent quarters. Summary statistics for the analyst forecast data are given in Table 2.2.

Table 2.2: Summary Statistics - Analyst Forecasts

| | (1) | | (2) | | (3) | | (4) | |
|---|---|---|---|---|---|---|---|---|
| | 1st Size Quartile | | 2nd Size Quartile | | 3rd Size Quartile | | 4th Size Quartile | |
| | mean | sd | mean | sd | mean | sd | mean | sd |
| Stdev of Forecasts | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.04 | 0.03 | 0.05 |
| Mean of Forecasts | 0.59 | 0.42 | 0.45 | 0.37 | 0.27 | 0.35 | 0.23 | 0.40 |
| # of Forecasts | 8.36 | 5.78 | 4.98 | 5.18 | 4.39 | 3.61 | 3.33 | 2.68 |
| Observations | 125 | | 125 | | 125 | | 125 | |

This table gives summary statistics for IBES analyst forecast data, by firm size quartile.

## Earnings Forecasts and Announcement Drifts in Russell members

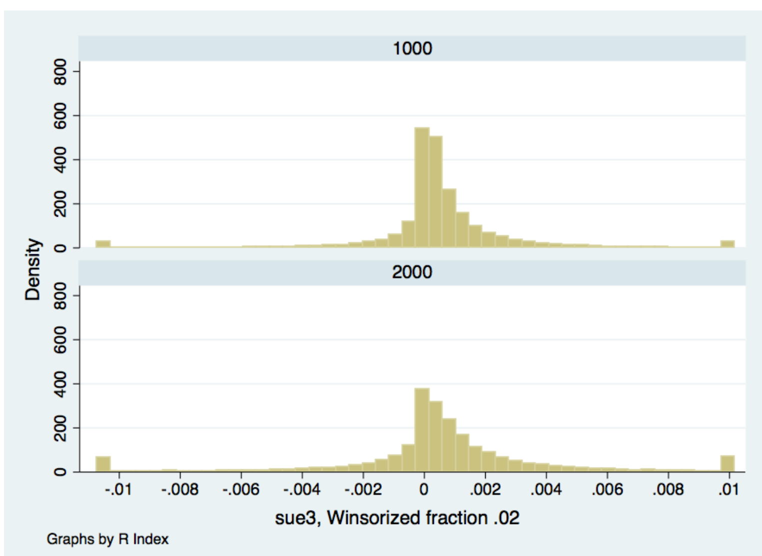
A simple test to examine how price efficiency might be affected by index inclusion would be to analyze the post earnings announcement drift effect (PEAD) around the Russell cutoff (Ball and Brown (1968)). PEAD implies that the returns to firms subject to an earnings surprise drift in the direction of the surprise over the following weeks. Under the hypothesis that price efficiency is improved among the Russell 2000 constituents, then the PEAD effect should be mitigated among these firms.

Using a bandwidth of 300 members around the cutoffs, I created 10 equal sized portfolios on each side of the cutoff. Figure 2.9 shows the standardized unexpected earnings (SUE) measure, defined as actual earnings per share minus the mean forecast, divided by price per share at the announcement date. As the distribution of SUE measures across the cutoff are similar, the portfolios experience comparable earnings surprises. Thus, portfolios across the cutoff can be meaningfully compared.

Figures 2.10 and 2.11 plots the results. For portfolios with large positive SUE measures, PEAD is higher for stocks in the Russell 1000 than for the Russell 2000. For portfolios 9 and 10, the 30 day cumulative excess returns are approximately 5% higher for Russell 1000 stocks. The difference is less for other portfolios, with the exception of Portfolio 3, but otherwise, the drift for Russell 2000 stocks is relatively flat across portfolios. This would be consistent with the idea that information is incorporated sooner into the Russell 2000 via trading.

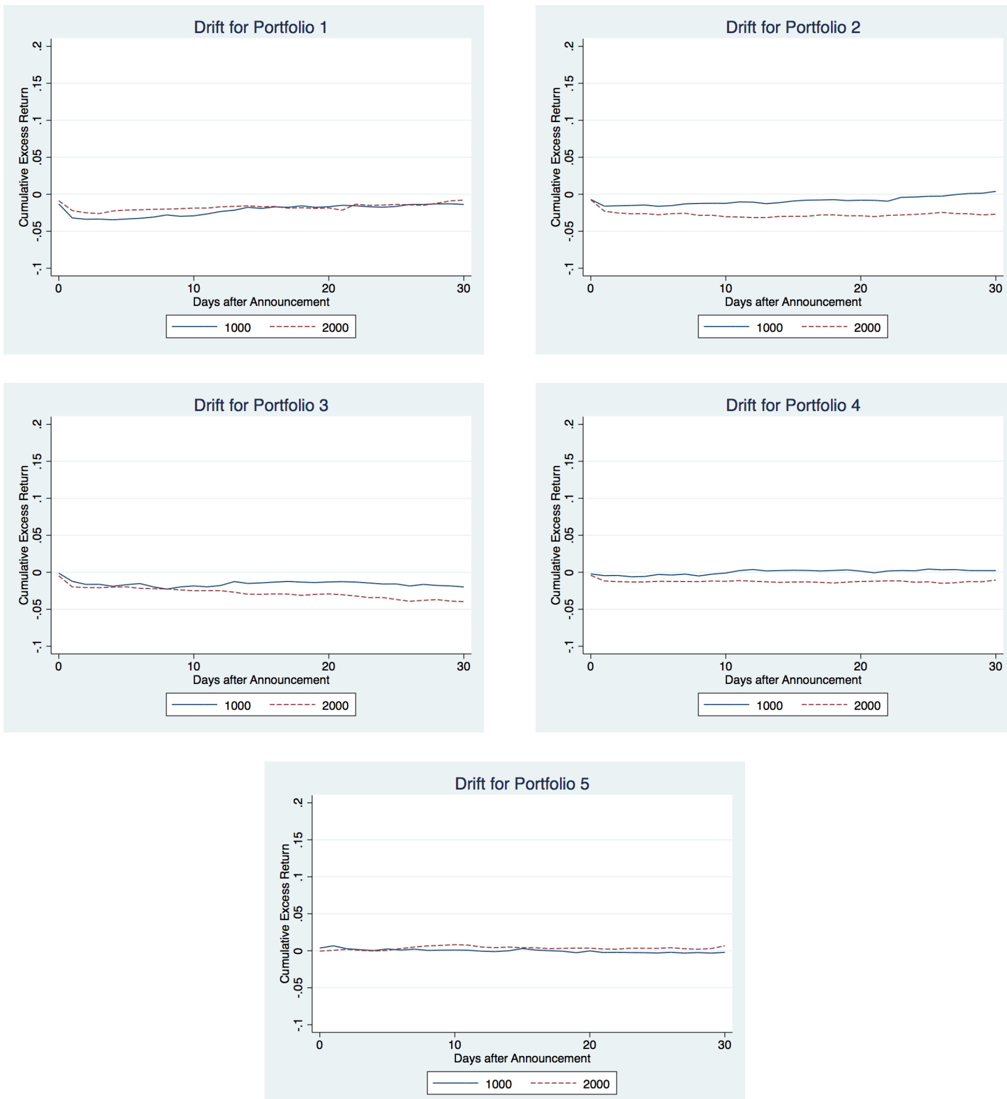
This result cannot be reconciled by either the size effect or noise trading explanations. A size-relate explanation might predict that returns should be higher for the smaller group of Russell 2000 stocks. Similarly, noise trading explanations might predict that Russell 2000 stocks should be slow to incorporate new information.

Figure 2.9: Standardized unexpected earnings (SUE) measure across Russell discontinuity




This figure shows the standardized unexpected earnings measures across the Russell discontinuity. The data is winsorized at the 2% level.

Figure 2.10: Post Earnings Announcement Drift - bottom five portfolios sorted by SUE measure



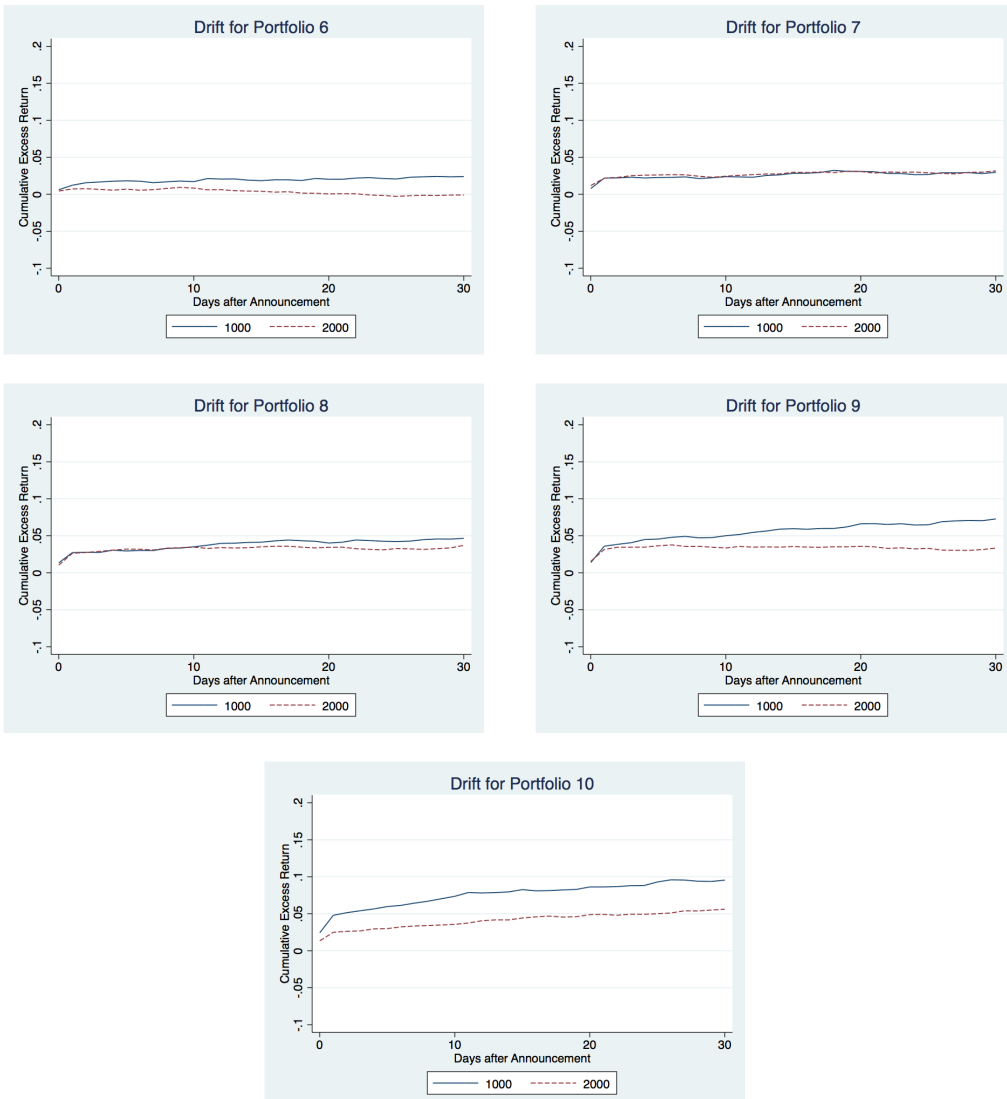

## Does earnings forecast dispersion predict quarterly institutional holdings changes?

For institution $j$ invested in firm $i$ in quarter $t$:

$$|\Delta \log(\text{holdings}_{i,j,t})| = \alpha_j + \delta_t + \beta_1 D_{i,t} + \beta_2 \psi_{i,t} + \gamma \Big[\psi_{i,t} \cdot D_{i,t}\Big] + X'\beta + \epsilon_{i,j,t}, \qquad (2.1)$$

where the dependent variable is the absolute value of the change in log dollar holdings in firm $i$ by institution $j$ in quarter $t$. This variable can be interpreted as trading intensity, as

Figure 2.11: Post Earnings Announcement Drift - top five portfolios sorted by SUE measure




it roughly measures the magnitude of the percentage change in stock holding every quarter. Variable $\alpha_j$ and $\delta_t$ are institution and time fixed-effects, and $D_{i,t}$ is a dummy variable that takes a value of 1 for firm $i$ if it is in the Russell 2000 index in quarter $t$, and 0 otherwise. The variable $\psi_{i,t}$ is any proxy for information uncertainty in firm $i$ prior to quarter $t$. In this case, I use the earnings forecast dispersion described in the previous section. The vector $X$ contains controls; in this case industry fixed-effects and the log market capitalization for firm $i$. Market capitalization is included as a control since Russell membership might be correlated with size. Regressions are estimated separately for the three major types of institutions as classified by Bushee (1998).

If we expect institutions to respond to short run information uncertainty, then we should expect $\beta_2 \neq 0$. Furthermore, $\gamma \neq 0$ if passive index ownership affects the ability of non-indexing institutions to trade on information. However, according to Bushee (1998) classifications, this should only be true for Transient, and not Blockholders or Quasi-Indexers.

Table 1 is consistent with this hypothesis, which shows the results for Equation 2.1 for 100 firms around the Russell cutoff. Column 3 shows that when dispersion increases by 1 point, holdings change by approximately 5%. However, when interacted with the Russell dummy, the holdings change is negated, even when controlling for market capitalization. This is consistent with the idea that while institutional traders respond to uncertainty in forecast earnings, they do not do so when the firms are held in the Russell 2000 index. For Columns 1 and 2 in Table 1, this effect is not present, which is consistent with only Transient investors with frequent portfolio turnovers responding to new information present in earnings forecasts.

Table 2.3: Absolute Value of Change in Log Holdings - 100 firms around cutoff

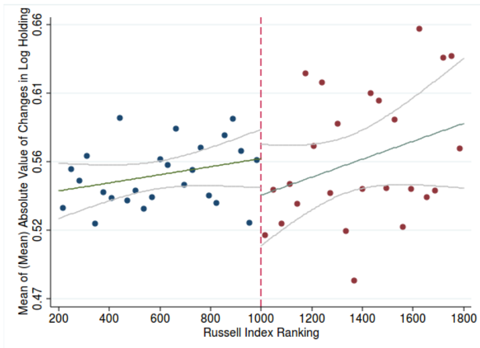
| | (1)<br>Blockholders | (2)<br>Quasi-Indexers | (3)<br>Transient |
|---|---|---|---|
| Russell Dummy=2000 | -0.112$^{**}$ | -0.0414 | -0.0418$^{**}$ |
| | (0.0437) | (0.0264) | (0.0210) |
| Demeaned Forecast Dispersion | -0.0551 | 0.00593 | 0.0502$^{***}$ |
| | (0.0492) | (0.0320) | (0.0183) |
| Russell Dummy=2000 × Demeaned Forecast forersion | 0.0847 | -0.0131 | -0.0596$^{**}$ |
| | (0.106) | (0.0355) | (0.0248) |
| Log Market Cap | -0.222$^{***}$ | -0.419$^{***}$ | -0.278$^{***}$ |
| | (0.0793) | (0.0807) | (0.0639) |
| Constant | 5.212$^{***}$ | 8.743$^{***}$ | 6.572$^{***}$ |
| | (1.627) | (1.656) | (1.325) |
| Quarter-year FE | Yes | Yes | Yes |
| Industry FE | Yes | Yes | Yes |
| Manager FE | Yes | Yes | Yes |
| Observations | 4931 | 157220 | 93306 |
| $R^2$ | 0.139 | 0.155 | 0.104 |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

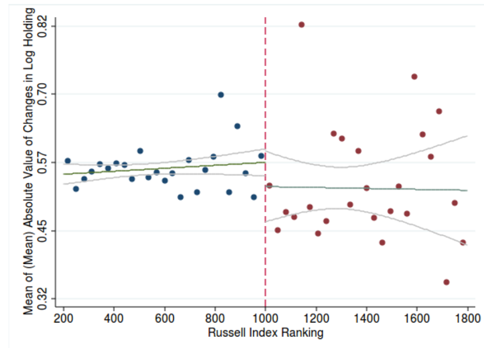
This table shows regression estimates of absolute value of change in log holdings around the Russell cutoff. I estimate three regressions, one for each Bushee (1998) category.

Figure 2.12 shows the regression results for Column 3 in regression discontinuity plots. Firms are sorted into dispersion measure by quartile, and $|\Delta \log(\text{holdings}_{i,j,t})|$ is plotted relative to the cutoff. The discontinuity is larger for firms in the higher dispersion quartiles.

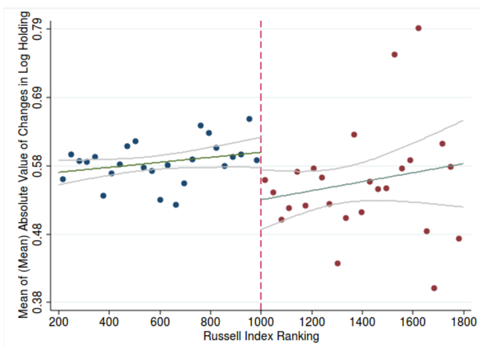
Figure 2.12: Regression Discontinuity Plots - $|\Delta \log(\text{holdings}_{i,j,t})|$ and Dispersion



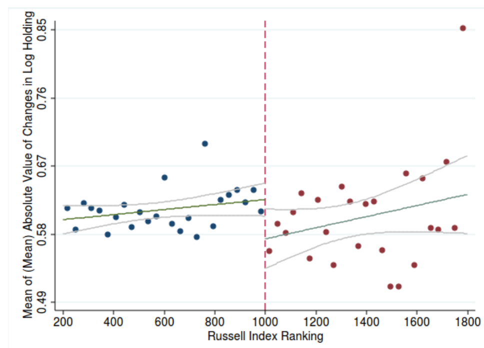

(a) Dispersion (1st Quartile)




(b) Dispersion (2nd Quartile)




(c) Dispersion (3rd Quartile)




(d) Dispersion (4th Quartile)

These figures show regression discontinuity plots across the Russell 2000 index for the four disperion quartiles as sorted by analyst forecast dispersion.

## Is the amount of analyst coverage influenced by the Russell cutoff?

In this section, I examine whether forecast dispersion varies across stocks around the cutoff. I specify the regression

$$N_{i,t} = \alpha_i + \delta_t + \beta D_{i,t} + \epsilon_{i,t}, \tag{2.2}$$

where $N_{i,t}$ is the number of analysts covering firm $i$ at quarter $t$, $\alpha_i$ is a firm fixed effect, $\delta_t$ is a quarter fixed effect, and $D_{i,t}$ is an indicator for Russell 2000 ownership. If attention devoted to producing earnings forecasts plays a role in information acquisition, then we should find a difference between the number of analysts across Russell cutoffs.

Table 2 shows that this is not the case. Controlling for log market capitalization, a regression discontinuity around the Russell for cutoff shows that being included results in,

Table 2.4: Number of Analyst Forecasts

| | (1) |
|---|---|
| | 100 |
| Russell Dummy=2000 | -0.608$^{***}$ |
| | (0.159) |
| Log Market Capitalization | -0.358$^{*}$ |
| | (0.200) |
| Constant | 7.470$^{***}$ |
| | (1.572) |
| Quarter-year FE | Yes |
| Observations | 2855 |
| $R^2$ | 0.052 |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

This table shows the regression estimates for number of analyst forecasts across the Russell 2000 cutoff.

on average, one fewer analyst. The mean number of analysts is seven. This difference is unlikely to produce significant differences in qualities of analyst forecasts.

## 2.6 Conclusion

In this paper, I analyze the information distortions to informed trading in stocks where stock comovements are induced by passive index trading. A simple model explains how profits to being informed can fall when stocks have passive index holdings, due to faster diffusion of information across stocks.

I provide several pieces of empirical evidence consistent with this effect. I use a regression discontinuity approach and analyze institutional holdings changes around the Russell 1000/2000 cutoff, and show that a) announcement drift effects after earnings surprises are mitigated in Russell 2000 stocks just below the cutoff; b) there are fewer changes in institutional holdings for these group of stocks, especially when earnings uncertainty is large; c) there is no material difference in analyst coverage across the cutoff. These observations are consistent with the idea that passive indices improve price efficiency of underlying stocks.

# Chapter 3

# Cryptocurrencies: The rise of a new asset class

## 3.1 Introduction

Cryptocurrencies are an emerging class of digital transactions that use decentralized cryptographic methods to verify ownership records, the most important among them being Bitcoin and Ethereum. As of November 2017, the total market capitalization of cryptocurrencies has reached over $300 billion. Their uses have ranged from crowdfunding projects, speculation, to online transactions in illicit materials. While brokerage, trading, and even financial derivatives for these currencies are beginning to flourish and receive widespread attention, little research is done regarding the asset pricing properties of this new class of assets.

In this paper, we discuss some of these asset pricing properties, namely the return distributions, and show that these returns are large and varied. The stunning growth in market capitalizations and use cases suggest that they shall soon represent an economically impactful class of assets. Next, we show that cryptocurrencies, in aggregate, carry a common source of systematic risk correlated to Bitcoin returns. This has important implications for portfolio diversification and risk assessment. Lastly, we reconcile these stylized facts in a simple model of returns by incorporating a “use value” of exchange.

This paper contributes to an emerging literature on the pricing of Bitcoin, and the larger question of the economic value of cryptocurrencies. A number of papers are concerned with explaining valuation and pricing of Bitcoin from economic principles. Athey et al. (2016) evaluate a model of adoption with Bitcoin prices up to 2015 and concludes that adoption cannot explain prices. Concerns about the speculative nature of Bitcoin are also posed in Yermack (2013). Ciaian et al. (2016) use an econometric approach to show that macrofinancials do not explain Bitcoin prices. Gandal and Halaburda (2014) suggest that a network effect is present that characterizes competition between different cryptocurrencies, and explains Bitcoin’s early dominant position. In this paper, we address some of these explanations, and benchmark cryptocurrency returns empirically with a framework that incorporates network externalities into their asset pricing.

A related strand of literature addresses market efficiency in cryptocurrencies, which is characterized by a high degree of decentralization in trading and in issuances. Kroeger and Sarkar (2017), for example, show persistent violations of the law of one price for Bitcoin, relating this to the microstructure of Bitcoin trading. We add to this literature by documenting the return patterns of alternative cryptocurrencies (altcoins) and initial coin offerings (ICO), documenting a number of remarkable price variations.

The remainder of this article is organized as follows. Section 3.2 discusses the market mechanisms of initial coin offerings and trading. Section 3.3 gives an account of the data source. Section 3.4 discusses the return distribution of cryptocurrencies, initial coin offerings, and correlation with other assets. Section 3.5 discusses the appropriate benchmark to evaluate returns. Section 3.6 concludes.

## 3.2 Market Mechanics of ICOs and Trading

The current landscape contains a wide and varied selection of cryptocurrencies. This paper examines a subset from a data set of over 1,000 cryptocurrencies, which includes Bitcoin, Ethereum, and other altcoins. By comparison, the number of currencies issued by sovereign governments is around 180. The pioneering cryptocurrency is Bitcoin, which is the first to use the technology of *blockchains* to decentralize a distributed ledger of ownership and transactions without the use of a central repository. Typically, supply of coins is predetermined in the long-run – fixed in the case of Bitcoin – or released under a pre-determined schedule. Many altcoins, such as Litecoin, are variations on Bitcoin technology, such as the precise cryptographic challenge employed or in the supply of coins. Others use new software on pre-existing transaction ledgers, called *forks* of existing cryptocurrencies.

An important application of blockchain technology is a distributed computing platform. Platforms such as the Ethereum platform and its currency, the Ether, allow not just distributed ledger verification, but scripts to be run on decentralized computing nodes. This allows more complicated applications to take advantage of the decentralized nature of this computing technology. Many of the digital assets that have ICOs function on the Ethereum blockchain.

There are two categories of tradable currencies alternative to Bitcoin, or altcoins: coins, and tokens. While the difference is often not entirely explicit, the consensus is that coins are used and traded exclusively as mediums of exchange or stores of value, akin to currency, while tokens additionally have some use for a specific application, such as one that requires a reward mechanism, or even representing a share of stake in a project similar to equity. For example, Litecoin has no application-specific use, other than being transacted like a currency, so it is considered a coin. An example of a token is the method of exchange in the decentralized blogging platform Steemit, which rewards content creators with tokens called STEEMs, which are then traded in the secondary market for monetary redemption.

The issuance of tokens are called token sales or initial coin offerings (ICOs), and we liken them to initial public offerings in equity markets. Typical ICOs proceed as follows. Prospective participants first acquire cryptocurrencies, usually Bitcoin or Ethereum's Ether (ETH)[1], to be exchanged for an ICO's tokens at the time of sale. The time, duration, and online address of an offering is publicized beforehand. Some offerings last for a few minutes, and others last for several months. At the time of the ICO, tokens can be purchased by participants with a delivery of the required cryptocurrency. The tokens can be offered for fixed prices by the issuing entities, or the entities can increase the offering prices during the period of the ICO depending on demand. Shortly after the ICO is completed, the tokens can be traded on a secondary market. In a few instances, the secondary market for the tokens begins prior to the ICO.

ICOs are publicized on a number of aggregators[2], to which ICO issuers can submit current or upcoming offerings. These aggregators then direct participants to a white paper which details the purpose of the offering and use of underlying funds, and provides a homepage where the ICO takes place. Homepages then require user registration, and provide a bidding mechanism for any number of tokens that are being offered. Participants typically deliver cryptocurrencies such as Bitcoin or Ether, in exchange for claiming any number of tokens.

Trading of cryptocurrencies occur year-round 24 hours per day on one of many electronic exchanges. The auction mechanism is similar to limit order books in equity markets, but without centralized regulation, market limits, or order size rules. The most prominent of these exchanges, as of November 2017, include Bitfinex and Kraken. Most of the exchange in altcoins occur against Ethereum or Bitcoin, but also occasionally against US dollars.

At the time of this writing, several asset management funds, including hedge funds, have spawned to invest in diversified offerings of cryptocurrencies and ICOs. A few of these are similar to common investment management organizations, which include the use of managers, and charge management, exit, and performance fees. Their focuses may differ, such as in targeting specific return asset pricing characteristics, or specific classes of cryptocurrencies. Organizations such as the CME and CBOE are working to develop cryptocurrency derivatives, which applies expertise in financial engineering to this new class of assets.

## 3.3 Data Sources

The main source of daily historical price data is from the website CoinMarketCap[3]. The website publishes live prices from all of the major cryptocurrency exchanges and records historical daily data. They include daily open, high, low, closing prices, trading volume, and market capitalization. The closing prices are collected at 23:59 UTC daily, including during weekends and US public holidays, and are calculated by calculating the weighted average of

[1] "Ether" is often used synonymously with "Ethereum" to refer to the currency, although the latter more correctly describes the entire platform.

[2] Some of these aggregators are: 99bitcoins.com, icowatchlist.com and icoalert.com

[3] https://www.coinmarketcap.com

prices from all markets. At the time of data collection, no historical data API was provided for the available 1324 currencies, thus we hand collected the historical prices. The set of cryptocurrencies is further separated into "coins" and "tokens".

We keep currencies for which there are at least one years' worth of data and have market capitalizations of at least $1 million by November 23, 2017. Due to the unregulated and decentralized nature of blockchain development, many coins exist that have little legitimacy, and we require at least the market to have allocated capital in coins with sufficient use value, development, and future prospects. Qualitatively, this may mean having a coherent non-technical white paper, a public team of known or anonymous technical specialists, evidence of ongoing updates and monitoring of a currency's development, as well as transparency in the innovation process.

The ICO data is hand collected from individual application white papers as of August 2017. This dataset includes the name of the application, the token abbreviation, the duration and time of the ICO, the total amount of funds raised, token supply, as well as average price during the ICO.

Table 3.1 and 3.2 (Tables 3.3 and 3.4) provide summary statistics for the top 50 coins (tokens), by market capitalization and volume, respectively. For coins, these include the market leader Bitcoin, as well as the well known coins Ethereum, Ripple, Dash, Litecoin, and Monero. Many coins begin their histories a few years prior to the writing of this article. The daily mean return is high compared to other asset classes, with prices moving up to several percentages a day. The total market capitalization of these cryptocurrencies sum to about $250 billion. All of the return figures presented in this paper represent closing price returns over daily or monthly frequencies; none of the returns are annualized.

A limitation of this data source is that it relies on the assumption that closing prices aggregated from multiple exchanges are good summary indications of prices. However as Kroeger and Sarkar (2017) demonstrate, persistent violations of the law of one price imply this may turn out to be a poor assumption. Thus it is necessary that such violations do not affect the correlational studies presented here.

Table 3.1: Top 50 Cryptocurrencies (Coins) by Market Capitalization

| Symbol | Coin Name | Avg Daily Turnover | Mkt Cap ($ million) | Mean Return (%) | Series Start | Series End |
|---|---|---|---|---|---|---|
| BTC | Bitcoin | 20.1814 | 137,444.00 | 0.338 | 28apr2013 | 23nov2017 |
| ETH | Ethereum | 2.4256 | 36,577.30 | 0.959 | 07aug2015 | 23nov2017 |
| BCH | Bitcoin Cash | 53.2955 | 21,899.40 | 1.996 | 23jul2017 | 23nov2017 |
| XRP | Ripple | 0.0008 | 9,230.52 | 0.521 | 04aug2013 | 23nov2017 |
| DASH | Dash | 1.4220 | 4,447.48 | 0.867 | 14feb2014 | 23nov2017 |
| LTC | Litecoin | 1.0301 | 3,882.95 | 0.410 | 28apr2013 | 23nov2017 |
| XMR | Monero | 0.5640 | 2,555.95 | 0.691 | 21may2014 | 23nov2017 |
| MIOTA | IOTA | 0.0078 | 2,494.42 | 0.733 | 13jun2017 | 23nov2017 |
| NEO | NEO | 0.6475 | 2,329.20 | 1.759 | 09sep2016 | 23nov2017 |
| XEM | NEM | 0.0003 | 1,837.21 | 1.090 | 01apr2015 | 23nov2017 |
| ETC | Ethereum Classic | 0.6835 | 1,760.67 | 0.919 | 24jul2016 | 23nov2017 |
| QTUM | Qtum | 1.2361 | 1,047.38 | 1.104 | 24may2017 | 23nov2017 |
| LSK | Lisk | 0.0543 | 997.92 | 1.381 | 06apr2016 | 23nov2017 |
| ZEC | Zcash | 68.0141 | 855.65 | 0.189 | 29oct2016 | 23nov2017 |
| XLM | Stellar Lumens | 0.0004 | 755.91 | 0.589 | 05aug2014 | 23nov2017 |
| ADA | Cardano | 0.0003 | 739.96 | 0.637 | 01oct2017 | 23nov2017 |
| HSR | Hshare | 1.4882 | 702.92 | 1.338 | 20aug2017 | 23nov2017 |
| BCC | BitConnect | 1.2667 | 639.91 | 2.979 | 20jan2017 | 23nov2017 |
| WAVES | Waves | 0.0244 | 529.43 | 0.677 | 02jun2016 | 23nov2017 |
| STRAT | Stratis | 0.0578 | 367.88 | 1.827 | 12aug2016 | 23nov2017 |
| BTS | BitShares | 0.0030 | 355.84 | 0.506 | 21jul2014 | 23nov2017 |
| ARK | Ark | 0.0251 | 304.62 | 2.690 | 22mar2017 | 23nov2017 |
| BCN | Bytecoin | 0.0000 | 283.90 | 0.851 | 17jun2014 | 23nov2017 |
| KMD | Komodo | 0.0147 | 252.32 | 1.928 | 05feb2017 | 23nov2017 |
| DCR | Decred | 0.1493 | 250.98 | 1.008 | 10feb2016 | 23nov2017 |
| STEEM | Steem | 0.0065 | 246.29 | 0.772 | 18apr2016 | 23nov2017 |
| DOGE | Dogecoin | 0.0000 | 203.82 | 0.444 | 15dec2013 | 23nov2017 |
| MONA | MonaCoin | 0.0069 | 199.31 | 0.653 | 20mar2014 | 23nov2017 |
| FCT | Factom | 0.2230 | 198.87 | 1.082 | 06oct2015 | 23nov2017 |
| VTC | Vertcoin | 0.0302 | 170.45 | 0.961 | 20jan2014 | 23nov2017 |
| PIVX | PIVX | 0.0091 | 170.08 | 2.144 | 13feb2016 | 23nov2017 |
| SC | Siacoin | 0.0001 | 168.73 | 1.263 | 26aug2015 | 23nov2017 |
| GBYTE | Byteball Bytes | 1.5395 | 162.86 | 1.674 | 27dec2016 | 23nov2017 |
| BTCD | BitcoinDark | 0.1077 | 144.09 | 0.920 | 16jul2014 | 23nov2017 |
| XZC | ZCoin | 0.4687 | 134.38 | 2.129 | 06oct2016 | 23nov2017 |
| ETP | Metaverse ETP | 0.2083 | 134.34 | 0.737 | 05jun2017 | 23nov2017 |
| GAME | GameCredits | 0.0090 | 127.35 | 2.352 | 01sep2014 | 23nov2017 |
| NXT | Nxt | 0.0017 | 126.35 | 0.588 | 04dec2013 | 23nov2017 |
| SYS | Syscoin | 0.0011 | 122.49 | 0.969 | 20aug2014 | 23nov2017 |
| PURA | Pura | 0.0002 | 111.72 | 5.677 | 27mar2015 | 23nov2017 |
| GXS | GXShares | 0.0449 | 111.69 | 0.390 | 25jun2017 | 23nov2017 |
| BLOCK | Blocknet | 0.0128 | 106.72 | 1.589 | 01nov2014 | 23nov2017 |
| LKK | Lykke | 0.0015 | 94.00 | 0.638 | 14nov2016 | 23nov2017 |
| B3 | B3Coin | 0.0029 | 93.69 | 4.555 | 03sep2016 | 23nov2017 |
| DGB | DigiByte | 0.0003 | 91.24 | 0.715 | 06feb2014 | 23nov2017 |
| XVG | Verge | 0.0000 | 82.27 | 2.336 | 25oct2014 | 23nov2017 |
| VEN | VeChain | 0.0036 | 72.61 | 1.224 | 22aug2017 | 23nov2017 |
| CNX | Cryptonex | 0.0024 | 69.03 | -0.303 | 07oct2017 | 23nov2017 |
| PART | Particl | 0.0375 | 65.91 | 0.525 | 20jul2017 | 23nov2017 |
| ZEN | ZenCash | 0.7330 | 61.00 | 1.437 | 01jun2017 | 23nov2017 |

Data collected from CoinMarketCap as of November 23, 2017. Daily returns are calculated from daily closing times at 11:59 UTC. Only currencies with market capitalization greater than $1 million as of November 23, 2017 were retained. Average turnover and mean return are calculated from all prices available, and market capitalization calculated for November 23, 2017. Returns are winsorized at the 1% level.

Table 3.2: Top 50 Cryptocurrencies (Coins) by Average Daily Turnover

| Symbol | Coin Name | Avg Daily Turnover | Mkt Cap ($ million) | Mean Return (%) | Series Start | Series End |
|---|---|---|---|---|---|---|
| ZEC | Zcash | 68.0141 | 855.65 | 0.189 | 29oct2016 | 23nov2017 |
| BCH | Bitcoin Cash | 53.2955 | 21,899.40 | 1.996 | 23jul2017 | 23nov2017 |
| BTC | Bitcoin | 20.1814 | 137,444.00 | 0.338 | 28apr2013 | 23nov2017 |
| XBC | Bitcoin Plus | 3.3731 | 8.34 | 2.366 | 04may2014 | 23nov2017 |
| ETH | Ethereum | 2.4256 | 36,577.30 | 0.959 | 07aug2015 | 23nov2017 |
| GBYTE | Byteball Bytes | 1.5395 | 162.86 | 1.674 | 27dec2016 | 23nov2017 |
| HSR | Hshare | 1.4882 | 702.92 | 1.338 | 20aug2017 | 23nov2017 |
| DASH | Dash | 1.4220 | 4,447.48 | 0.867 | 14feb2014 | 23nov2017 |
| BCC | BitConnect | 1.2667 | 639.91 | 2.979 | 20jan2017 | 23nov2017 |
| QTUM | Qtum | 1.2361 | 1,047.38 | 1.104 | 24may2017 | 23nov2017 |
| LTC | Litecoin | 1.0301 | 3,882.95 | 0.410 | 28apr2013 | 23nov2017 |
| IOP | Internet of People | 0.9041 | 4.30 | 1.814 | 29nov2016 | 23nov2017 |
| ZEN | ZenCash | 0.7330 | 61.00 | 1.437 | 01jun2017 | 23nov2017 |
| ETC | Ethereum Classic | 0.6835 | 1,760.67 | 0.919 | 24jul2016 | 23nov2017 |
| NEO | NEO | 0.6475 | 2,329.20 | 1.759 | 09sep2016 | 23nov2017 |
| BTX | Bitcore | 0.5945 | 32.32 | 2.429 | 27apr2017 | 23nov2017 |
| XMR | Monero | 0.5640 | 2,555.95 | 0.691 | 21may2014 | 23nov2017 |
| XZC | ZCoin | 0.4687 | 134.38 | 2.129 | 06oct2016 | 23nov2017 |
| OMNI | Omni | 0.3044 | 13.45 | 0.895 | 24dec2013 | 23nov2017 |
| FCT | Factom | 0.2230 | 198.87 | 1.082 | 06oct2015 | 23nov2017 |
| UNO | Unobtanium | 0.2108 | 12.92 | 0.681 | 21dec2013 | 23nov2017 |
| ETP | Metaverse ETP | 0.2083 | 134.34 | 0.737 | 05jun2017 | 23nov2017 |
| INN | Innova | 0.2002 | 2.58 | -0.348 | 08nov2017 | 23nov2017 |
| VIVO | VIVO | 0.1741 | 2.57 | 1.667 | 12sep2017 | 23nov2017 |
| DCR | Decred | 0.1493 | 250.98 | 1.008 | 10feb2016 | 23nov2017 |
| DBIX | DubaiCoin | 0.1338 | 13.88 | 1.704 | 06apr2017 | 23nov2017 |
| CRDNC | Credence Coin | 0.1290 | 13.07 | 9.833 | 06nov2017 | 23nov2017 |
| ZCL | ZClassic | 0.1233 | 3.33 | 0.680 | 10nov2016 | 23nov2017 |
| EBST | eBoost | 0.1098 | 1.29 | 0.864 | 06jun2017 | 23nov2017 |
| SBD | Steem Dollars | 0.1096 | 5.56 | 0.260 | 18jul2016 | 23nov2017 |
| BTCD | BitcoinDark | 0.1077 | 144.09 | 0.920 | 16jul2014 | 23nov2017 |
| CLAM | Clams | 0.1001 | 17.03 | 1.018 | 26aug2014 | 23nov2017 |
| NVC | Novacoin | 0.0915 | 10.96 | 0.287 | 28apr2013 | 23nov2017 |
| EXP | Expanse | 0.0773 | 14.48 | 1.036 | 15sep2015 | 23nov2017 |
| STRAT | Stratis | 0.0578 | 367.88 | 1.827 | 12aug2016 | 23nov2017 |
| LSK | Lisk | 0.0543 | 997.92 | 1.381 | 06apr2016 | 23nov2017 |
| VRM | VeriumReserve | 0.0538 | 3.22 | 1.905 | 17sep2016 | 23nov2017 |
| RADS | Radium | 0.0507 | 12.25 | 1.545 | 18jan2016 | 23nov2017 |
| DFT | DraftCoin | 0.0477 | 1.85 | 1.064 | 23nov2015 | 23nov2017 |
| GXS | GXShares | 0.0449 | 111.69 | 0.390 | 25jun2017 | 23nov2017 |
| XCP | Counterparty | 0.0437 | 36.93 | 0.550 | 15feb2014 | 23nov2017 |
| SLS | SaluS | 0.0427 | 13.73 | 3.274 | 22jan2016 | 23nov2017 |
| PASC | Pascal Coin | 0.0403 | 9.97 | 2.046 | 07dec2016 | 23nov2017 |
| PPY | Peerplays | 0.0400 | 16.83 | 0.969 | 13jun2017 | 23nov2017 |
| HUSH | Hush | 0.0391 | 2.91 | 3.720 | 01dec2016 | 23nov2017 |
| PART | Particl | 0.0375 | 65.91 | 0.525 | 20jul2017 | 23nov2017 |
| KORE | Kore | 0.0373 | 7.31 | 2.489 | 21jun2014 | 23nov2017 |
| PKB | ParkByte | 0.0360 | 5.37 | 4.056 | 19may2015 | 23nov2017 |
| XAS | Asch | 0.0358 | 53.09 | 3.036 | 01apr2017 | 23nov2017 |
| DYN | Dynamic | 0.0339 | 6.44 | 1.516 | 26mar2017 | 23nov2017 |

Data collected from CoinMarketCap as of November 23, 2017. Daily returns are calculated from daily closing times at 11:59 UTC. Only currencies with market capitalization greater than $1 million as of November 23, 2017 were retained. Average turnover and mean return are calculated from all prices available, and market capitalization calculated for November 23, 2017. Returns are winsorized at the 1% level.

Table 3.3: Top 50 Cryptocurrencies (Tokens) by Market Capitalization

| **Symbol** | **Coin Name** | **Avg Daily Turnover** | **Mkt Cap ($ million)** | **Mean Return (%)** | **Series Start** | **Series End** |
|---|---|---|---|---|---|---|
| EOS | EOS | 0.1257 | 929.84 | 1.247 | 01jul2017 | 23nov2017 |
| OMG | OmiseGO | 0.5141 | 814.28 | 2.938 | 14jul2017 | 23nov2017 |
| USDT | Tether | 0.2689 | 674.06 | 0.001 | 25feb2015 | 23nov2017 |
| PPT | Populous | 0.0160 | 417.12 | 2.412 | 11jul2017 | 23nov2017 |
| POWR | Power Ledger | 0.1646 | 297.01 | 15.706 | 01nov2017 | 23nov2017 |
| ARDR | Ardor | 0.0014 | 295.37 | 0.873 | 23jul2016 | 23nov2017 |
| REP | Augur | 0.1729 | 276.16 | 0.850 | 27oct2015 | 23nov2017 |
| PAY | TenX | 0.0956 | 198.18 | 1.204 | 27jun2017 | 23nov2017 |
| RDN | Raiden Network Token | 0.1213 | 196.15 | 10.845 | 08nov2017 | 23nov2017 |
| MAID | MaidSafeCoin | 0.0014 | 192.91 | 0.495 | 28apr2014 | 23nov2017 |
| GNT | Golem | 0.0050 | 192.90 | 1.307 | 18nov2016 | 23nov2017 |
| VERI | Veritaseum | 0.3573 | 187.70 | 2.515 | 08jun2017 | 23nov2017 |
| GAS | Gas | 0.4201 | 182.28 | 3.225 | 06jul2017 | 23nov2017 |
| SALT | SALT | 0.1351 | 179.84 | 0.368 | 29sep2017 | 23nov2017 |
| AE | Aeternity | 0.0029 | 169.02 | 2.375 | 01jun2017 | 23nov2017 |
| BATa | Basic Attention Token | 0.0038 | 166.95 | 0.482 | 01jun2017 | 23nov2017 |
| BNB | Binance Coin | 0.1246 | 159.47 | 3.352 | 25jul2017 | 23nov2017 |
| DGD | DigixDAO | 0.1378 | 152.49 | 0.581 | 18apr2016 | 23nov2017 |
| TRX | TRON | 0.0000 | 151.66 | 1.450 | 13sep2017 | 23nov2017 |
| KNCa | Kyber Network | 0.0274 | 149.48 | -0.607 | 24sep2017 | 23nov2017 |
| ICNa | Iconomi | 0.0107 | 147.20 | 1.179 | 30sep2016 | 23nov2017 |
| ETHOS | Ethos | 0.0149 | 128.14 | 4.123 | 18jul2017 | 23nov2017 |
| SNT | Status | 0.0029 | 123.13 | 0.365 | 28jun2017 | 23nov2017 |
| BTMa | Bytom | 0.0088 | 116.09 | 1.023 | 08aug2017 | 23nov2017 |
| WTC | Walton | 0.4696 | 116.03 | 3.519 | 27aug2017 | 23nov2017 |
| ZRX | 0x | 0.0078 | 114.44 | 0.501 | 16aug2017 | 23nov2017 |
| QSP | Quantstamp | 0.0837 | 110.64 | 41.937 | 21nov2017 | 23nov2017 |
| CVC | Civic | 0.0160 | 104.83 | 1.072 | 17jul2017 | 23nov2017 |
| FUN | FunFair | 0.0003 | 94.59 | 1.505 | 27jun2017 | 23nov2017 |
| BNT | Bancor | 0.0717 | 91.58 | -0.445 | 18jun2017 | 23nov2017 |
| MTL | Metal | 0.1781 | 85.02 | 1.488 | 09jul2017 | 23nov2017 |
| ATM | ATMChain | 0.0004 | 83.47 | 1.561 | 04oct2017 | 23nov2017 |
| GNO | Gnosis | 1.8703 | 79.74 | 0.415 | 01may2017 | 23nov2017 |
| SNGLS | SingularDTV | 0.0007 | 78.43 | 1.083 | 03oct2016 | 23nov2017 |
| STORJ | Storj | 0.0454 | 77.56 | 1.271 | 02jul2017 | 23nov2017 |
| MGO | MobileGo | 0.0052 | 63.36 | -0.194 | 11jun2017 | 23nov2017 |
| ADX | AdEx | 0.1581 | 62.02 | 2.566 | 01jul2017 | 23nov2017 |
| EDG | Edgeless | 0.0155 | 60.12 | 2.246 | 30mar2017 | 23nov2017 |
| LINK | ChainLink | 0.0132 | 59.38 | 0.473 | 20sep2017 | 23nov2017 |
| ANT | Aragon | 0.0227 | 58.75 | 0.539 | 18may2017 | 23nov2017 |
| PPP | PayPie | 0.0021 | 58.56 | 4.133 | 10oct2017 | 23nov2017 |
| RCNa | Ripio Credit Network | 0.0141 | 55.74 | 2.751 | 26oct2017 | 23nov2017 |
| MCO | Monaco | 0.7551 | 52.86 | 2.202 | 03jul2017 | 23nov2017 |
| LRC | Loopring | 0.0032 | 50.36 | 1.928 | 30aug2017 | 23nov2017 |
| DATA | Streamr DATAcoin | 0.0019 | 50.22 | 1.095 | 03nov2017 | 23nov2017 |
| QRL | Quantum Resistant Ledger | 0.0144 | 49.92 | 0.808 | 10jun2017 | 23nov2017 |
| ZSC | Zeusshield | 0.0006 | 48.74 | 0.694 | 13oct2017 | 23nov2017 |
| KIN | Kin | 0.0000 | 48.24 | -0.575 | 27sep2017 | 23nov2017 |
| RLC | iExec RLC | 0.0059 | 46.14 | 0.886 | 20apr2017 | 23nov2017 |
| WINGS | Wings | 0.0072 | 45.51 | 2.032 | 11jan2017 | 23nov2017 |

Data collected from CoinMarketCap as of November 23, 2017. Daily returns are calculated from daily closing times at 11:59 UTC. Only currencies with market capitalization greater than $1 million as of November 23, 2017 were retained. Average turnover and mean return are calculated from all prices available, and market capitalization calculated for November 23, 2017. Returns are winsorized at the 1% level.

Table 3.4: Top 50 Cryptocurrencies (Tokens) by Average Daily Turnover

| **Symbol** | **Coin Name** | **Avg Daily Turnover** | **Mkt Cap ($ million)** | **Mean Return (%)** | **Series Start** | **Series End** |
|---|---|---|---|---|---|---|
| PBT | Primalbase Token | 26.7742 | 3.72 | 0.646 | 27jul2017 | 23nov2017 |
| OTN | Open Trading Network | 6.6446 | 40.79 | 13.808 | 25oct2017 | 23nov2017 |
| ETBS | Ethbits | 2.9753 | 1.09 | 5.711 | 01jun2017 | 23nov2017 |
| GNO | Gnosis | 1.8703 | 79.74 | 0.415 | 01may2017 | 23nov2017 |
| NMR | Numeraire | 0.9766 | 15.80 | 0.099 | 23jun2017 | 23nov2017 |
| MCO | Monaco | 0.7551 | 52.86 | 2.202 | 03jul2017 | 23nov2017 |
| TIME | Chronobank | 0.7430 | 15.65 | 0.615 | 27feb2017 | 23nov2017 |
| MLN | Melon | 0.6312 | 39.40 | 0.877 | 22feb2017 | 23nov2017 |
| KLN | Kolion | 0.6169 | 2.28 | 1.757 | 29sep2017 | 23nov2017 |
| OMG | OmiseGO | 0.5141 | 814.28 | 2.938 | 14jul2017 | 23nov2017 |
| XAUR | Xaurum | 0.5016 | 22.27 | 0.837 | 21apr2015 | 23nov2017 |
| WTC | Walton | 0.4696 | 116.03 | 3.519 | 27aug2017 | 23nov2017 |
| LUN | Lunyr | 0.4294 | 11.48 | 1.018 | 01may2017 | 23nov2017 |
| GAS | Gas | 0.4201 | 182.28 | 3.225 | 06jul2017 | 23nov2017 |
| VERI | Veritaseum | 0.3573 | 187.70 | 2.515 | 08jun2017 | 23nov2017 |
| GVT | Genesis Vision | 0.3546 | 10.23 | 0.368 | 15nov2017 | 23nov2017 |
| LGD | Legends Room | 0.3305 | 1.52 | 0.090 | 16jun2017 | 23nov2017 |
| USDT | Tether | 0.2689 | 674.06 | 0.001 | 25feb2015 | 23nov2017 |
| AVT | Aventus | 0.1995 | 16.45 | 0.365 | 06sep2017 | 23nov2017 |
| MTL | Metal | 0.1781 | 85.02 | 1.488 | 09jul2017 | 23nov2017 |
| REP | Augur | 0.1729 | 276.16 | 0.850 | 27oct2015 | 23nov2017 |
| POWR | Power Ledger | 0.1646 | 297.01 | 15.706 | 01nov2017 | 23nov2017 |
| APX | APX | 0.1583 | 5.01 | 1.237 | 25may2017 | 23nov2017 |
| ADX | AdEx | 0.1581 | 62.02 | 2.566 | 01jul2017 | 23nov2017 |
| MCAP | MCAP | 0.1428 | 13.84 | -0.639 | 30may2017 | 23nov2017 |
| DGD | DigixDAO | 0.1378 | 152.49 | 0.581 | 18apr2016 | 23nov2017 |
| ICOS | ICOS | 0.1352 | 30.81 | 1.331 | 23oct2017 | 23nov2017 |
| SALT | SALT | 0.1351 | 179.84 | 0.368 | 29sep2017 | 23nov2017 |
| PLU | Pluton | 0.1297 | 6.02 | 2.846 | 17sep2016 | 23nov2017 |
| EOS | EOS | 0.1257 | 929.84 | 1.247 | 01jul2017 | 23nov2017 |
| BNB | Binance Coin | 0.1246 | 159.47 | 3.352 | 25jul2017 | 23nov2017 |
| RDN | Raiden Network Token | 0.1213 | 196.15 | 10.845 | 08nov2017 | 23nov2017 |
| EDO | Eidoo | 0.1134 | 24.48 | 0.240 | 17oct2017 | 23nov2017 |
| EVX | Everex | 0.0994 | 25.44 | 0.606 | 10oct2017 | 23nov2017 |
| PAY | TenX | 0.0956 | 198.18 | 1.204 | 27jun2017 | 23nov2017 |
| MOD | Modum | 0.0907 | 31.08 | 5.186 | 23oct2017 | 23nov2017 |
| QSP | Quantstamp | 0.0837 | 110.64 | 41.937 | 21nov2017 | 23nov2017 |
| BNT | Bancor | 0.0717 | 91.58 | -0.445 | 18jun2017 | 23nov2017 |
| INXT | Internxt | 0.0625 | 1.15 | 0.863 | 07oct2017 | 23nov2017 |
| TKS | Tokes | 0.0621 | 2.03 | 1.206 | 26mar2017 | 23nov2017 |
| BITUSD | bitUSD | 0.0616 | 4.34 | 0.158 | 23sep2014 | 23nov2017 |
| ARN | Aeron | 0.0596 | 6.00 | 3.050 | 07nov2017 | 23nov2017 |
| MDA | Moeda Loyalty Points | 0.0585 | 39.11 | 0.447 | 11sep2017 | 23nov2017 |
| 1ST | FirstBlood | 0.0471 | 32.54 | 0.760 | 28sep2016 | 23nov2017 |
| STORJ | Storj | 0.0454 | 77.56 | 1.271 | 02jul2017 | 23nov2017 |
| ELIX | Elixir | 0.0442 | 1.41 | 3.064 | 23sep2017 | 23nov2017 |
| HDG | Hedge | 0.0350 | 3.54 | 4.050 | 22oct2017 | 23nov2017 |
| PRIX | Privatix | 0.0350 | 5.20 | 3.747 | 16nov2017 | 23nov2017 |
| TRIG | Triggers | 0.0333 | 23.06 | 2.272 | 11oct2016 | 23nov2017 |
| STX | Stox | 0.0332 | 22.67 | -0.048 | 05aug2017 | 23nov2017 |

Data collected from CoinMarketCap as of November 23, 2017. Monthly returns are calculated from daily closing times at 11:59 UTC. Only currencies with market capitalization greater than $1 million as of November 23, 2017 were retained. Average turnover and mean return are calculated from all prices available, and market capitalization calculated for November 23, 2017. Returns are winsorized at the 1% level.

## 3.4 Return Characteristics

### Summary of Return Properties

Summary statistics for cryptocurrency returns are given in Table 3.5. The mean of the population of mean daily returns for cryptocurrencies is 1.93%. Applying filters referenced in the previous section result in 222 cryptocurrencies. Returns are winsorized at the 1% level so that large outliers leave the distributions interpretable. Figure 3.1 shows that mean daily returns roughly follow a power law distribution, with a significant amount of daily returns as high as 5-10% per day, but with most falling around 1-2% per day. Log market capitalizations show an even greater degree of skewness.

Bitcoin returns, as a representative cryptocurrency, displays basic time-series econometric properties that are similar to stock prices. Table 3.6 shows that, at both the daily and monthly frequencies, we can confidently reject a unit root for returns, but not for prices. Table 3.7 shows this approximately in regression form. With no constant, one period lagged returns do not predict returns next period returns for either daily or monthly returns.

### Market portfolio returns

We consider a portfolio of diversified cryptocurrencies and their performance over time. Figure 3.5 compares the value of $100 invested in an equal-weighted and market capitalization-weighted portfolio of cryptocurrencies. To be included in the portfolio at November 23, 2014, we consider only cryptocurrencies with daily value traded greater than $100. This leaves 37 cryptocurrencies. As shown, the equal-weighted portfolio outperforms the market-weighted portfolio, due to the currency exposures aside from Bitcoin. A three-year investment held in such portfolios would lead to a 31-fold and 22-fold increase in the initial investment value, respectively. Nonetheless, there is a clear degree of co-variation among the two portfolios.

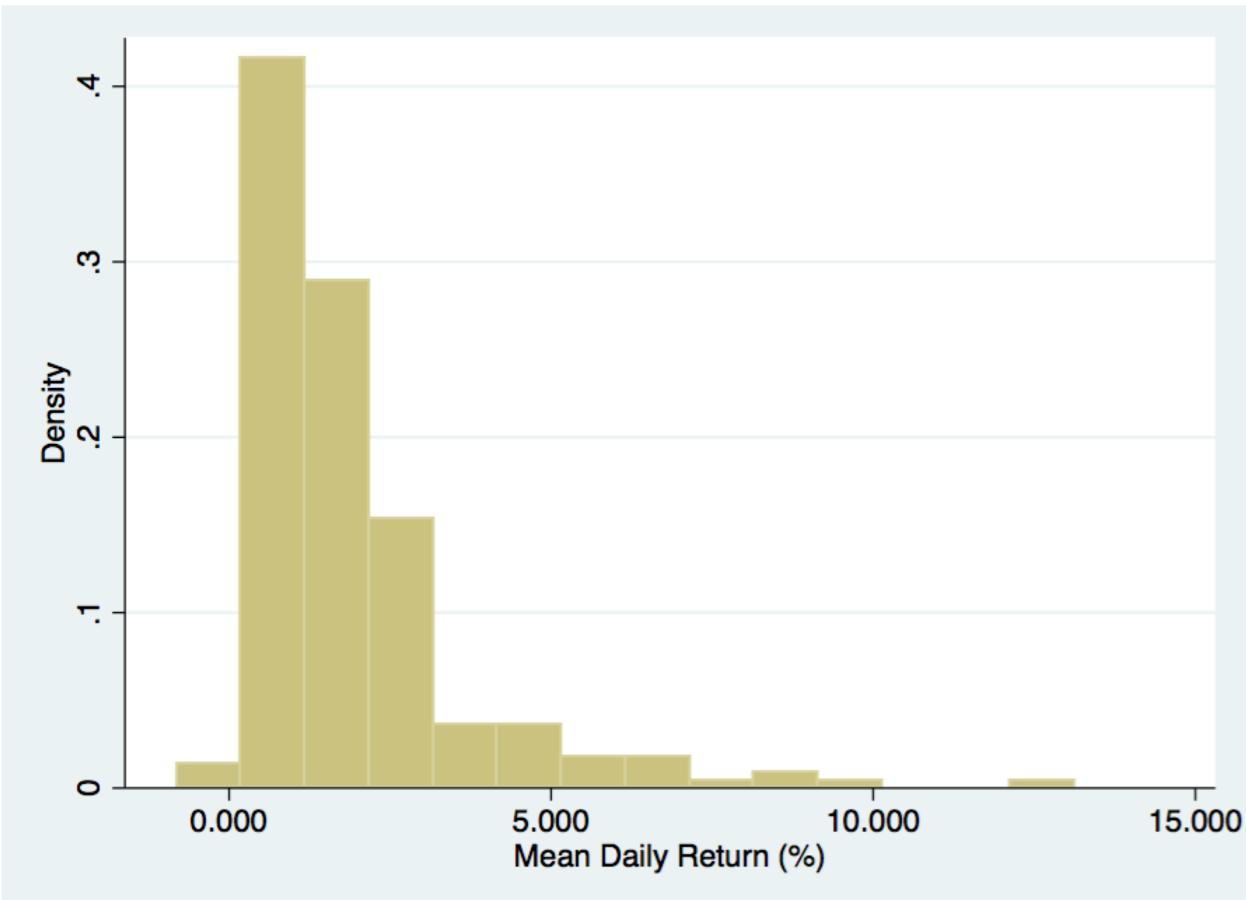
To examine the risk-return tradeoff, we plot the efficient frontier of daily returns for a two-year horizon, formed by the top 50 cryptocurrencies by daily turnover as of November 23, 2015. The global minimum variance portfolio has a daily volatility of 3.22% and daily mean return of 0.68%. This implies an annualized Sharpe ratio[4] of 4.02. This minimum

[4] Annualized Sharpe ratio is calculated by multiplying the Sharpe ratio from daily data by $\sqrt{365}$

Table 3.5: Summary Statistics of Cryptocurrencies

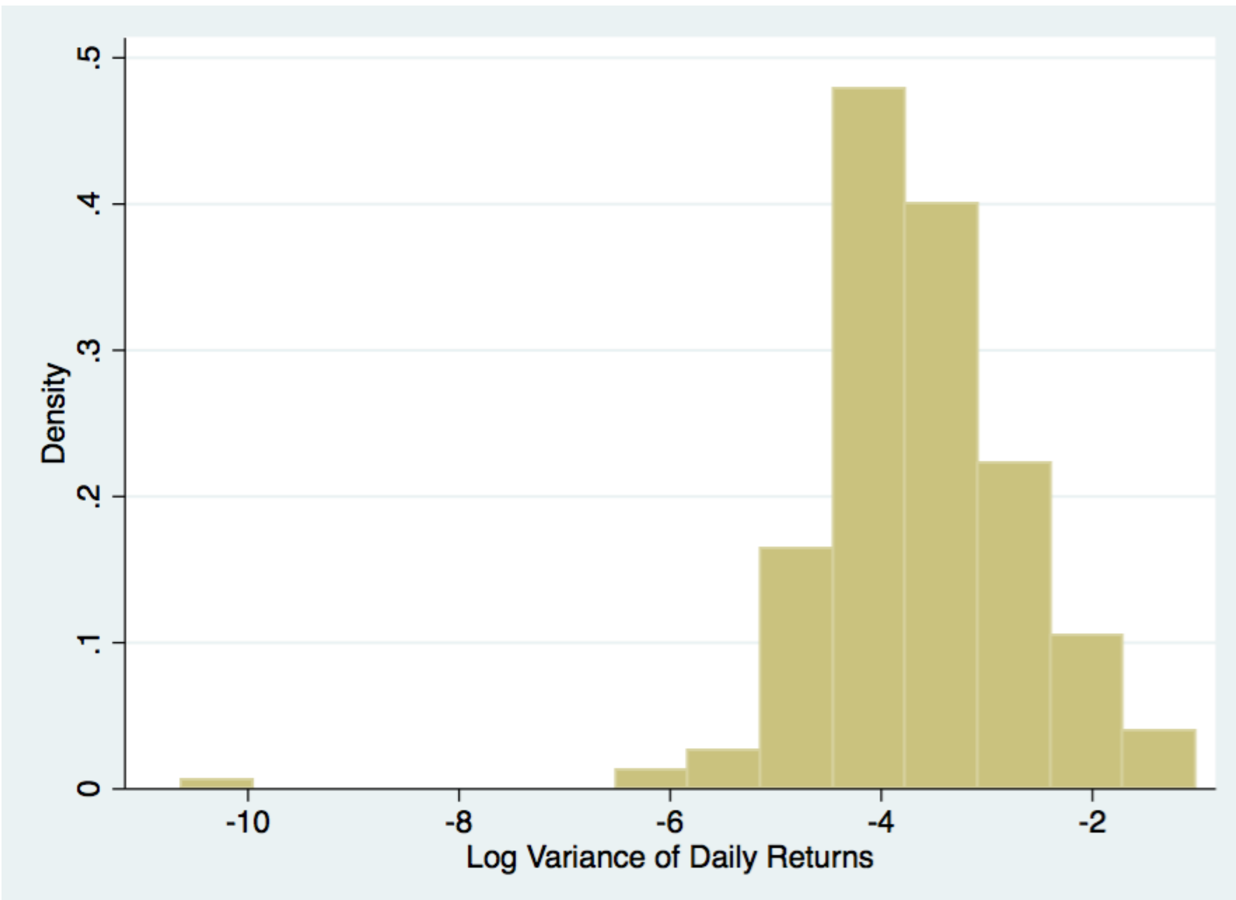
| | (1) | | | | | | |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | p25 | p50 | p75 | Max |
| Mean Daily Returns | .0193 | .018 | -.00809 | .00869 | .0133 | .0234 | .131 |
| Variance of Daily Returns | .041 | .0469 | .0000241 | .015 | .0254 | .0464 | .36 |
| Daily Turnover | .463 | 4.76 | 3.59e-09 | .000195 | .00219 | .0115 | 68 |
| Market Capitalization ($ 1000s) | 950,986 | 9,553,533 | 1,026 | 3,328 | 8,701 | 36,777 | 137,444,000 |
| Observations | 222 | | | | | | |

Figure 3.1: Histogram of mean daily returns



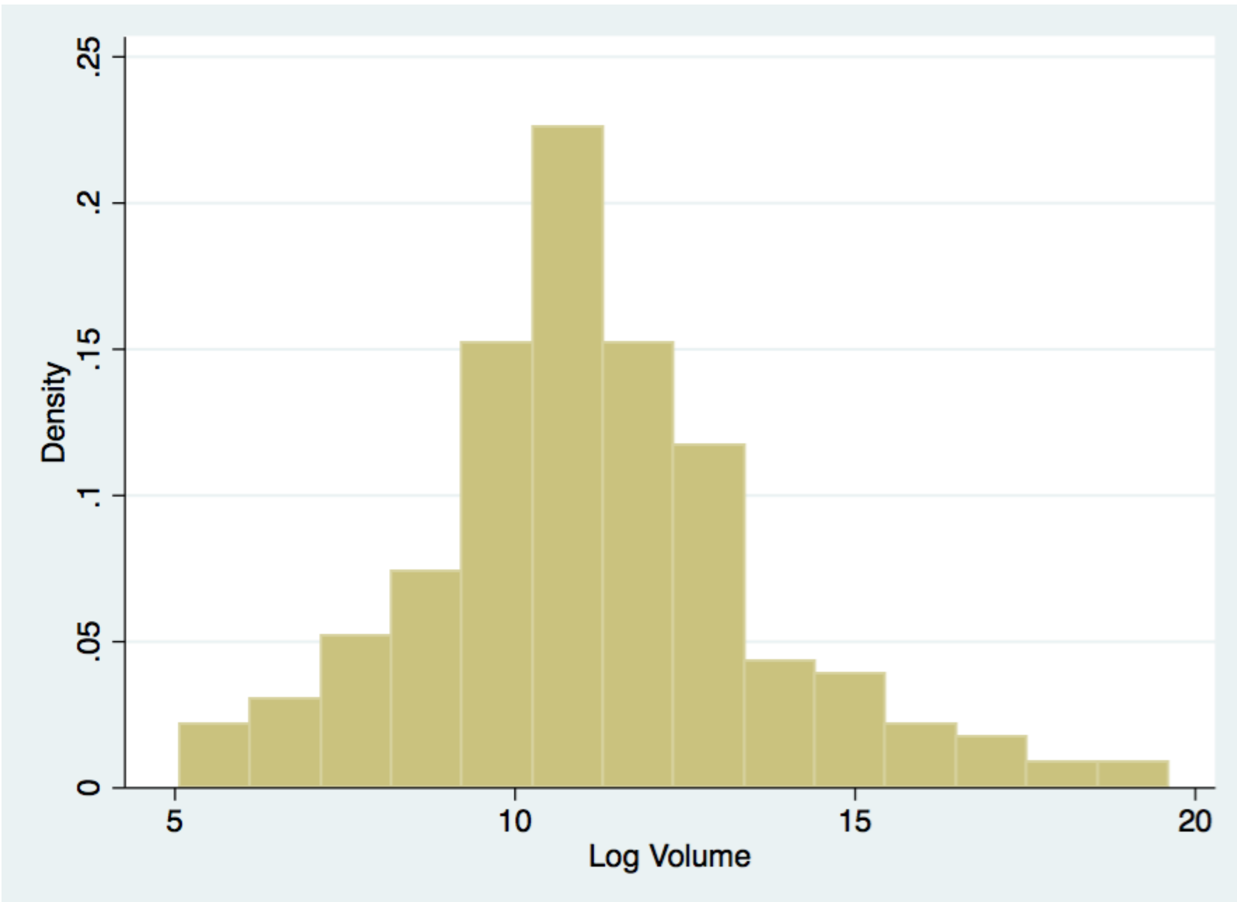

Data collected from CoinMarketCap as of November 23, 2017. Plotted is the histogram for mean daily returns. Daily returns are calculated from daily closing times at 11:59 UTC. Returns used to calculate means are winsorized at the 1% level.

Figure 3.2: Histogram of log variance of daily returns



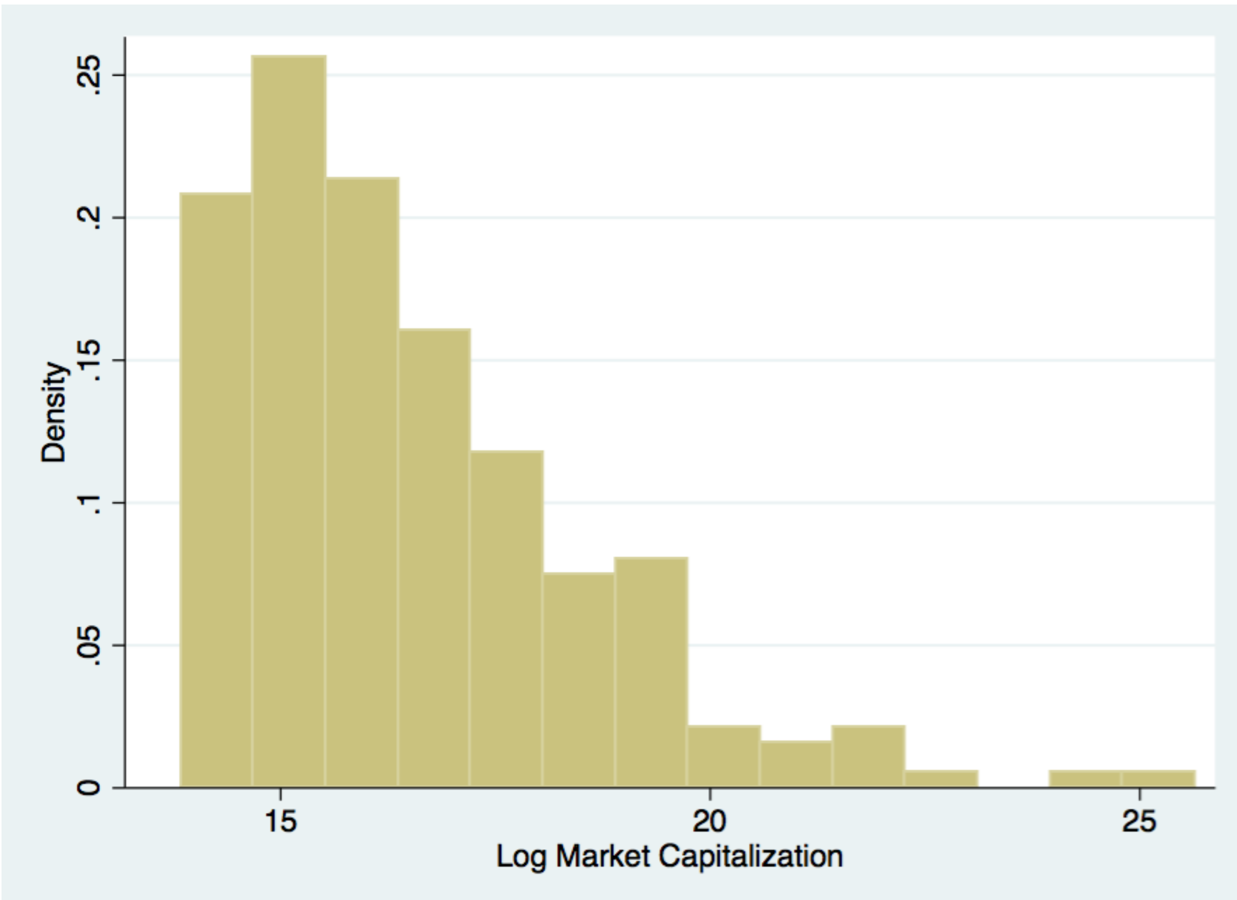

Data collected from CoinMarketCap as of November 23, 2017. Plotted is the histogram for log variances of daily returns. Daily returns are calculated from daily closing times at 11:59 UTC. Returns used to calculate means are winsorized at the 1% level.

Figure 3.3: Histogram of log volume



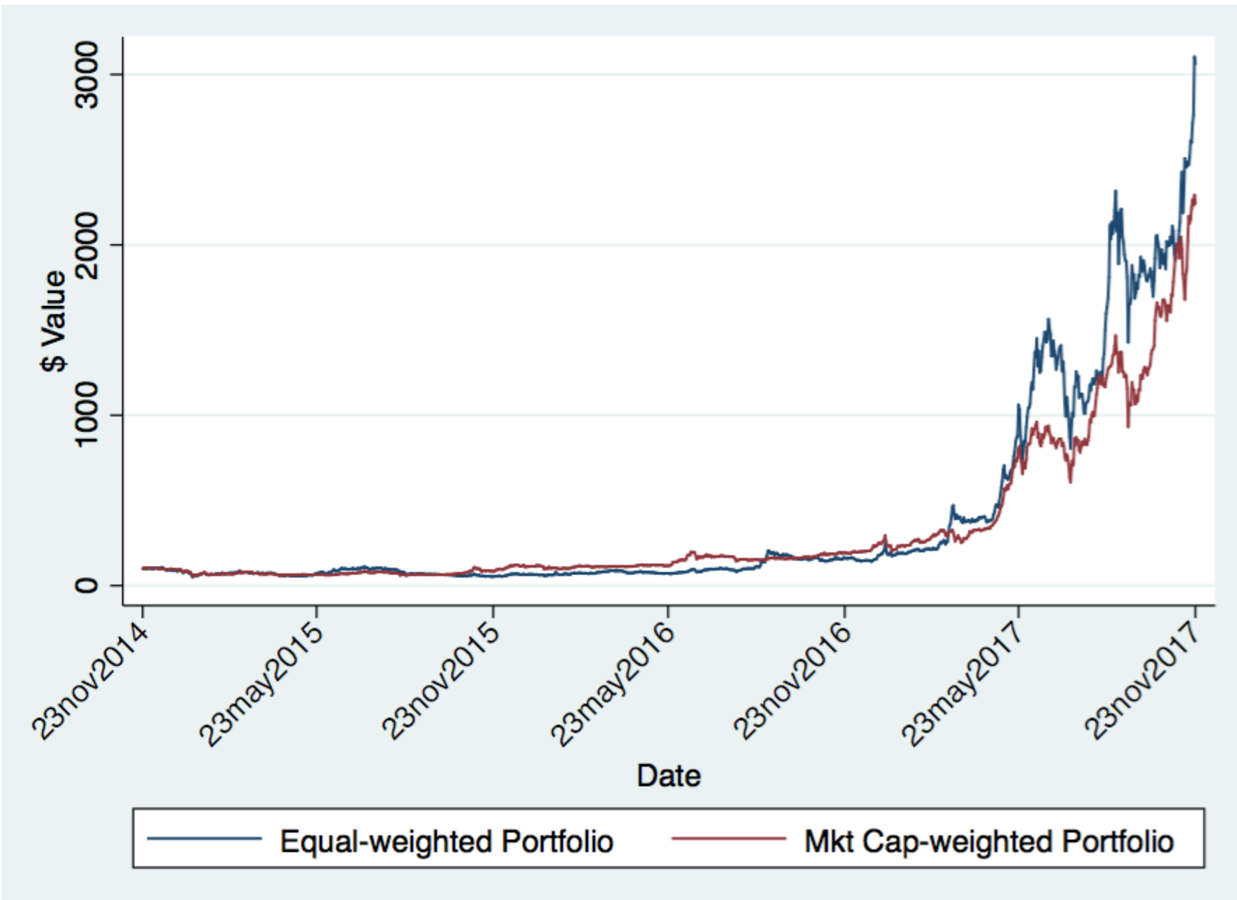

Data collected from CoinMarketCap as of November 23, 2017. Plotted is the histogram for log daily average volume. Daily returns are calculated from daily closing times at 11:59 UTC. Returns used to calculate means are winsorized at the 1% level.

Figure 3.4: Histogram of log market capitalization




Data collected from CoinMarketCap as of November 23, 2017. Plotted is the histogram for log market capitalization, calculated at November 23, 2017. Daily returns are calculated from daily closing times at 11:59 UTC. Returns used to calculate means are winsorized at the 1% level.

Figure 3.5: Value of $100 Invested in Diversified Cryptocurrency Portfolio




This plot considers the change in the value of $100 if invested into an equal-weighted portfolio and a market-capitalization weighted portfolio on November 23, 2014. These portfolios consist of all tradable cryptocurrencies in the data. To be considered tradable, we require the daily value traded to be greater than $100. The total size of the portfolio is 37 currencies. Approximately 89% of the market-capitalization weighted portfolio consists of Bitcoin.

Table 3.6: Dickey-Fuller Tests of Unit Root for Bitcoin Returns and Prices

| | | Critical Value | | | | |
|---|---|---|---|---|---|---|
| **Frequency/Price** | **Dickey-Fuller Test Statistic** | **1%** | **5%** | **10%** | **N** | **p-value** |
| Daily Return | -41.09 | -3.96 | -3.41 | -3.12 | 1,669 | 0.000 |
| Monthly Return | -7.10 | -4.14 | -3.49 | -3.18 | 55 | 0.000 |
| Daily Price | 4.35 | -3.96 | -3.41 | -3.12 | 1,669 | 1.000 |
| Monthly Price | 4.28 | -4.14 | -3.49 | -3.18 | 55 | 1.000 |

This table calculates the Dickey-Fuller test statistics to test for unit roots in the Bitcoin return series. Test statistics were calculated separately for daily and monthly returns and prices. These series include a trend.

Table 3.7: One Period Lagged Bitcoin Return Predictability Regressions

| **Frequency** | **Coef.** | **Std. Err** | **t** | **N** | $R^2$ |
|---|---|---|---|---|---|
| Daily | 0.001 | 0.053 | 0.02 | 1,669 | 0.00 |
| Monthly | -0.137 | 0.132 | -1.04 | 55 | 0.02 |

This table presents coefficients for regressions of Bitcoin returns on one period lag of itself. No constant is included in the regressions. Results are provided for daily and monthly series.

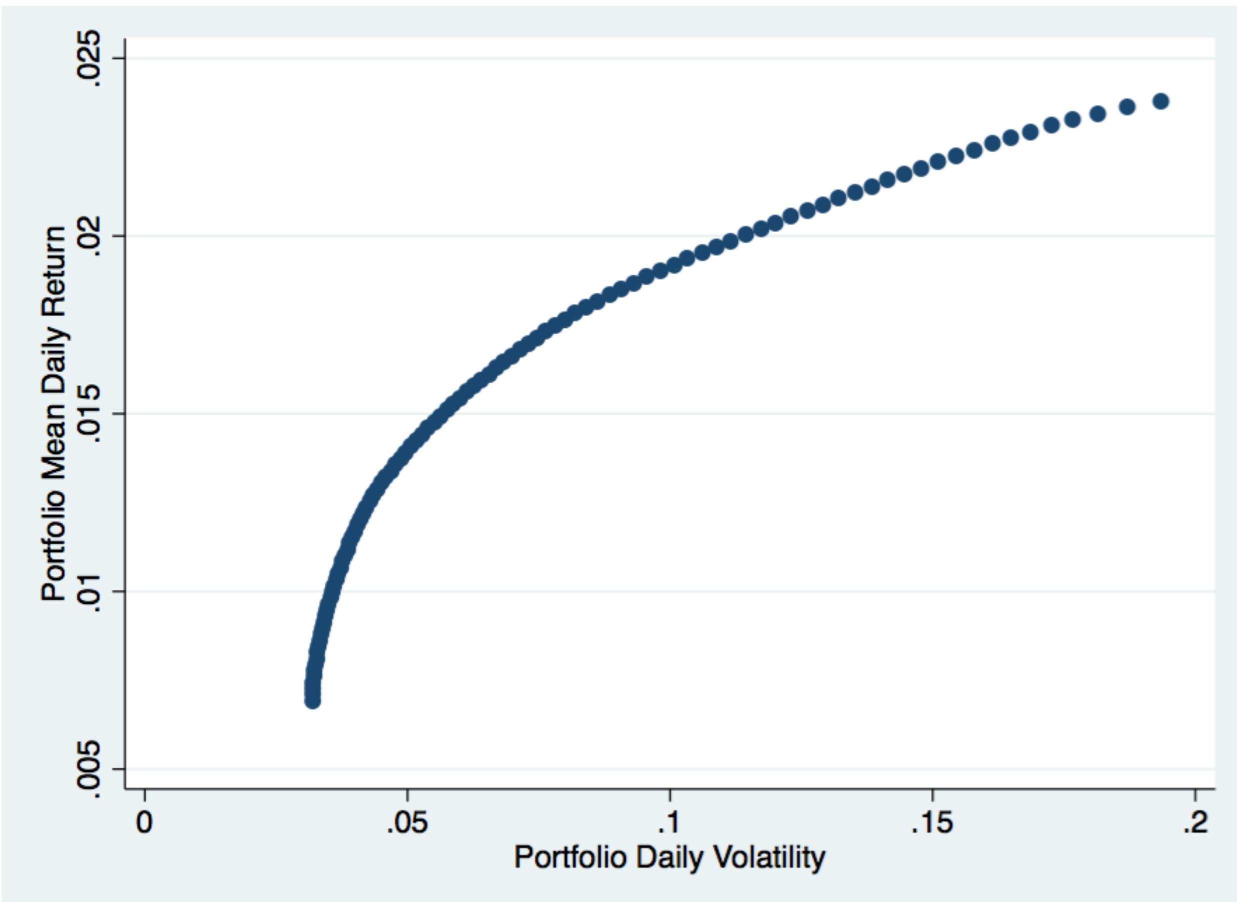
variance portfolio is 56% weighted in Bitcoin, where the weight decreases towards zero as one increases the portfolio volatility.

## ICO returns

In this section we document the return to initial coin offerings. In Table 3.8, we examine first week and first month returns from participating in token offerings, either during the ICO period, or in the secondary market. Without excluding outliers, the mean return is 4,746%, with a median of 115%, on the very first day of secondary market trading. By contrast, the mean return of only participating in the secondary market is much lower. This data contains a number of large outlier returns, with first trading day returns above 1000%, given in Table 3.9.

We exclude these outliers to give a more reasonable graphical representation of the returns. Figure 3.7 depicts the mean percentage return of participating directly in a token offering, versus the returns of only trading in the secondary market (Figure 3.8). These suggest that tokens are issued at steep discounts to secondary market trading. This appears similar to underpricing in IPO markets.

Figure 3.6: Efficient Frontier of Daily Returns




This plot considers 50 portfolios on an efficient frontier of the top 50 cryptocurrencies by turnover, as of November 23, 2015, with market capitalizations of at least $1 million. Portfolio weights were required to be weakly positive. The global minimum variance portfolio is 56% weighted in Bitcoin.

Table 3.8: Summary Statistics of Initial Coin Offerings (%)

| | (1) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | p25 | p50 | p75 | Max | N |
| First Day Return (Post ICO) | 4,746 | 31,652 | -46.5 | 32.3 | 115 | 375 | 226,300 | 51 |
| First Week Return (Post ICO) | 2,815 | 18,319 | -54.6 | 17.7 | 94.8 | 277 | 129,733 | 50 |
| First Month Return (Post ICO) | 19,999 | 140,818 | -78.4 | 7.06 | 144 | 368 | 1,005,917 | 51 |
| First Week Return (Secondary Market) | 1.75 | 58.4 | -94.3 | -31.4 | -10.3 | 28 | 274 | 64 |
| First Month Return (Secondary Market) | 46.3 | 191 | -94 | -57.8 | -16.1 | 49.6 | 1,091 | 64 |

This table shows the returns to investing in ICOs. The first three rows are assuming purchasing ICOs at the average price, and liquidating at a time relative to the beginning of secondary market trading. The last two rows show returns simply from investing at the start of secondary market trading. All numbers are in percents.

Table 3.9: ICO Returns at First Trading Day, with Returns > 1000%

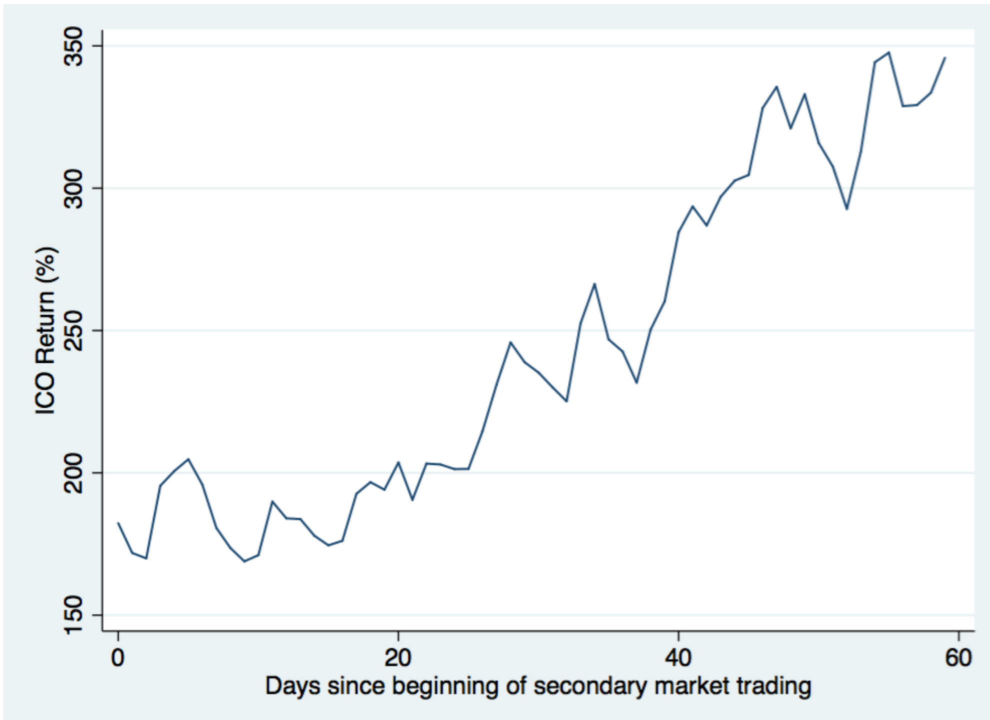
| **ICO Name** | **Token Symbol** | **First Trading Day Return (%)** |
|---|---|---|
| Nxt | NXT | 226,300.00 |
| Lisk | LSK | 4,948.88 |
| Particl | PART | 1,180.16 |
| Counterparty | XCP | 1,050.71 |

## Cross-Asset Correlations

In an effort to understand sources of common variation, this section provides a summary of cross-correlations of various cryptocurrencies with other assets. Correlations between return series are derived for each cryptocurrency and each of i) Bitcoin returns, ii) gold returns, iii) S&P 500 Excess Return, at both daily and monthly frequencies. The confidence intervals for these are provided in Table 3.10.

The only asset for which cross-correlations are noticeably above zero are for Bitcoin returns. The histograms for daily correlations are given in Figures 3.9 through 3.11, where the distribution for Bitcoin correlations show many observations demonstrating high correlations between 0.30 and 0.50. Table 3.11 lists currencies with highest (positive and negative) correlations, and those that exhibit the lowest correlations, at monthly frequencies. The top 10 most positive and negative correlated are mostly coins. Coins that exhibit highest positive correlation tend to be ones with longer history, usually more than four years, and include some coins with high turnover and high market capitalization. These correlations are very close to one, again which may suggest some source of systematic risk. Coins that exhibit correlations closest to zero tend to be for certain gaming-specific purposes, such as NoLimitCoin, a coin dedicated to fantasy sports, or GameCredits, a universal coin for various other types of gaming transactions.

Figure 3.7: Mean Return from ICO Participation



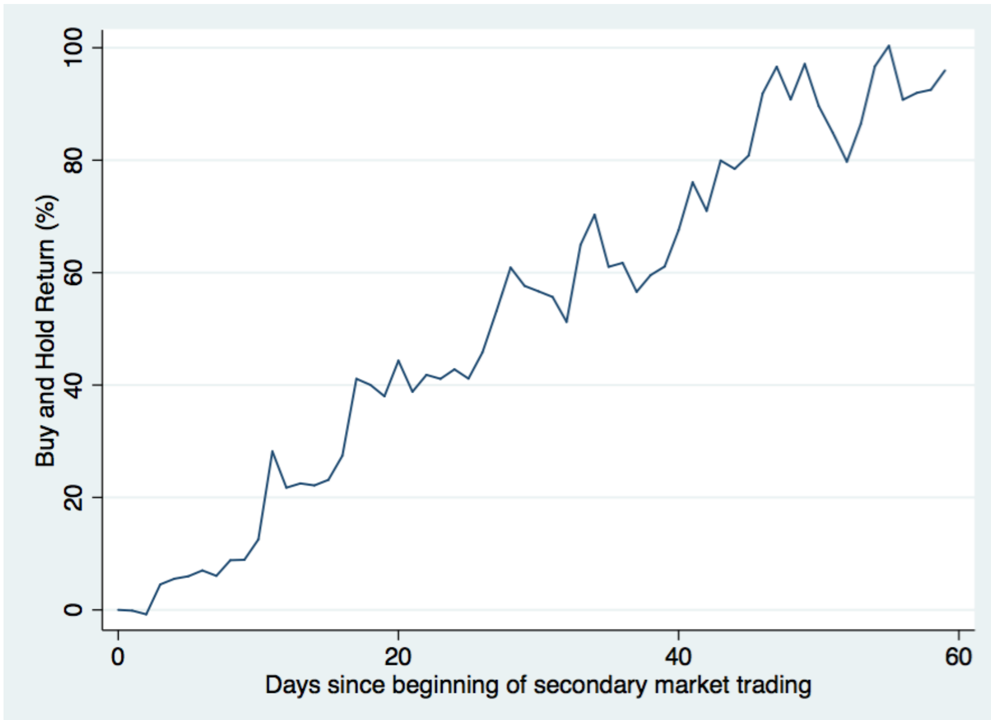

These high correlations explain the high apparent comovement between the equal-weighted and market-weighted portfolio in Figure 3.5, where the latter portfolio is weighted 89% in Bitcoin. Since Bitcoin is a much smaller fraction of the equal-weighted portfolio, if there were no correlation between Bitcoin and non-Bitcoin cryptocurrencies, comovement between the two portfolio returns would be much less.

## CAPM Performance

Despite the lack of strong correlation with the market excess return, for altcoins that do demonstrate positive correlation with the market, there is a demonstrable relationship between market betas and expected return in the sense of CAPM. Figures 3.12 and 3.13 show the daily and monthly betas for values between 0 and 10, sorted into 10 deciles. Their portfolio average betas are plotted against mean returns of the portfolios, showing a positive relationship.

What is the source of correlation with Bitcoin? One possible explanation is that many altcoins do not trade directly against fiat currencies, but against Bitcoin itself. Purchasing any of these altcoins thus may require purchases in Bitcoin, which may drive the common price movement.

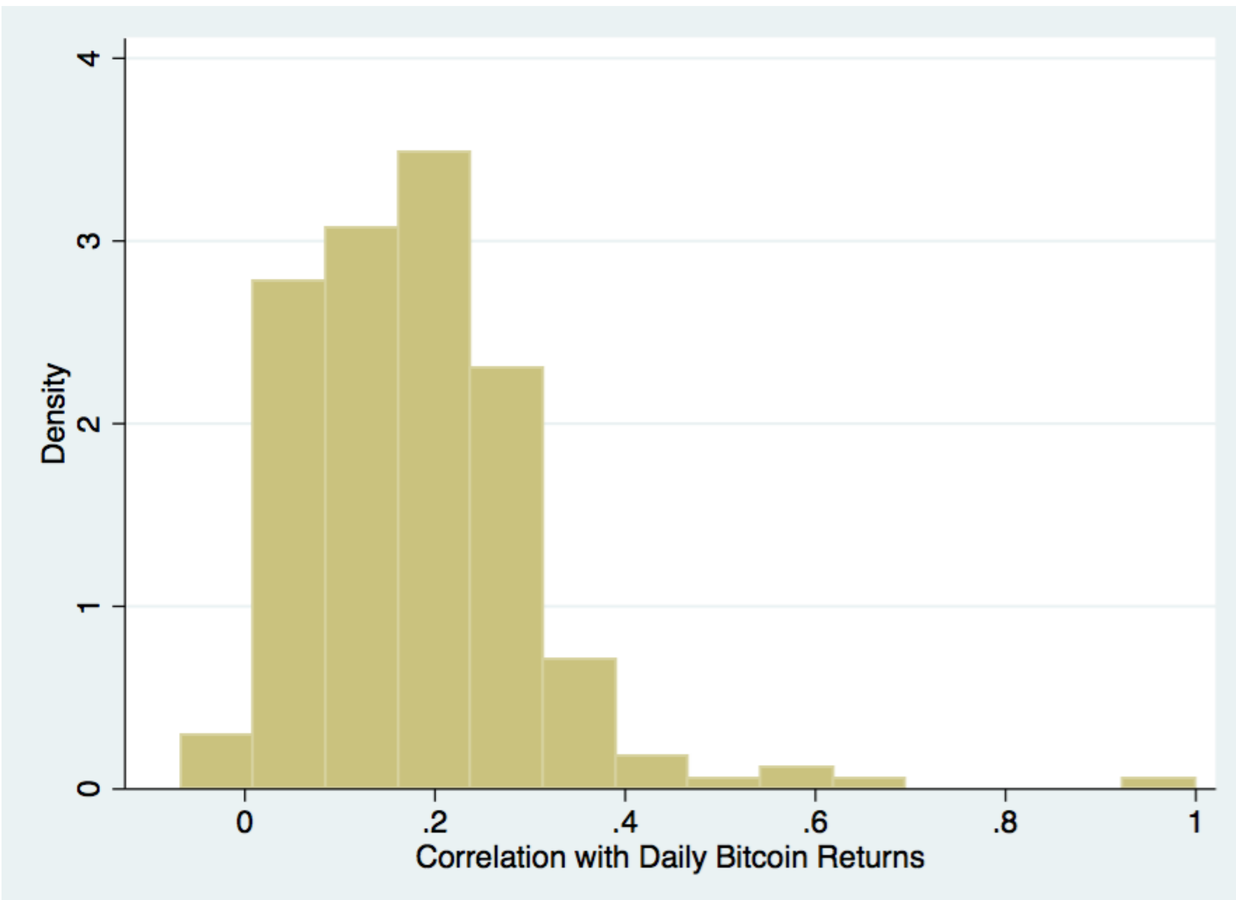
Figure 3.8: Mean Return from Secondary Market Participation (Post-ICO)




## Principal Component Analysis

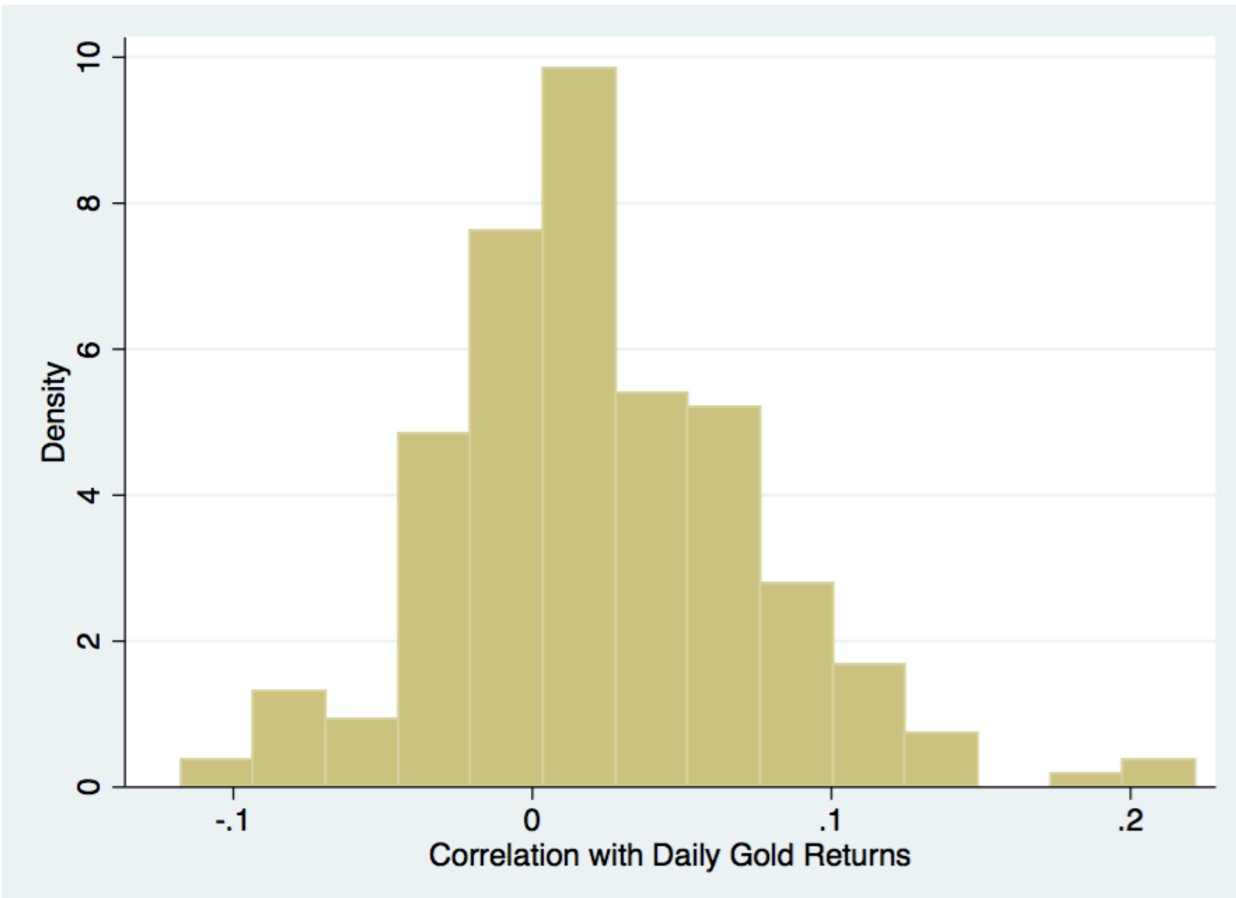
A second method of showing altcoin relationships with Bitcoin is to show that the first principal component is correlated with Bitcoin returns. Table 3.12 gives the first five components from PCA results for daily and monthly returns. To make sure a sufficiently complete time series is included, we restrict the universe of cryptocurrencies in this analysis to coins for which at least two years worth of price history are available. The first principal component explains 11.4% of daily returns, while the first principal component for monthly returns explains 31.7% of daily returns. For both these frequencies, the first principal component is positively related to Bitcoin returns, as demonstrated by Figures 3.16 and 3.17.

Figure 3.9: Correlations with Daily Bitcoin returns.



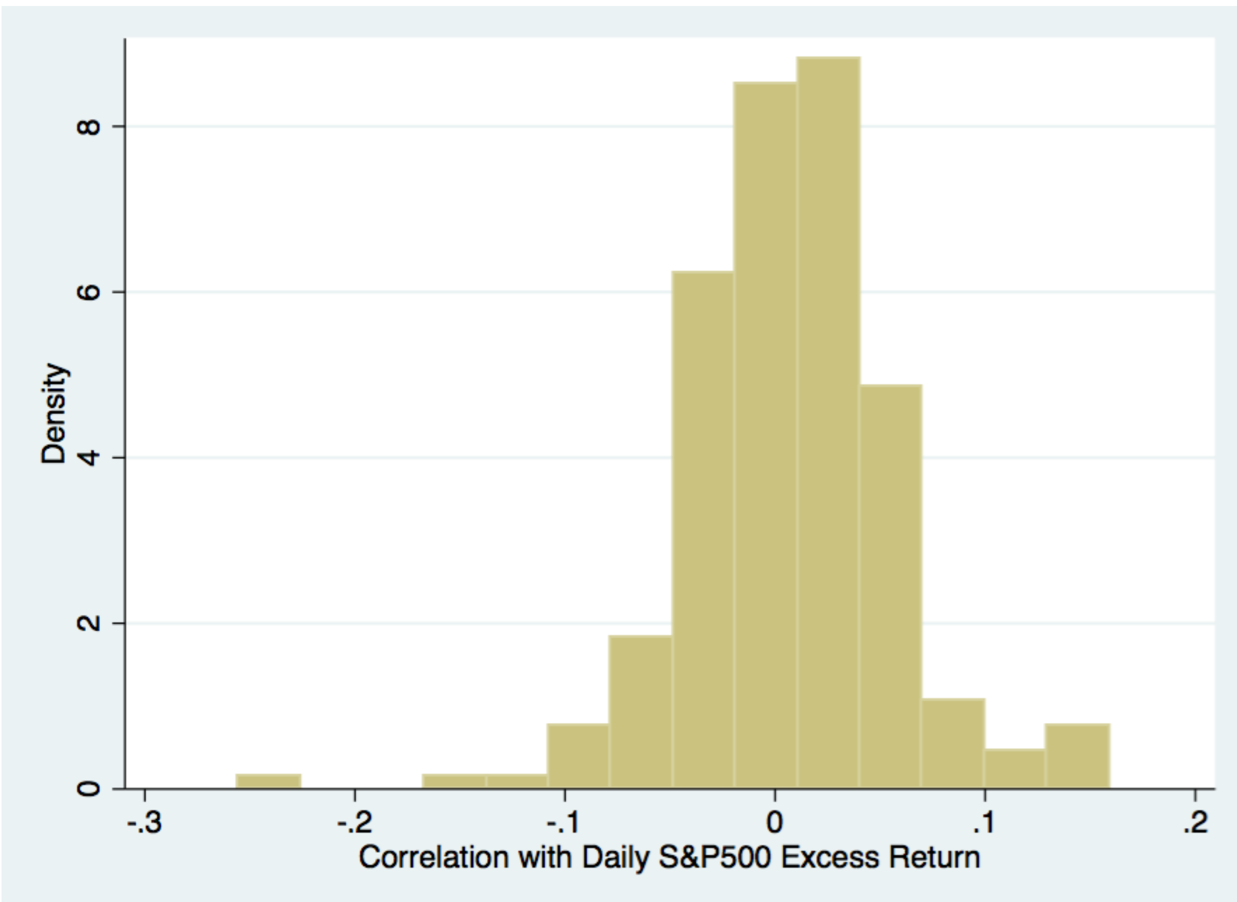

This plot shows a histogram of pairwise correlation coefficients between individual cryptocurrencies with daily bitcoin returns. The correlation coefficients are calculated from daily returns from the first price of every month to the next.

Figure 3.10: Correlations with Daily Gold returns.



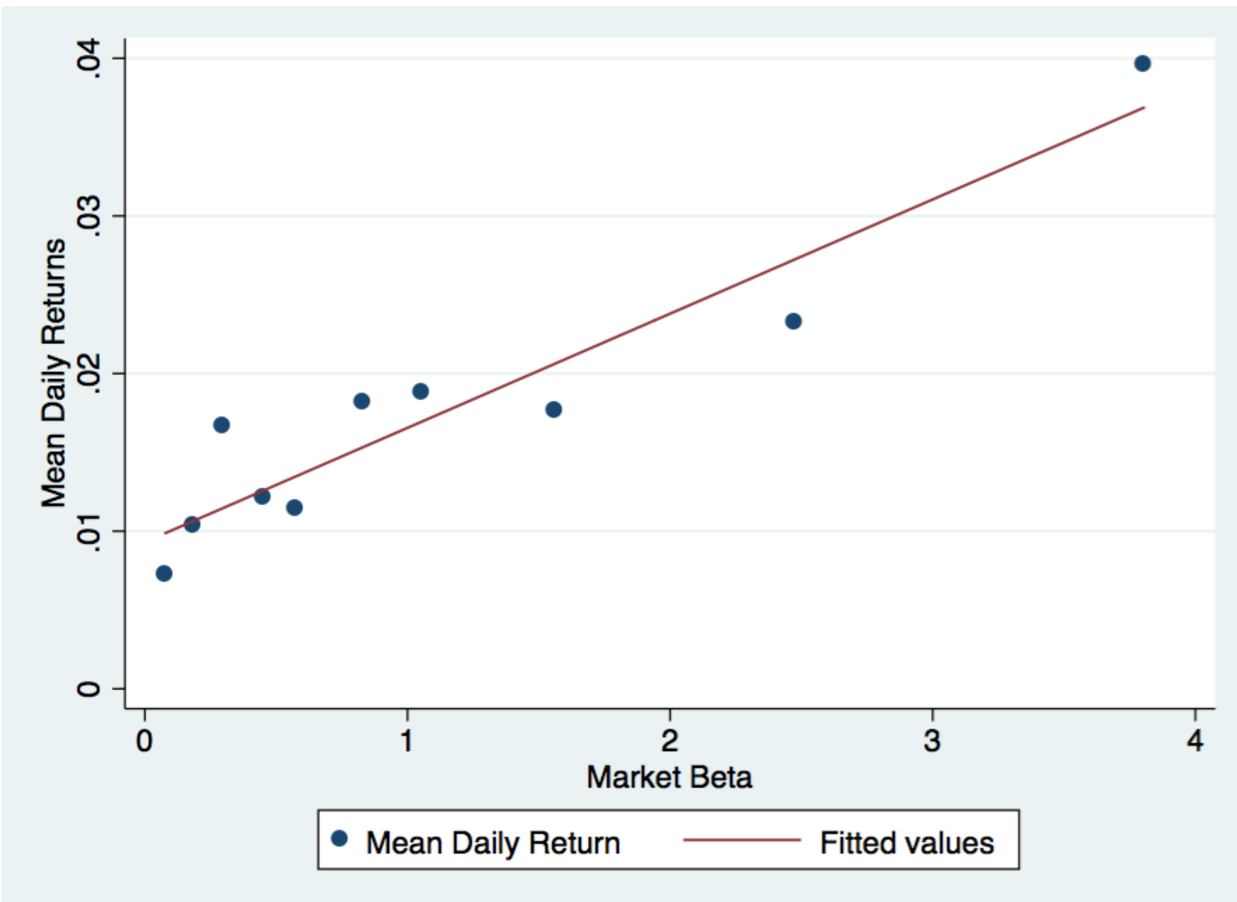

This plot shows a histogram of pairwise correlation coefficients between individual cryptocurrencies with daily gold returns. The correlation coefficients are calculated from daily returns from the first price of every month to the next.

Figure 3.11: Correlations with Daily S&P500 excess returns.



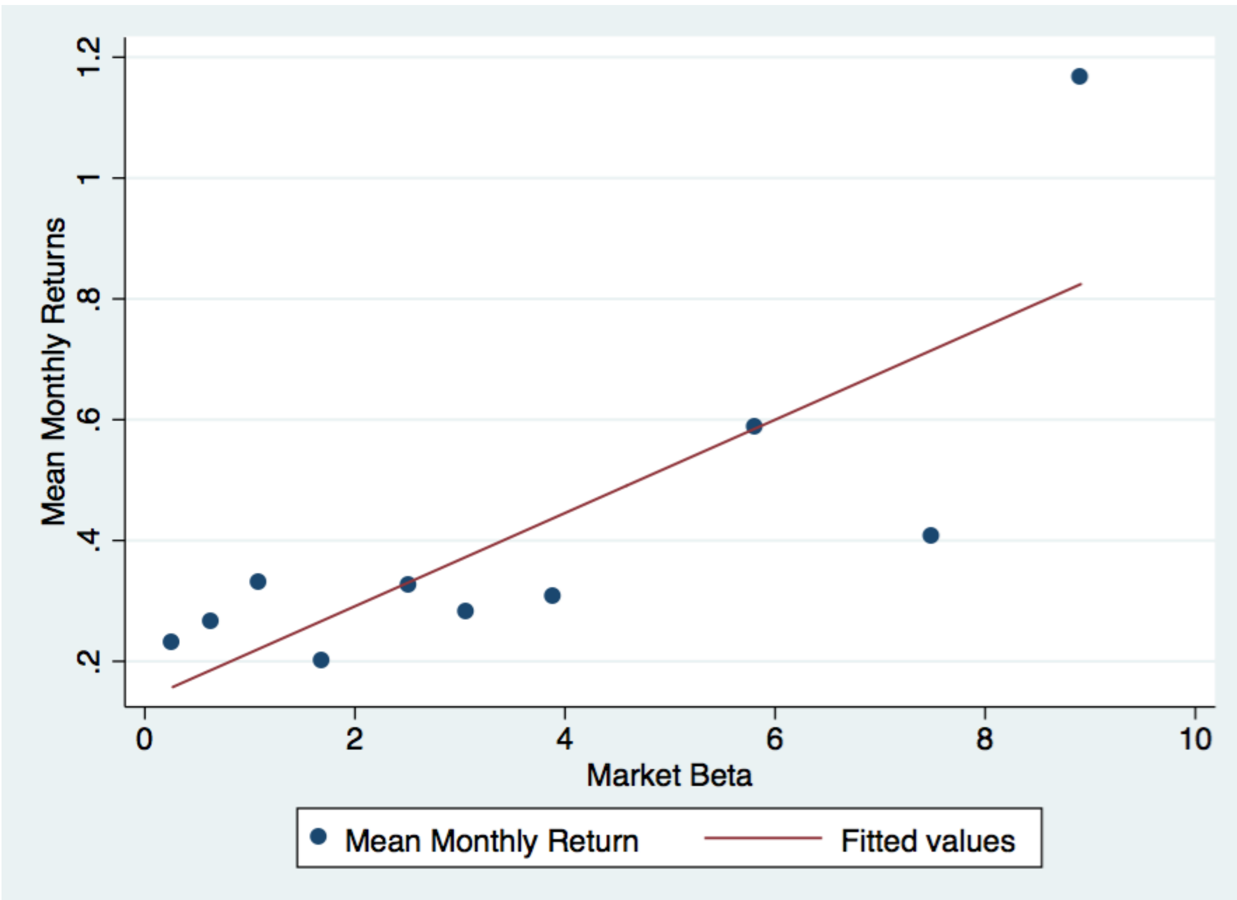

This plot shows a histogram of pairwise correlation coefficients between individual cryptocurrencies with daily excess returns for the S&P500 index. The correlation coefficients are calculated from daily returns from the first price of every month to the next.

Figure 3.12: Daily Return Portfolios formed on Market Betas




This plot shows the average betas of beta deciles, for cryptocurrency betas calculated using the daily S&P500 excess return. Only betas between 0 and 10 were included. The total number of currencies is 257, distributed roughly evenly across 10 decile portfolios.

Figure 3.13: Monthly Return Portfolios formed on Market Betas




This plot shows the average betas of beta deciles, for cryptocurrency betas calculated using the monthly S&P500 excess return. Only betas between 0 and 10 were included. The total number of currencies is 207, distributed roughly evenly across 10 decile portfolios.

Table 3.10: 95% Confidence Intervals of Correlations for Individual Cryptocurrencies with Other Assets

| **Correlations - Monthly** | **95% Confidence Interval** | |
|---|---|---|
| Bitcoin Return | 0.178 | 0.256 |
| Gold Return | 0.046 | 0.096 |
| S&P500 Excess Return | -0.025 | 0.025 |
| **Correlations - Daily** | **95% Confidence Interval** | |
| Bitcoin Return | 0.163 | 0.195 |
| Gold Return | 0.015 | 0.029 |
| S&P500 Excess Return | 0.000 | 0.013 |

This table gives the 95% confidence intervals for altcoin correlations with Bitcoin, gold, and the S&P500 excess return.

Table 3.11: Top 10 Positive, Negative, and Low Correlation Coins versus Bitcoin (Monthly)

| **Name** | **Symbol** | **Corr.** | **Type** | **Turnover** | **Series Start** | **Series End** |
|---|---|---|---|---|---|---|
| Litecoin | LTC | 0.95 | Coin | 1.0301 | 28apr2013 | 23nov2017 |
| Peercoin | PPC | 0.95 | Coin | 0.0187 | 28apr2013 | 23nov2017 |
| Namecoin | NMC | 0.94 | Coin | 0.0271 | 28apr2013 | 23nov2017 |
| Infinitecoin | IFC | 0.94 | Coin | 0.0000 | 10jul2013 | 23nov2017 |
| Quark | QRK | 0.94 | Coin | 0.0002 | 25aug2013 | 23nov2017 |
| Megacoin | MEC | 0.94 | Coin | 0.0040 | 07jul2013 | 23nov2017 |
| DraftCoin | DFT | 0.94 | Coin | 0.0477 | 23nov2015 | 23nov2017 |
| Crypto Bullion | CBX | 0.94 | Coin | 0.0015 | 04aug2013 | 23nov2017 |
| Terracoin | TRC | 0.93 | Coin | 0.0005 | 28apr2013 | 23nov2017 |
| GoldCoin | GLD | 0.93 | Coin | 0.0003 | 14jun2013 | 23nov2017 |
| SuperCoin | SUPER | 0.01 | Coin | 0.0005 | 26may2014 | 23nov2017 |
| LoMoCoin | LMC | 0.01 | Coin | 0.0029 | 09sep2016 | 23nov2017 |
| AudioCoin | ADC | 0.01 | Coin | 0.0000 | 05jun2015 | 23nov2017 |
| GameCredits | GAME | 0.01 | Coin | 0.0090 | 01sep2014 | 23nov2017 |
| RussiaCoin | RC | 0.00 | Coin | 0.0001 | 05jan2016 | 23nov2017 |
| Nexus | NXS | 0.00 | Coin | 0.0021 | 25jan2015 | 23nov2017 |
| Iconomi | ICNa | -0.00 | Token | 0.0107 | 30sep2016 | 23nov2017 |
| NeosCoin | NEOS | -0.01 | Coin | 0.0279 | 26aug2014 | 23nov2017 |
| NoLimitCoin | NLC2 | -0.01 | Coin | 0.0002 | 12sep2016 | 23nov2017 |
| VeriCoin | VRC | -0.01 | Coin | 0.0051 | 16may2014 | 23nov2017 |
| EverGreenCoin | EGC | -0.23 | Coin | 0.0045 | 10jan2016 | 23nov2017 |
| Emercoin | EMC | -0.23 | Coin | 0.0041 | 23aug2014 | 23nov2017 |
| Bitcloud | BTDX | -0.24 | Coin | 0.0004 | 12sep2016 | 23nov2017 |
| Golem | GNT | -0.24 | Token | 0.0050 | 18nov2016 | 23nov2017 |
| PutinCoin | PUT | -0.24 | Coin | 0.0001 | 07jul2016 | 23nov2017 |
| PIVX | PIVX | -0.25 | Coin | 0.0091 | 13feb2016 | 23nov2017 |
| FirstBlood | 1ST | -0.28 | Token | 0.0471 | 28sep2016 | 23nov2017 |
| B3Coin | B3 | -0.37 | Coin | 0.0029 | 03sep2016 | 23nov2017 |
| ArcticCoin | ARC | -0.39 | Coin | 0.0009 | 19oct2016 | 23nov2017 |
| Zoin | ZOI | -0.41 | Coin | 0.0007 | 10nov2016 | 23nov2017 |

This table shows the top 10 most positively correlated, top 10 most uncorrelated, and top 10 most negative correlated coins with Bitcoin. Correlations are calculated at the monthly frequency using all available data.

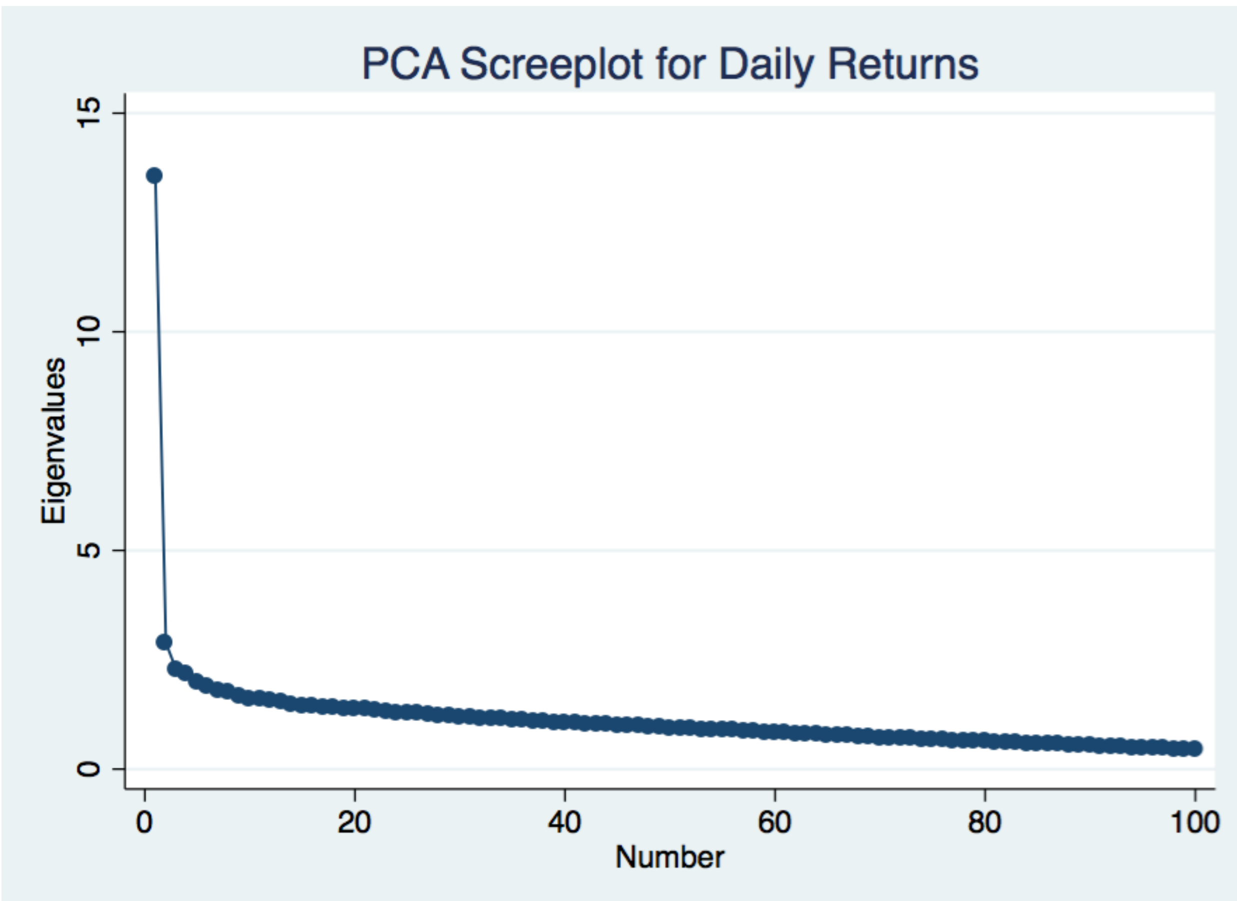
Table 3.12: Principal Component Analysis of Cryptocurrency Returns

| **Daily Returns** | | | | |
|---|---|---|---|---|
| **Component** | **Eigenvalue** | **Difference** | **Proportion** | **Cumulative** |
| Component 1 | 13.535 | 10.648 | 0.114 | 0.114 |
| Component 2 | 2.886 | 0.624 | 0.024 | 0.138 |
| Component 3 | 2.262 | 0.103 | 0.019 | 0.157 |
| Component 4 | 2.159 | 0.193 | 0.018 | 0.175 |
| Component 5 | 1.965 | 0.072 | 0.017 | 0.192 |
| Observations | | | | 724 |
| No. Components | | | | 119 |

| **Monthly Returns** | | | | |
|---|---|---|---|---|
| **Component** | **Eigenvalue** | **Difference** | **Proportion** | **Cumulative** |
| Component 1 | 37.753 | 22.545 | 0.317 | 0.317 |
| Component 2 | 15.208 | 7.611 | 0.128 | 0.445 |
| Component 3 | 7.597 | 0.525 | 0.064 | 0.509 |
| Component 4 | 7.072 | 0.363 | 0.059 | 0.568 |
| Component 5 | 6.708 | 1.195 | 0.056 | 0.625 |
| Observations | | | | 24 |
| No. Components | | | | 23 |

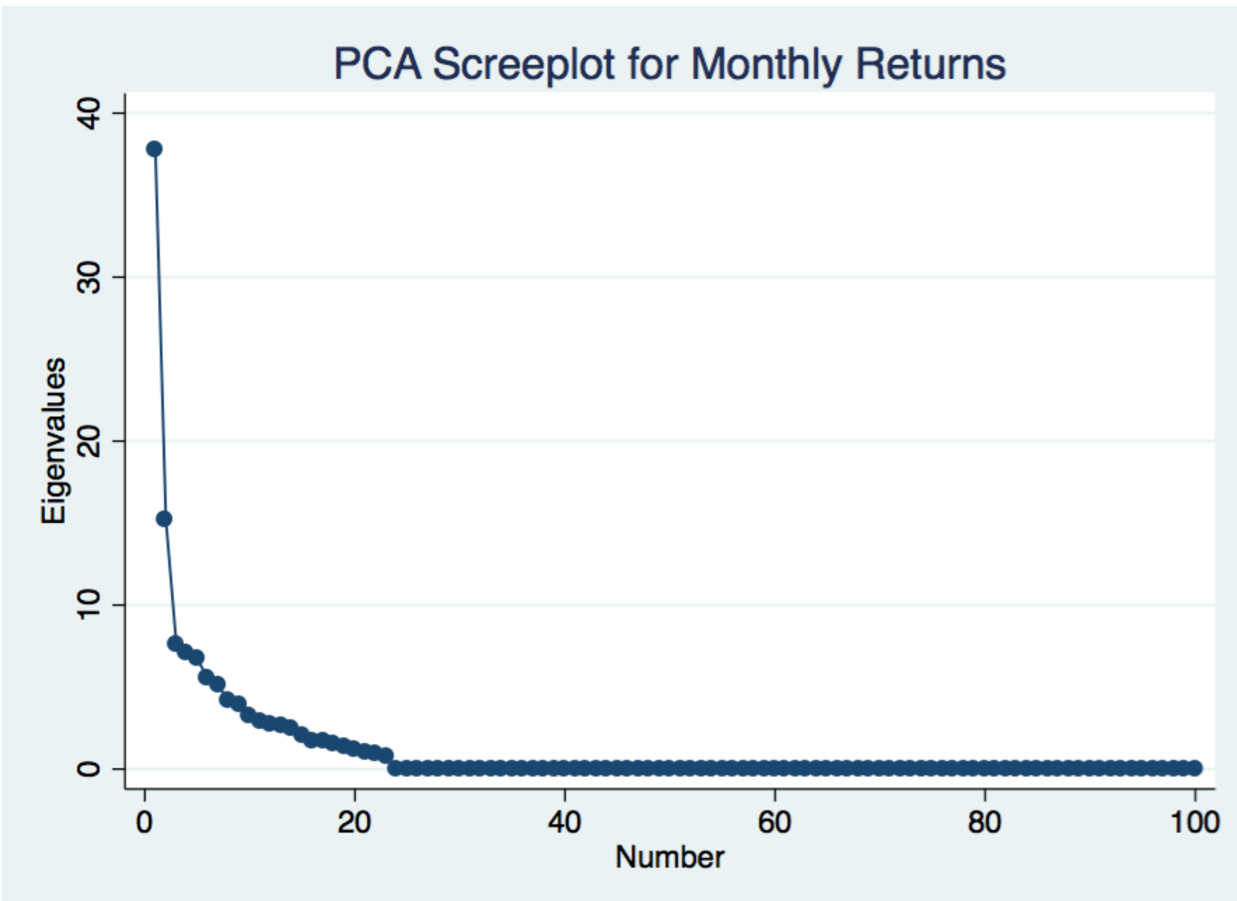
This table shows results for the first five principal components for daily and monthly returns. We limit the analysis to only currencies that have at least two years worth of time series data.

Figure 3.14: Screeplot of PCA for daily returns



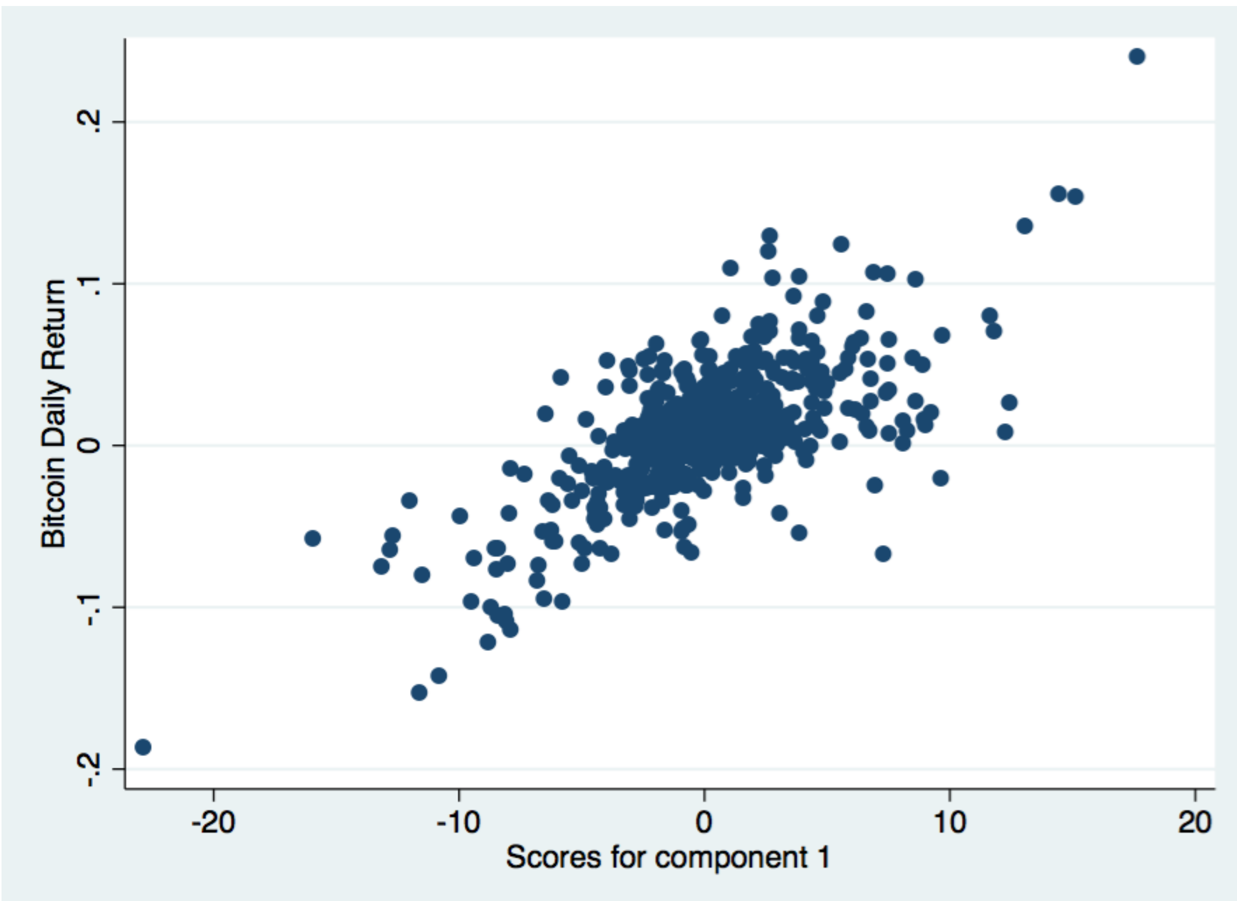

This figure shows a screeplot for the daily PCA, of the principal component number against its eigenvalue. The first principal component explains 11.4% of daily return variation.

Figure 3.15: Screeplot of PCA for weekly returns



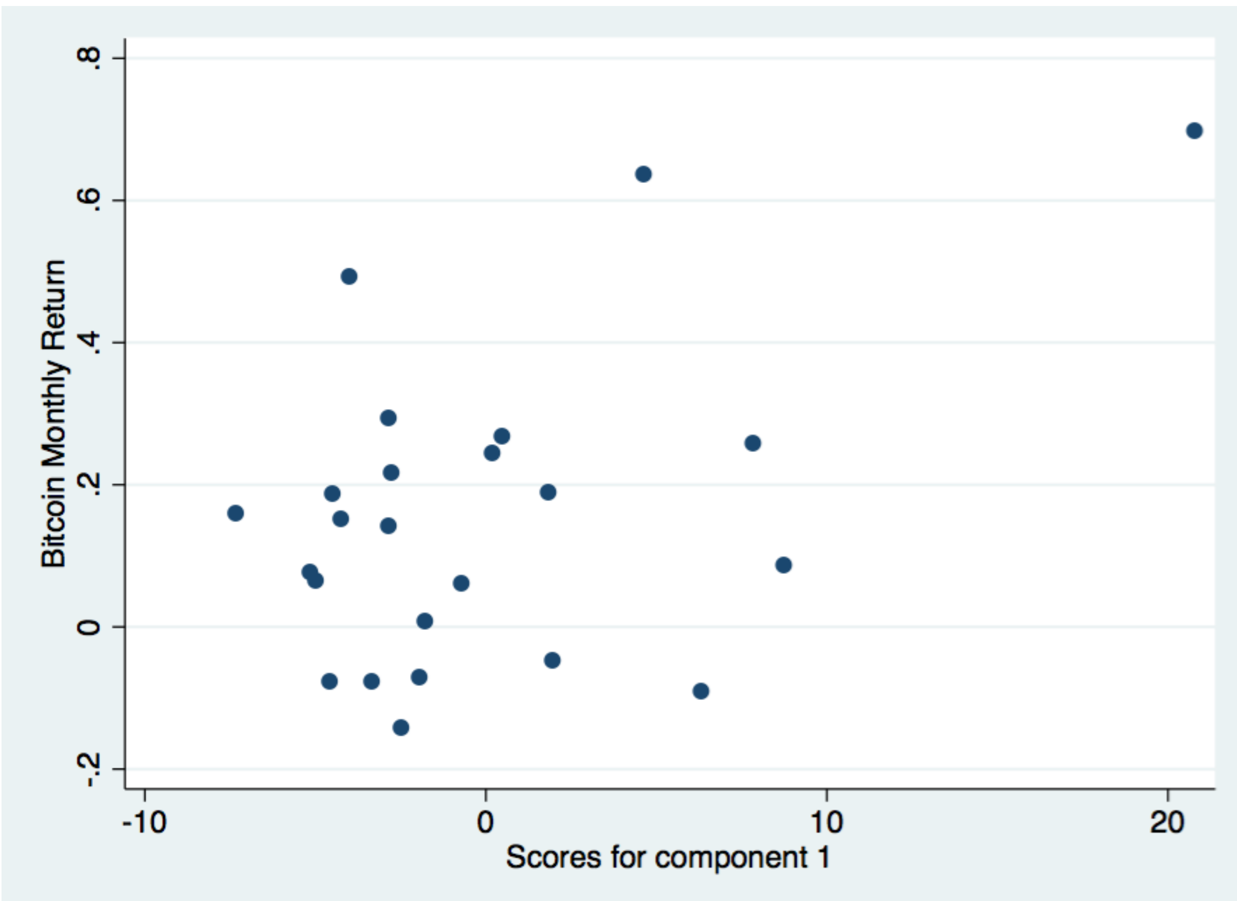

This figure shows a screeplot for the daily PCA, of the principal component number against its eigenvalue. The first principal component explains 31.7% of monthly return variation.

Figure 3.16: Plot of Daily Bitcoin Returns vs. First Principal Component Scores




This figure shows a scatterplot of the scores of the first principal component from the daily PCA against daily Bitcoin returns.

Figure 3.17: Plot of Monthly Bitcoin Returns vs. First Principal Component Scores




This figure shows a scatterplot of the scores of the first principal component from the monthly PCA against monthly Bitcoin returns.

## 3.5 Evaluating Coin Returns

To evaluate the return properties of coins, it is useful to have a benchmark. First, observe, because there is currently little or no regulation around coins, they are not comparable to listed equities in which there are both stringent disclosure requirements and listing requirements. Indeed, the listing requirements for all exchanges mean that the analogy between ICOs and IPOs is literally semantic.

The more relevant comparison group is probably venture capital. However, even in this case, a notable difference is that given the structure of the VC industry, the supply of capital is restricted, whereas in the ICO model, it is not. We would thus expect the observed distribution of projects (coins) to be different. Cochrane (2005) estimates the characteristics of VC risk and return. He observes that data are only available if firms either go public or are awarded a second round. Without adjusting for the observation bias, he obtains mean log returns of 108%. These are associated with very high volatilities and thus high arithmetic returns.

However, VC returns (ignoring the selection at the VC level) are not an exact benchmark. This is because another way in which coins differ from equities is their use value. To see this, notice that in as much as there is a liquid market for coins, they may be valued as any other financial security. Specifically, if other assets that have a similar riskiness have an expected return of $Er_t$ at time $t$, then the price of a coin at time $t$ should be

$$P_t \quad = \quad \frac{E_t P_{t+1}}{1 + Er_t}.$$

This naturally implies an empirical proxy for the (unobserved) expected return of

$$1 + r_t \quad = \quad \frac{P_{t+1}}{P_t} + \epsilon_t,$$

where the independence of the error term stems from the unbiased expectations.

Now, observe that in addition to being a store of value (in the sense that it can be sold), most coins have a use component, denote this by $\nu$. Suppose that a fraction $\alpha$ of a coin is required to use it. If the personal discount rate is $\delta$, then we have

$$P_t \quad = \quad \max\left[\frac{E_t P_{t+1}}{1 + Er_t}\ , (1-\alpha)\frac{E_t P_{t+1}}{1 + Er_t} + \frac{E\nu_{t+1}}{1 + \delta_t}\right].$$

Clearly, the empirical proxy, $\frac{P_{t+1}}{P_t}$ will systematically underestimate the expected return.

This formulation, similar to a convenience yield in commodities, depends on the specific implementation of the coin. Further note, that in many cases the use value (convenience yield) could exhibit a network externality. Specifically, the value to a user of owning the coin is increasing in the distribution of ownership of the coin, i.e., on how many people join the network. In this case, variables such as volume should be correlated with the estimation error induced by using $\frac{P_{t+1}}{P_t}$.

## 3.6 Conclusion

In this paper, we provide summary statistics for returns of over 200 cryptocurrencies. We provide data for both the universe of currencies and for those involved in initial coin offerings. There is a large degree of skewness and volatility in the population of returns. A principal risk factor is the return of Bitcoin itself, which is highly correlated with many altcoins. This is demonstrable through examining simple correlations with Bitcoin returns at the daily and monthly frequencies, as well as through a principal component analysis. The existence of this risk factor has implications for asset management and regulation in cryptocurrencies.

In prior years, traditional finance theories have avoided explanations of the cryptocurrency landscape due to its decentralized nature, volatility, and high technological barrier. However, the entry of institutional market participants such as ICO issuers, asset managers, and traditional derivatives exchanges in this area suggest that the time is right for a financial treatment of this topic. Revelations in this paper may help introduce finance to this new class of assets by summoning traditional financial concepts.

# Bibliography

Adelino, Manuel, Igor Cunha, and Miguel A. Ferreira (2017), "The economic effects of public financing: Evidence from municipal bond ratings recalibration." *The Review of Financial Studies*, 30, 3223–3268.

Angelides, Philip (2002), "Bond insurance as a form of credit enhancement in california's municipal bond market." *California Debt and Investment Advisory Commission*.

Appel, Ian R., Todd A. Gormley, and Donald B. Keim (2016), "Passive investors, not passive owners." *Journal of Financial Economics*, 121, 111 – 141.

Athey, Susan, Ivo Parashkevov, Vishnu Sarukkai, and Jing Xia (2016), "Bitcoin pricing, adoption, and usage: Theory and evidence." *Stanford University Graduate School of Business Research Paper*.

Azar, José, Martin C Schmalz, and Isabel Tecu (2015), "Anti-competitive effects of common ownership." *Ross School of Business Paper*.

Ball, Ray and Philip Brown (1968), "An empirical evaluation of accounting income numbers." *Journal of Accounting Research*, 6, 159–178.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler (2005), "Comovement." *Journal of Financial Economics*, 75, 283–317.

Berk, Jonathan B. and Richard C. Green (2004), "Mutual fund flows and performance in rational markets." *Journal of Political Economy*, 112, 1269–1295.

Bernanke, Ben S. (2015), *The Courage to Act: a memoir of a crisis and its aftermath*. W. W. Norton & Company, Inc.

Broman, Markus S. (2014), "Liquidity clienteles, correlated demand and excess comovement of exchange-traded fund returns." *Working Paper*.

Bushee, Brian J (1998), "The influence of institutional investors on myopic r&d investment behavior." *Accounting review*, 305–333.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein (2010), "The value of school facility investments: Evidence from a dynamic regression discontinuity design." *The Quarterly Journal of Economics*, 125, 215–261.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich (2015), "Regression discontinuity and the price effects of stock market indexing." *The Review of Financial Studies*, 28, 212–246.

Chodorow-Reich, Gabriel (2014), "The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis *." *The Quarterly Journal of Economics*, 129, 1–59.

Ciaian, Pavel, Miroslava Rajcaniova, and d'Artis Kancs (2016), "The economics of bitcoin price formation." *Applied Economics*, 48, 1799–1815.

Cirillo, Donald King (2008), *The Handbook of Municipal Bonds*, chapter How to Analyze the Municipal Bond Insurers and the Bonds They Insure. John Wiley & Sons, Inc.

Cochrane, John H. (2005), "The risk and return of venture capital." *Journal of Financial Economics*, 75, 3 – 52.

Cornaggia, Jess, Kimberly Rodgers Cornaggia, and Ryan D. Israelsen (2017), "Credit ratings and the cost of municipal financing." *Review of Financial Studies, Forthcoming.*

Cuny, Christine (2016), "Voluntary disclosure incentives: Evidence from the municipal bond market." *Journal of Accounting and Economics*, 62, 87–102.

Da, Zhi and Sophie Shive (2013), "When the bellwether dances to noise: Evidence from exchange-traded funds." *Working Paper*.

Edmans, Alex (2014), "Blockholders." *Annual Review of Financial Economics*.

Fan, C. Simon, Chen Lin, and Daniel Treisman (2009), "Political decentralization and corruption: Evidence from around the world." *Journal of Public Economics*, 93, 14 – 34.

Gandal, Neil and Hanna Halaburda (2014), "Competition in the Cryptocurrency Market." CESifo Working Paper Series 4980, CESifo Group Munich.

Gârleanu, Nicolae B and Lasse H Pedersen (2015), "Efficiently inefficient markets for assets and asset management." Technical report, National Bureau of Economic Research.

Glosten, Lawrence R and Paul R Milgrom (1985), "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders." *Journal of financial economics*, 14, 71–100.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen (2014), "Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and normal economic times." *Working Paper, University of Chicago.*

Grossman, Sanford J and Joseph E Stiglitz (1980), "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, 70, 393–408.

Jensen, Michael C. (1968), "The performance of mutual funds in the period 1945–1964." *Journal of Finance.*

Kisgen, Darren J. (2006), "Credit ratings and capital structure." *The Journal of Finance*, 61, 1035–1072.

Kriz, Kenneth A. and Marc D. Joffe (2017), *Municipal Bond Insurance after the Financial Crisis: Can It Help Reduce Borrowing Costs for Local Governments?* Mercatus Research, Mercatus Center at George Mason University, Arlington, VA.

Kroeger, Alexander and Asani Sarkar (2017), "Law of one bitcoin price?" *Federal Reserve Bank of Philadelphia.*

Malkiel, Burton G. and Eugene F. Fama (1970), "Efficient capital markets: A review of theory and empirical work*." *The Journal of Finance*, 25, 383–417.

Marlowe, Justin (2015), "Municipal bonds and infrastructure development – past, present, and future." *International City/County Management Association (ICMA) and Government Finance Officers Association (GFOA) White Paper.*

Moldogaziev, Tima Tilek-uulu (2013), "The collapse of the municipal bond insurance market: How did we get here and is there life for the monoline industry beyond the great recession?" *Journal of Public Budgeting, Accounting and Financial Management*, 25, 199–233.

Mullins, William (2014), "The governance impact of index funds: Evidence from regression discontinuity."

Nanda, Vikram and Rajdeep Singh (2004), "Bond insurance: What is special about munis?" *The Journal of Finance*, 59, 2253–2280.

Samuelson, Paul A. (1954), "The pure theory of public expenditure." *The Review of Economics and Statistics*, 36, 387–389.

Von Hagen, Jürgen (2007), "Achieving economic stabilization by sharing risk within countries." *Intergovernmental Fiscal Transfers*, 107–132.

Whitbeck, Nicholas (2013), "Dexia - rise and fall of a banking giant." *Les Etudes Du Club, HEC Paris Club Finance.*

Wilkoff, Sean M. (2012), *Essays on the Municipal Bond Market*. Ph.D. thesis, UC Berkeley.

Wilson, John D. (1986), "A theory of interregional tax competition." *Journal of Urban Economics*, 19, 296 – 315.

Yermack, David (2013), "Is bitcoin a real currency? an economic appraisal." Working Paper 19747, National Bureau of Economic Research.

# Appendix A

# Index Comovement and Informed Trading

## A.1 Proof of Proposition 2

The unconditional probability of observing a buy or sell is

$$\begin{aligned} \mathrm{P(buy)} &= \mu\Big((1-\delta)(1-\theta)+\delta\theta\Big)+\nu(1-\delta)+\frac{(1-\mu-\nu)}{2} \\ \mathrm{P(sell)} &= \mu\Big((1-\delta)\theta+\delta(1-\theta)\Big)+\nu\delta+\frac{(1-\mu-\nu)}{2}. \end{aligned}$$

Using Bayes' rule, we can calculate the probability of the state of $Y$, conditional on observing a buy or sell order:

$$\begin{aligned} \mathrm{P}(\overline{Y}|\mathrm{buy}) &= \frac{\Big(\mu+\nu(1-\delta)+\frac{(1-\mu-\nu)}{2}\Big)\Big((1-\delta)(1-\theta)+\delta\theta\Big)}{\mu\Big((1-\delta)(1-\theta)+\delta\theta\Big)+\nu(1-\delta)+\frac{(1-\mu-\nu)}{2}} \\ \mathrm{P}(\overline{Y}|\mathrm{sell}) &= \frac{\Big(\nu\delta+\frac{(1-\mu-\nu)}{2}\Big)\Big((1-\delta)(1-\theta)+\delta\theta\Big)}{\mu\Big((1-\delta)\theta+\delta(1-\theta)\Big)+\nu\delta+\frac{(1-\mu-\nu)}{2}} \\ \mathrm{P}(\underline{Y}|\mathrm{buy}) &= \frac{\Big(\nu(1-\delta)+\frac{(1-\mu-\nu)}{2}\Big)\Big((1-\delta)\theta+\delta(1-\theta)\Big)}{\mu\Big((1-\delta)(1-\theta)+\delta\theta\Big)+\nu(1-\delta)+\frac{(1-\mu-\nu)}{2}} \\ \mathrm{P}(\underline{Y}|\mathrm{sell}) &= \frac{\Big(\mu+\nu\delta+\frac{(1-\mu-\nu)}{2}\Big)\Big((1-\delta)\theta+\delta(1-\theta)\Big)}{\mu\Big((1-\delta)\theta+\delta(1-\theta)\Big)+\nu\delta+\frac{(1-\mu-\nu)}{2}} \end{aligned}$$

Thus the bid price for asset $Y$ is:

$$\begin{aligned}
B &= \mathrm{E}[Y|\text{sell}] \\
&= \overline{Y} \cdot \mathrm{P}(\overline{Y}|\text{sell}) + \underline{Y} \cdot \mathrm{P}(\underline{Y}|\text{sell}) \\
&= \overline{Y}\left(\frac{\left(\nu\delta + \frac{(1-\mu-\nu)}{2}\right)\overline{w}}{\mu\underline{w} + \nu\delta + \frac{(1-\mu-\nu)}{2}}\right) + \underline{Y}\left(\frac{\left(\mu + \nu\delta + \frac{(1-\mu-\nu)}{2}\right)\underline{w}}{\mu\underline{w} + \nu\delta + \frac{(1-\mu-\nu)}{2}}\right) \\
&= \frac{\underline{Y}\mu\underline{w} + (\overline{w}\overline{Y} + \underline{wY})\nu\delta + (\overline{w}\overline{Y} + \underline{wY})\left(\frac{1-\mu-\nu}{2}\right)}{\mu\underline{w} + \nu\delta + \frac{(1-\mu-\nu)}{2}},
\end{aligned}$$

and similarly, the ask price:

$$\begin{aligned}
A &= \mathrm{E}[Y|\text{buy}] \\
&= \frac{\overline{Y}\mu\overline{w} + (\overline{w}\overline{Y} + \underline{wY})\nu(1-\delta) + (\overline{w}\overline{Y} + \underline{wY})\left(\frac{1-\mu-\nu}{2}\right)}{\mu\overline{w} + \nu(1-\delta) + \frac{(1-\mu-\nu)}{2}}.
\end{aligned}$$