Essays on Corporate Finance and Asset Management

by

Paulo M. Martins Barbosa Fortes Manoel

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 Gustavo Manso, Co-chair
Professor Annette Vissing-Jorgensen, Co-chair
Professor David Sraer
Professor Prasad Krishnamurthy

Spring 2019

Essays on Corporate Finance and Asset Management

Abstract

Essays on Corporate Finance and Asset Management

by

Paulo M. Martins Barbosa Fortes Manoel

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Gustavo Manso, Co-chair

Professor Annette Vissing-Jorgensen, Co-chair

This dissertation consists of three chapters spanning two areas of finance: corporate finance and asset management. A unifying theme across all chapters is the identification of issues that calls for the action of policymakers in order to improve the conditions of the less empowered, with a special focus on two themes: (1) factors affecting small entrepreneurs in the main street, and (2) agency and transparency problems underlying the management of small investors' wealth. Understanding these issues and how to mitigate them is of fundamental importance for regulations and policies aiming at the reduction of income and wealth inequality. I quantify the economic losses generated by (i) the underperformance of funds subject to agency problems, and by (ii) the business underactivity in low-income regions, carefully measuring the benefits of actual and hypothetical interventions by policymakers.

In the first chapter, "Crime Rates, Law Enforcement, and Business Activity", I ask if regions with prevalent violent and property crimes can promote business activity by reducing crime rates with more law enforcement. A short-term increase in the police force leads to a reduction in crime rates and an increase in the total sales of the retail sector. Surprisingly, the economic gains stemming from the reduction in murders is similar to the additional value added by the retail sector, highlighting the importance of accounting for business activity in the law enforcement cost-benefit analysis.

This causal effect from crime rates to business activity, taken together with the finding of the crime literature that increased business activity leads to fewer crimes, implies a feedback between crime and business, which suggests the existence of multiple Pareto-ranked equilibria. I provide evidence that strong yet temporary police shocks can create a permanent reduction in crime and a permanent increase in entrepreneurship, consistent with a shift away from the undesirable equilibria. The findings highlight how crime prevention law enforcement could be used as a tool to achieve economic development for economies stuck in a perverse poverty trap equilibrium with high-crime and low-business-activity.

In the second chapter, "Outraged by Compensation: Implications for Public Pension Performance", based on joint work with Adair Morse and Alexander Dyck, we analyze the

frictions in optimal contracting emerging from board members' sensitivity to employee and public outrage over high management compensation. We show that relaxing outrage constraints results in substantial incremental value added. The case highlights the importance of governance structures that insulates boards from external political pressures in state pension funds.

In the third chapter, "Mutual Fund Portfolios: The Case of the Missing Value Funds", based on joint work with Martin Lettau and Sydney Ludvigson, we assess the portfolios of active mutual funds and ETFs from the lens of risk (anomaly) factors that have been identified by the asset pricing literature. Our main finding is that mutual funds' portfolios are highly concentrated in growth stocks. Surprisingly, this is true even for funds that advertise themselves as value-oriented. This finding is of great importance for the evaluation of the transparency of the investment strategies in this industry, given that non-sector funds are usually classified in the "square" composed by the four combinations of small cap/ large cap and value/growth. We conclude that U.S. retail investors cannot trust the official labels of mutual funds and ETFs.

To my mother, Laura Andrea Martins Barbosa Silveira.

Contents

Contents **ii**

List of Figures **iv**

List of Tables **vi**

1 Crime Rates and Business Activity **1**
1.1 Effect of Short-Lived Police Shocks . . . 5
1.2 Effect of a More Persistent (Yet Temporary) Police Shock . . . 9
1.3 A Model for the Crime Rate and Business Activity . . . 18
1.4 Conclusion . . . 21

2 Outrage and Public Pensions **50**
2.1 Introduction . . . 50
2.2 Model of Portfolio Choice with Political Agency Costs . . . 54
2.3 Empirical Methodology . . . 60
2.4 Data . . . 61
2.5 Results . . . 65
2.6 Conclusion . . . 71

3 Mutual Fund Portfolios **88**
3.1 Introduction . . . 88
3.2 Data Construction . . . 93
3.3 Mutual Fund Portfolios . . . 96
3.4 Loadings vs. Holdings . . . 102
3.5 Portfolio Composition by Quintiles . . . 105
3.6 Joint Distribution of Mutual Fund Characteristics . . . 107
3.7 Characteristics and Return . . . 108
3.8 Conclusion . . . 109

Bibliography **134**

A Chapter 1 Appendix **139**

A.1 Proofs 140

B Chapter 2 Appendix **142**
B.1 Model Solution 143
B.2 Comparative Statics Computations 144

List of Figures

1.1 Esprito Santo State Police Strike 23
1.2 Law Enforcement Strike Map 24
1.3 Law Enforcement Strikes' Characteristics 25
1.4 The Effect of State-Level Law Enforcement Strikes 26
1.5 Rio de Janeiro State Revenues and Police Budget 27
1.6 % Variation in the Number of Police Officers per 100 People 28
1.7 Number of Pacifying Police Units (UPPs) 29
1.8 Location of Favelas in the City of Rio de Janeiro 30
1.9 Favelas' Territorial Command 31
1.10 Favelas' Firm Sector Breakdown 32
1.11 Impact of Pacification on Crime 33
1.12 Effect of Pacifying Police Units on the Number of Firms 34
1.13 Firm Creation Induced by the UPPs in Each Sector 35
1.14 Alternative Specifications 36
1.15 Effect of Pacification on the Number of Firms in Neighboring favelas 37
1.16 Business Activity in Regions with Implemented and Canceled Units 38
1.17 Simulated Effect of a Temporary Increase in Law Enforcement 39

2.1 Compensation of Investment Manager by Prior Profession 73
2.2 Compensation by Tertitle of Trustee Occupation-Outrage Variables 74
2.3 Compensation by Tertitle of Reference Wage-Outrage Variables 75

3.1 Distribution of Book-to-Market Ratios of Mutual Funds 111
3.2 Characteristics of "The Investment Company of America Fund" (AIVSX) 112
3.3 Characteristics of Mutual Funds 113
3.4 Histograms - Characteristics of Mutual Funds and Stocks 114
3.5 Histograms - Characteristics of Hedge Funds and ETFs 115
3.6 Characteristics of Mutual Funds – Robustness 116
3.7 Time-series of Characteristics of two large Mutual Funds 117
3.8 Loadings in Rolling 4-Factor Regressions 118
3.9 HML and SMB Loadings in Rolling 4-Factor Regressions 119
3.10 Histograms - Loadings of Mutuals Funds and Hedge Funds 120

3.11 BM-Quintile Portfolio Shares of Mutual Funds, Hedge Funds and ETFs 121
3.12 Joint Characteristics Distributions 122
3.13 Joint Characteristics Distributions: 1999Q1 and 2001Q2 123

List of Tables

1.1 Descriptive Statistics (2000-2017) 40
1.2 Losses Under 30 Law Enforcement Strikes 41
1.3 Placebo Test: Effect of Teachers' Strikes 42
1.4 Descriptive Statistics (2003-2017) 43
1.5 Pacified and Nonpacified Regions at the Baseline 44
1.6 Test for Reversion of Crime After the Fiscal Crisis 45
1.7 Test for Reversion of Business Activity After the Fiscal Crisis 46
1.8 Effect of Pacification on Business: Regression Discontinuity Evidence 47
1.9 Effect of Pacification on Household Income 48
1.10 Welfare Analysis (2008-2017) 49

2.1 Comparative Statics: Political Agency Variables Role 76
2.2 Variable Definitions 77
2.3 Pension Fund Profile Statistics 78
2.4 Performance and Allocation Statistics 79
2.5 Compensation, Trustee Occupation, Wages and Other Agency Statistics 80
2.6 Professions of Investment Managers and Trustees 81
2.7 Effect of Outrage and Political Agency on Manager Compensation 82
2.8 Effect of Outrage on Returns 83
2.9 Effect of Outrage on Returns 84
2.10 Effect of Outrage on Realized Tracking Error 85
2.11 Effect of Outrage on Asset Class Weights Results 86
2.12 Effect of Outrage on Portfolio Delegation 87

3.1 Descriptive Statistics of Mutual Funds and Hedge Funds 124
3.2 Characteristics of Passive Benchmark Portfolios 125
3.3 Characteristics of Mutual Funds and Hedge Funds 126
3.4 Characteristics of highest/lowest BM Mutual Funds 127
3.5 Probit Regressions 128
3.6 4-Factor Regressions of Passive Benchmark Portfolios 129
3.7 Loadings of 25 ME-BM sorted Portfolios 130
3.8 Portfolio Composition of Mutual Funds by Quintiles 131

3.9 Returns of Stocks and Mutual Funds 132
3.10 Fama-MacBeth Regressions 133

Acknowledgments

In this journey, I am indebted to many. First and foremost, I owe a special debt to my family: my mother Laura, my brothers Daniel and Andre, and my stepfather Josino for their unconditional support throughout the hardships of the doctoral process.

Second, I'm specially grateful to my advisors Gustavo Manso, Annette Vissing-Jorgensen, David Sraer, and Prasad Krishnamurthy for their patience and dedication during the long hours we spent discussing my research projects. I'm also in special debt to the entire faculty of the Finance department at the Haas School of Business, but I must offer my special thanks to Adair Morse, Martin Lettau, Brett Green, and Terrance Odean.

Many others have been integral to my intellectual development during the past years in Berkeley. I also to give a special thank to my fellow Haas Finance graduates Carlos Avenancio-Len, Andrew Schwartzman, Troup Howard, and Tristam Fitzgerald for generously sharing their knowledge and presence when needed. I am also grateful for great intellectual discussions and friendship with doctoral students from other departments, among them: Sibo Lu, Nan Chen, and Auyon Siddiq. I also want to acknowledge the support provided by Haas. Many thanks to Melissa Hacker, Bradley Jong, Kim Guilfoyle, Lisa Sanders Villalba, for all their help and support over the past six years.

Chapter 1

Crime Rates, Law Enforcement, and Business Activity

Previous research has spent a great deal of effort on uncovering the reasons behind the sharp decrease in the crime rates in the United States in the 1990s, culminating in a rich and extensive empirical literature about the effects of business activity on crime (Freeman 1999; Raphael and Winter-Ember 2001; Gould, Weinberg, and Mustard 2002; Dix-Carneiro, Soares, and Ulyssea 2018). However, the *consequences* of high crime for business activity are still unexplored. Understanding this missing link is critical for policymaking, as it would establish a negative loop between crime and business, potentially leading to multiple equilibria. In this case, a temporary law enforcement intervention could induce a permanent shift from the undesirable high-crime low-business equilibrium. This paper explores the missing link from crime to business, providing evidence that crime shocks have strong effects on business activity, even over the long run.

Several mechanisms might shape the business response to shocks in crime rates. First, the demand for goods and services in a given region could be affected by the perceived risk of crime victimization. On the supply side, entrepreneurs could be reluctant to start or expand their business if they do not expect to keep the resulting fruits (Besley 1995; Goldstein and Udry 2008; Hornbeck 2010). High crime rates might also reduce the collateral value of housing and of durable goods owned by entrepreneurs, intensifying credit constraints (Besley and Mueller 2012) and reducing business activity (Schmalz, Sraer, and Thesmar 2017). Furthermore, a profitable crime career might divert labor from productive uses, reducing the supply of entrepreneurs and workers to local business (Raphael and Winter-Ember 2001; Gould et al. 2002). Finally, high crime rates are likely to create distortions in investment decisions, given that resources otherwise efficiently allocated could be drained by security expenses.

I measure the effect of crime on business by taking advantage of variations in crime stemming from 30 state-level law enforcement strikes in Brazil between 2000 and 2017 .[1]

[1] In Brazil, preventive police work related to violent crimes and property crimes is the responsibility of

My methodology is based on a generalized difference-in-difference model with monthly data, using a continuous treatment variable (rather than binary[2]) that scales by the percentage of police participating in the strike. I find that a full-participation strike lasting for a full month in a given state triples the crime rates and reduces the overall sales of the retail sector by 14 percentage points.[3] This large effect exists only during the strike month: the treatment and control states move in parallel before and after the strike month. Given that strikes are a result of failed negotiations and usually last for several months, it seems unlikely that these results are being driven by reverse causality. Overall, I estimate economic losses of approximately $580 million for the 30 strikes in my sample, from which $280 million comes from a reduction in the value added by the retail sector, while the remaining $300 million comes from additional homicides during a strike.

A corollary from this first finding is the existence of a negative loop between crime and business, given that previous research has already documented a causal effect of business on crime (Freeman 1999; Raphael and Winter-Ember 2001; Gould et al. 2002). This feedback could, theoretically, induce the existence of multiple equilibria: one equilibrium with low business and high crime, which I call the "poverty trap", and one equilibrium with high business activity and low crime, which I call the "prosperous equilibrium". A natural policy question arising from this observation is how to shift economies away from the poverty trap. Intuitively, a sufficiently strong law enforcement shock could work as a "coordination device", creating incentives for some agents to migrate from crime to formal labor or non-criminal self-employment. If so, a temporary increase in the number of police in a given poor and violent region could lead to a permanent increase in the local business activity.

I test the equilibrium shift hypothesis by analyzing a law enforcement program introduced in the city of Rio de Janeiro in 2008. Named the Pacifying Police Unit (UPP in the Portuguese abbreviation), this program aimed at reclaiming territories known as favelas from the control of criminal gangs through the installation of several police units in nearby violent communities, and initially got a positive reception.[4] However, an unexpected fiscal crisis faced by the local government in 2015 drastically reduced the resources available to the program, compromising its capacity to prevent crimes. Therefore, the police shock in pacified favelas was not permanent, which is an ideal setup to test the predictions of the equilibrium shift hypothesis.

According to a past coordinator of the Pacifying Police Unit,[5] the program would mainly select favelas ruled by the two most violent drug-dealing gangs,[6] with the ultimate objective

---

the executive branch of each state.

[2]Acemoglu, Autor, and Lyle (2004) employs a similar strategy in a different context.

[3]Homicide rate increases by 117 log points, or 222 percentage points.

[4]Favelas are low-income informal urban areas in Brazil, formed over the years by poor citizens that cannot afford house prices in the more urbanized regions. These areas were originally built by former enslaved Africans in the 19th century but received large migratory inflows with the Brazilian rural exodus that occurred in the 1970s.

[5]Robson Rodrigues da Silva, former coordinator of the Pacifying Police Unit and former chief of staff of the Rio de Janeiro Military Police.

[6]Namely *Comando Vermelho* and *Amigos dos Amigos* or Red Command and Friends of Friends in

of taking all their territories. These plans were frustrated by the fiscal crisis faced by the local government in 2015, and all planned units that would have been implemented after 2014 were aborted. I take advantage of this fact when assessing the effect of pacification on business activity, measured by the number of local firms, and estimate a generalized difference-in-differences model that compares pacified and would-have-been pacified favelas. As for the effect of the Pacifying Police Units on crime, I face the limitation of observing crime rates only at the geographical level of the police districts, which are larger geographical units than favelas. For this reason, I measure the effect on crime by analyzing the number of murders and robberies after the introduction of the first police unit in a given district, using districts without any units as a control group.

Violent crime was reduced by 27 percentage points in the two years following the implementation of a police unit in a given police district. However, this effect reverts to zero seven years after the implementation year. This reversion aligns with the reduction in the police budget in 2015. Given that the reduction in violent crimes was the primary objective of the program, it is safe to conclude that the crime prevention effectiveness of the police units was only temporary. The effect on property crimes, however, exhibits a different pattern: the implementation of a police unit decreases local robbery persistently, reaching a reduction of 57 percentage points after seven years.

I argue that the difference between the results for violent crimes and for property crimes can be explained by the permanent expansion of business activity, given that violent crimes are insensitive to local business activity (Raphael and Winter-Ember 2001). Indeed, the implementation of one police unit in a given favela permanently increases the number of firms by approximately 33 percentage points. The effect is widespread in almost all sectors, including manufacturing, but is stronger for the retail sector, transportation, hotels, and restaurants. This result is consistent with a negative feedback loop between business and property crimes and with multiple possible equilibrium values for these variables: one poverty trap and one (or several) equilibria with low property crime rates and a large number of firms. Using this terminology, the Pacifying Police Unit program was able to move favelas away from the poverty trap. However, violent crimes are more sensitive to law enforcement than to local economic conditions, and they spiked when the resources of the program were reduced.

I run several robustness checks to address potential identification and interpretation issues. One possible concern is that the implementation of a police unit in a given favela might induce the migration of criminals to neighboring favelas. Similarly, pacification could induce a flow of firms from unpacified favelas to the newly pacified region, without a real change in the total stock of firms in the city of Rio de Janeiro. I address these two possibilities by showing that the pacification of a given favela does not change the number of firms in neighboring favelas. Another important concern is that comparing pacified and would-have-been pacified favelas might not address all possible endogeneities, as the sorting of the favelas in the program implementation schedule could depend on unobservables that are correlated with trends of business activity. I address this possibility by using an ordered

English.

list of the first favelas that would have been pacified if more resources had been made available to the program.[7] I sort favelas based on the actual or planned implementation dates, and find that the growth in the number of firms (calculated from the planned/actual implementation date to two years after) is abruptly lower above the cancellation cutoff, using a regression discontinuity design.[8] Finally, I provide evidence that the increase in the number of firms is not a mechanical consequence of formalization by showing that the implementation of Pacifying Police Units has real impact on the favela's average household income using Brazilian Census data.

To understand the consistency of my results with the equilibrium shift hypothesis, I build a continuous time-overlapping generations model featuring business activity and crime. Motivated by the findings of the previous literature, I assume that a given generation takes into account the possible gains from formal employment or self-employment before deciding whether they will become criminals or formal workers (Cornwell and Trumbull 1994; Raphael and Winter-Ember 2001; Gould et al. 2002). On the other hand, entrepreneurs' profits are reduced by property crimes, which shrink the gains from labor markets and self-employment. I characterize the long-run equilibria set of the model, showing that the feedback between business profitability and crime rates can lead to multiple equilibria,[9] depending on how large the property crime gains are when compared to the probability of imprisonment. If crime returns are large, then the poverty trap equilibrium with large crime rates and a small number of firms is the single equilibrium. On the other hand, if crime returns are low, then there is a single prosperous equilibrium with no crime and many firms. Finally, if crime returns are moderate, then three equilibria are possible: the poverty trap equilibrium, the prosperous equilibrium, and an intermediate equilibrium with a moderate level of both crime and business activity. Next, I characterize the effect of a temporary law enforcement intervention, inspired by the Pacifying Police Unit program. I show that if a given economy starts in the poverty trap, then a temporary intervention can shift the economy to the prosperous equilibrium, provided that there is a sufficiently intense police shock. The results of such an intervention are everlasting: property crimes are permanently reduced, while business activity is permanently increased. This result is consistent with my findings regarding the effects of the Pacifying Police Units, highlighting possible mechanisms that might be coming into play: the effect of crimes on the profitability of local business, and the intertemporal choices of agents between crime, self-employment, and formal employment.

My findings are related to the extensive literature addressing the economic and financial

[7] This list was provided by a past coordinator of the program, although the group of favelas is well known and was widely reported by the media at the time.

[8] I do not use the regression discontinuity as my main result for two reasons. First, it is applied to the last Pacifying Police Units implemented in 2014, which reduces the possibility of analyzing the long-run effects. Second, the regression is reasonably underpowered, given the reduced number of units according to RDD standards.

[9] Other papers presented models of multiple equilibrium featuring a similar interaction between crime and labor markets. See Mehlum, Moene, and Torvik (2005) and Mauro and Carmeci (2007). The novelty of my model is the result that a temporary law enforcement intervention can lead to an equilibrium shift.

effects of crime and violence. Several works on this topic focus on the effects of corruption (Shleifer and Vishny 1993), corporate crimes (Karpoff, Lee, and Martin 2008; Guiso, Sapienza, and Zingales 2008; Dyck, Morse, and Zingales 2013) and civil wars (Bellows and Miguel 2009; Voors, Nillesen, Verwimp, Bulte, Lensink, and Soest 2012; Besley and Mueller 2012). Studies specifically about urban crime rates are more incipient, focusing mainly on the value of lost human lives (Soares 2006) and on urban flight (Cullen and Levitt 1999). I contribute to this debate by showing that urban crime rates (i.e., violent crime and property crimes) have strong effects on the level of local business activity, both over the short run and long run. Additionally, given that property crimes are violations related to property rights, this paper also adds to the literature about the effects of property rights in a dimension not analyzed by previous works, which focused mainly on agriculture (Besley 1995; Hornbeck 2010) and on the risk of expropriation by governments and elites (Acemoglu, Johnson, and Robinson 2001). This paper deals with the effects of stronger property rights for small entrepreneurs operating on *main street*, where the law is of little value if it is not enforced.

This paper is structured as follows. Section 1.1 quantifies the effect of short-lived police shocks on the sales of the retail sector. Section 1.2 addresses the effect of more persistent – yet temporary – police shocks on the expansion of local business. Section 1.3 presents the model featuring crime rates and business activity. Section 1.4 concludes.

1.1 Effect of Short-Lived Police Shocks

Is this section, I analyze how variations in crime – induced by shocks that reduce police activity – affect business activity, as measured by the total sales of the retail sector. Crime rates could affect retail sales for several reasons. Consumers might take into account the probability of crime victimization when deciding on places to shop. High crime rates might also shrink the retail sector's working hours, which, in turn, reduces sales to passerby customers. I measure the effect of a temporary variation in crime rates on the retail sector's sales by exploring the variations in crime induced by 30 law enforcement strikes in Brazil.

Institutional Setting

In Brazil, crime prevention and maintenance of the public order is the responsibility of the state Military Police. Each of the 27 Brazilian states has its own independent Military Police,[10] which are operated and financed by the respective Executive Branch. The poor fiscal condition of most Brazilian states since the 1980s has led to frequent and lengthy wage negotiations between the Military Police unions and the state governors, resulting in illegal police strikes on several occasions.[11]

[10] More precisely, 26 states and one federal district.

[11] The Brazilian Constitution states that Military Police strikes are illegal. As such, the strikes are performed through some action that prevent police from working, such as family members locking the garage gates.

The Esprito Santo State Military Police Strike

A dramatic example of a Military Police strike occurred in the Brazilian state of Esprito Santo in 2017. The police union demanded a wage correction accounting for cumulative inflation since the previous adjustment in 2010, which was refused by the then governor Paulo Hartung. After several months of failed negotiations, the strike started on February 2, with the participation of 75% of the police force in the entire state[12]. Hartung reacted by claiming that he would not "give in to blackmail against the citizens of Esprito Santo." As a result, the strike was contained on February 25, when the police force decided to return to work because it was faced with the threat of prosecution.

Esprito Santo was filled with the "disturbing silence of the deserted streets"[13] during the 23 days of the strike: the increased violence in urban areas greatly reduced the flow of passersby, according to Brazilian media. The number of monthly homicides in the state jumped from approximately 100 to 229 in the month of the strike (see Figure 1.1(A)). There are no available data on robbery rates, but the narrative of the local media suggests a similar spike in property crimes during the strike. As a result, the sales of the retail sector fell 14%, the largest drop ever registered in this series for the Esprito Santo state (see Figure 1.1(B)). Surprisingly, the sales in the months following the strike reverted back to the January level, which was not large enough to compensate for the loss in consumption in February. Consequently, total 2017 sales in the Esprito Santo state were reduced.

Effect of State-Level Police Strikes

To make sure that this pattern is not specific to the Esprito Santo state strike, I hand-collected data on a sample of 30 state-level strikes between 2000 and 2017 through searches in the main Brazilian newspapers and through information requests to the Secretariat for Public Security of each state.

Figure 1.2 shows that the strikes are geographically dispersed, with a mild concentration in the northeast of the country. There were strikes in the richest state of the country (Distrito Federal, in the Midwest) and in the poorest state of the country (Maranho, in the Northeast), which illustrates the diversity of my sample.

For each strike, I collected the starting date, the ending date, and the percentage of the police force participating in the strike. Most strikes are clustered around the years 2010-2015, which can be explained by the low economic activity in that period (see Figure 1.3(A)). Strikes are usually short-lived, given that authorities tend to react quickly by offering better wages or by arresting the union leaders. The average strike duration is 5 days or 17% of the days of the strike month (see Figure 1.3(B)). Strike participation (as a percentage of the total police force) was on average 58%, and there were several cases of almost full (100%) participation (see Figure 1.3(C)). These data indicate that unions usually decide to

[12]According to *O Globo*, one of the main Brazilian newspapers, there were "only 500 Military Police officers on the streets across the state. Normally, 2,000 officers would be patrolling".

[13]*Revista poca* 02/11/2017, *A greve da polcia no Esprito Santo passou do ponto.*

undertake a strike when their goal reflects the opinion of the majority of their members. For the sake of parsimony, I summarize the potential impact of a strike in a given state i and in a given month t with a single variable, which I call $StrikeIntensity_{it}$, defined as the product of the percentage of participation in the strike and the percentage of the month under strike:

$$StrikeIntensity_{it} = \left(\frac{\text{\# Police Participation}_{it}}{\text{\# All Police}_{it}} \right) \times \left(\frac{\text{\# Days of the Strike}_{it}}{\text{\# Days in the Month}_{it}} \right) \quad (1.1)$$

If a strike in a given state was longer than one calendar month, then I collapse the affected months into a single time observation t. This definition implies that $StrikeIntensity_{it}$ is equal to 1 if and only if the strike has full participation and a duration of one entire month. It is clear from the histogram 1.3(D) that two strikes are outliers, with values for $StrikeIntensity_{it}$ between 0.4 and 0.5. The next results do not depend on these extreme strikes.

I assess the effect of a given strike, summarized by the variable $StrikeIntensity_{it}$, on the log number of homicides and on the log of total retail sector sales[14]. Ideally, I would also use a measure of property crime as outcome variable, but there are no available state-month robbery or theft data in Brazil. Faced with this limitation, I use the homicide rate as my proxy for crime, keeping in mind the caveat that this is not the best measure of crimes affecting business.

The monthly number of homicides in each state comes from the Mortality Information System, a national database of obituaries managed by the Federal Health Department.[15] This database covers only the period up to 2016, so I filled in the missing values with data from the Security Secretariat of each state when possible. The state-level monthly sales comes from the Brazilian Institute of Geography and Statistics. All monthly variables were seasonally adjusted by a regression with month dummies. According to Table 1.1, the average Brazilian state has approximately 163 murders every month, with a strong cross-sectional variation that indicates the heterogeneous distribution of crime across the country. The log variation of monthly retail sales has an average of 0.003 and a standard deviation of 0.03, so the 14% drop in sales during the Esprito Santo strike is highly atypical and cannot be explained by the standard noise in the data.

To measure the effect of strikes on murders and retail sales, I use a generalized difference-in-difference model with monthly data and a continuous treatment variable (rather than binary) that exploits the $StrikeIntensity_{it}$ variable. My event study for the outcome variable $y \in \{\log(homicides), \log(sales)\}$ is as follows, where i is the state and t is the calendar month:

$$y_{it} = \gamma_i + \theta_t + \sum_{\substack{-6 \leq k \leq 6 \\ k \neq -1}} \beta_k \times 1(k \text{ months after the strike } s) \times StrikeIntensity_{it} + \varepsilon_{it} \quad (1.2)$$

[14] I do not use services in my analysis because the monthly survey of services (conducted by the Brazilian Institute of Geography and Statistics) is too short. Industrial production is also not included because it is not likely to react to such a short term shock.

[15] More precisely, I use the number of "deaths caused by aggression" in the Mortality Information System. Ideally, monthly data on robbery would be included, but there are no robbery data available at the month-state level in Brazil.

The coefficients measuring the treatment effect are β_k for $k \geq 0$. Underlying this model is the identification assumption that the estimated coefficients $\{\hat{\beta}_k \mid k \geq 0\}$ would be zero in the absence of a strike, which is motivated by two facts. First, strikes are the result of months of negotiation, so it is unlikely that unions are timing the strikes to match periods of high violence and low sales, which could result in reverse casualty. Second, police strikes are illegal in Brazil, so the police force usually stay inside the police stations during the strike period to avoid criminal prosecution. Hence, street manifestations that could also indirectly affect the retail sector – thus potentially creating a problem of omitted variable bias in the estimation of the model (1.2) – are unheard of. To rule out the possibility that the strikes have indirect effects on sales and murders by channels other than the shock in law enforcement, I look to the effects of placebo strikes of the largest group of state workers: state teachers. As displayed in Table 1.3, teachers' strikes do not impact crime or the retail sector.

The impact of law enforcement strikes, displayed in Figure 1.4, are very similar to what happened during the Esprito Santo strike. A 10% increase in $StrikeIntensity_{it}$ is followed by a sudden increase of 11.7 percentage points in the number of homicides. However, the number of homicides reverts to the pretreatment level one month after the strike (see Figure 1.4(A)). Unfortunately, there are no available data to measure how this directly affected the behavior of shoppers and retailers, but it is natural to assume that this spike in crime could have reduced both the flow of buyers in the streets and the working hours of the retail sector. As a result, the retail sector's sales suddenly fall by 1.4% in the month of the strike, partially reverting to the pre-strike level after that. Interestingly, sales after the strike do not compensate for the period of reduced purchases, indicating a real reduction on total trade rather than a mere intertemporal reallocation.

Next, I use the previous estimates to calculate the total economic losses caused by the 30 police strikes in my sample. First, I calculate the number of additional murders as the product between the coefficient estimates in the event study (1.2) and the pretreatment number of murders in the same state. Then, I estimate the losses caused by the additional murders by multiplying the additional number of murders with the Value of Statistical Life (VSL) in Brazil estimated by the Special Secretariat for Strategic Affairs of Brazil (2018)[16]. In mathematical terms, the losses due to additional murders are

$$LossesMurders = \sum_{s=1}^{30} VSL \times (e^{\beta_0} - 1) \times (PreTreatment\#Murders_s) \tag{1.3}$$

Similarly, the losses caused by the reduction in sales are calculated as the product of the event study coefficient and the pre-treatment amount of sales:

$$LossesSales = \sum_{s=1}^{30} (e^{\beta_0} - 1) \times (PreTreatment\$Sales_s) \tag{1.4}$$

[16]The VSL is an estimate of the amount of money the public wants to spend to reduce the loss of one life.

In terms of the number of deaths, the strikes resulted in 1,229 additional murders – which is equivalent to 27% of the total number of murders in Brazil in a typical month. Using the estimate of the Statistical Value of Life from Special Secretariat for Strategic Affairs of Brazil (2018), this amounts to an economic loss of 676 million reais (301 million dollars). The total loss of retail revenues is 3.13 billion of reais (approximately 1.40 billion of dollars, see Table 1.2). A better measure of economic loss can be obtained by removing the cost of the goods sold, from which we obtain a loss 634 million reais (or 282 million dollars) in terms of the value added by the retail sector. How much would it cost to hire substitute police to avoid such losses, assuming it would be possible in such a short period of time? Using the average police wages *after* the strike as a reference, this would cost approximately 151 million reais (67 million dollars). Therefore, the net benefit of hiring substitute police would be (taking into account the value added by the retail sector and the losses due to murders) approximately 1.16 billion reais (515 million dollars) for only 30 strikes, accounting for 157 days.

1.2 Effect of a More Persistent (Yet Temporary) Police Shock

Law enforcement strikes lead to sharp but temporary changes in crime and business activity. It is only natural that such consequences are not persistent: criminals might take advantage of the strike by intensifying their activities, and non-criminals might opportunistically commit crimes, but one would not expect changes in long-term choices in response to such a short event. In this section, I explore a more persistent – yet temporary – police shock, taking advantage of the introduction of a law enforcement program called the Pacifying Police Unit.

Pacifying Police Unit

The Pacifying Police Unit is a law enforcement program implemented in the Brazilian city of Rio de Janeiro. This program aimed to reclaim low-income territories known as favelas from the control of armed drug-dealing gangs through two steps: repression and pacification. During the repression period, usually lasting approximately one month, special police forces would invade the targeted favelas by surprise, arresting the gang members and confiscating guns. After that, a police unit would be installed nearby the group of selected favelas, with 1 to 3 police officers for each 100 people living in the pacified region.

The first Pacifying Police Unit was implemented at the end of 2008. The public reception of the program was, initially, very positive. Favelas that were out of reach of most citizens became tourist attractions, and the Pacifying Police Unit was studied abroad as a successful case of proximity police patrolling. However, an unexpected fiscal crisis faced by the Rio de Janeiro state in 2015 drastically reduced the resources available to the program. As reported in Figure 1.5(A), the total tax revenues of the Rio de Janeiro state fell by 36% in 2015 and another 29% in 2016. Naturally, this reduction translated into a reduction of the overall police budget (Figure 1.5(B)). The reduction in the police operating budget (which includes the maintenance of police stations, guns, ammunition, and cars) was dramatic, from 594 million reais in 2014 to 257 million reais in 2015, a 57% reduction (Figure 1.5(C)). The number of police officers allocated to the program also fell considerably: Figure 1.6 plots the percentage reduction in the number of officers (per 100 people) across different police units between the inception date and July 2018 (the only dates when these data are available) and shows that virtually all units lost a large percentage of their police force: on average, police stations lost 27% of their officers. Without the means to operate, the ability of the police units to prevent crimes in the pacified favelas was greatly compromised. These data and implications motivated my use of this program as a persistent yet temporary police shock in the pacified favelas.

The number of units grew consistently after 2008, when the first unit was implemented, reaching a total of 38 when the last unit was implemented in 2014 (Figure 1.7), accounting for 153 favelas covered under the program umbrella, or 20% of the facelas in the Rio de Janeiro city. According to a past coordinator of the program, several of the units planned for 2015 had to be canceled because of the lack of resources at the time.

Figure 1.8 shows the locations of favelas in the city of Rio de Janeiro, identifying all favelas with grey filling and the pacified favelas with blue borders. There is a clear concentration of pacified favelas in the southeast of the city, which is a consequence of the concentration of the two main drug dealing gangs – Red Command and Friend of Friends – in that region. According to a past coordinator of the program, this did not occur by coincidence: the main objective of the program was to take all the territories of these gangs, given that they were known to be killers of police officers. This plan was partially frustrated by the 2015 fiscal crisis, which implied that many favelas ruled by these gangs were not pacified. This can be clearly seen in the Figure 1.9: of 150 Red Command favelas, 102 were pacified, and of 50 Friend of Friends favelas, 18 were pacified. This data suggest that there is a natural control group that can be used to measure the effect of pacification on business activity: the nonpacified favelas controlled by the two main drug dealing gangs. Hereafter, I refer to these favelas as would-have-been pacified favelas. The following empirical analysis to measure the effect of pacification in entrepreneurship is based on a comparison of the number of firms operating in pacified and in would-have-been pacified favelas around the pacification year. To measure the effect of the police units in crime, I compare crime rates in police districts receiving a first unit and in districts without any unit around the pacification year.

Data

My analysis builds on several databases with information about crime, local business, and demographics between 2003 and 2017.

Annual data on crime come from the Public Security Institute of the Rio de Janeiro state, which discloses aggregated statistics at the level of the 37 police districts in the Rio de Janeiro city. Police district borders are represented by the dark lines on the Figure 1.8. The average homicide rate across police districts, approximately 50 per 100,000 people according Table 1.4, is comparable to the most violent countries of the world and was one of the motivating factors for the creation of the Pacifying Police Unit. Property crimes are also endemic: there were 3,601 annual robberies for each 100,000 people. Approximately 8.4% of these robberies targeted retail stores, banks, and trucks, indicating a potentially detrimental effect on local business.

Data on the firms that ever operated in the Rio de Janeiro city between 2003 and 2017 come from the Department of Federal Revenue, the Brazilian tax authority. This data includes the addresses, operating years, and taxpayer numbers (CNPJ). Firms addresses are matched to favelas through the National Address Book, which includes a list of all the addresses of each favela and is maintained by the Brazilian Institute of Geography and Statistics. My resulting sample consists of 59,381 firms that ever operated in one of the 763 favelas in the city of Rio de Janeiro between 2003 and 2017. As reported by Figure 1.10, there is a large concentration of firms in the retail sector but a reasonable dispersion in other sectors, such as manufacturers, hotels, restaurants, and other service providers. According to Table 1.4, the average favela has 35 firms per 1,000 adults, which is slightly higher than the Latin American average (see Klapper, Amit, and Guilln 2010). However, it is important to note that this is the average between 2003 and 2017, so it partially includes the benefit from the Pacifying Police Unit Program.

Demographic and household data at the favela level come from the Brazilian Census of 2000 and 2010, which is collected by the Brazilian Institute of Geography and Statistics. The average favela was inhabited by 1,565 people in 2000 and by 1,827 people in 2010, which matches the order of magnitude of a census tract in the United States. Average monthly household income in 2000 was R$323 (approximately $162). Strikingly, more than 10% of the households did not have any source of income in 2000, and approximately 10% of the adults could not read or write, while 9% of the houses were not equipped with running water. These numbers speak to the state of semi-urbanization in the favelas.

Baseline Comparison

In my empirical strategy, I estimate the impact of pacification on business by comparing pacified and would-be pacified favelas, while the effect on crime is estimated by comparing pacified and non-pacified police districts[17]. Underlying this methodology is the assumption of parallel trends in outcome variables in the treatment and control regions. I provide evidence of the plausibility of this assumption in two different ways. First, in this subsection, I show that the treatment and control regions are indistinguishable in terms of the observable variables at the baseline. In this case, shocks that load on these baseline characteristics should affect the treatment and control regions similarly. Second, in the next subsections,

[17] I say that a police district is pacified if it received at least one Pacifying Police Unit. The borders of the police districts can be found in the figure map 1.8.

I show that the trends of the outcome variables for the treatment and control regions are statistically indistinguishable before the Pacifying Police Unit program started.

Table 1.5 compares the averages of a set of variables calculated for the treatment and control regions at the baseline. The third column reports the differences, while the last column reports the p-values for the null that this difference is zero. For each variable, I use the latest data available before 2008 (when the first Pacifying Police Unit was implemented), which is 2000 for Census data and 2007 for the other variables. According to panel A, the pacified and non-pacified districts are statistically indistinguishable in terms of homicide rates (61 in the pacified districts, 49 in the non-pacified districts, p-value of 0.17) and robbery rates (1,257 in the pacified districts, 1,503 in the non-pacified districts, p-value of of 0.66). Panel B reports the results for the same tests for the favela level variables, telling a similar story: business density (defined as the number of firms per 1,000 adults), formal workers per firm, household income, and other characteristics are indistinguishable between the pacified and would-have-been pacified favelas at the baseline.

Impact of Pacification on the Crime Rate

My estimates for the causal effect of the police units on the crime rate are based on a generalized difference-in-differences model. If i denotes the police districts, t denotes the year, and $y_{it} \in \{\log(1+\text{homicides}_{it}), \log(1+\text{robberies}_{it})\}$ is the crime outcome variable in year t of the district i, then the regression model estimated with both pacified and non-pacified districts is

$$y_{it} = \gamma_i + \theta_t + \sum_{\substack{-5 \le k \le 7 \\ k \ne -1}} \beta_k \times 1(k \text{ years after police unit in } i) + \varepsilon_{it} \tag{1.5}$$

Figure 1.11(A) plots the event study described in (1.5) for murders. The results show that the number of homicides in the pacified and non-pacified pacified districts move in parallel before the implementation of the police unit. Pacification reduces the number of homicides by 25 percentage points in the year after the implementation, but this reduction is lost in the following years. Figure 1.11(B) shows that the number of robberies in the pacified and would-be pacified regions also move in parallel before the implementation of a police unit. However, the reduction of 57 percentage points does not revert to the pre-pacification level.

Does the reversion in the homicide rate after pacification coincide with the 2015 fiscal crisis? I answer this question by performing a formal test of reversion after the 2015 fiscal crisis using the following regression model:

$$y_{it} = \gamma_i + \theta_t + \beta \times 1(\text{UPP in } i \text{ at } t) + \delta \times 1(\text{UPP in } i \text{ at } t) \times 1(t \ge 2015) + \varepsilon_{it} \tag{1.6}$$

In this model, a reversion during the fiscal crisis would be associated with a positive coefficient δ. The results displayed in Table 1.6 show that there is a reversion in the number of homicides during the fiscal crisis but not in the number of robberies. Hence, the crime prevention effectiveness of the police units was affected by the fiscal crisis, but this did not affect the reduction in the number of robberies.

The difference between the results for violent crimes and those for property crimes implies that the police units changed the incentives to commit robbery for reasons other than police crime prevention. I argue in the next section that this difference can be explained by the expansion of the local business activity in the pacified favelas, given that property crimes are more sensitive to the local economic conditions than violent crimes according to previous research (Raphael and Winter-Ember 2001).

Impact of Pacification on Firm Creation

According to a past coordinator of the Pacifying Police Unit, the program focused mainly in favelas ruled by the two main drug dealing gangs of the city, Red Command and Friends of Friends, with the ultimate objective of taking all their territories.

Because of the 2015 fiscal crisis, this goal could not be reached, and several favelas controlled by these gangs were left outside the umbrella of the program. I use this group of would-have-been pacified favelas as a control group when estimating the impact of pacification on business activity. If i denotes the favela, t denotes the year, and $\log(1 + \#Firms_{igt})$ is the log of the number of firms in the favela i at t, then the regression model – estimated with both pacified and would-have-been pacified favelas – used is:

$$\log(1 + \#Firms_{it}) = \gamma_i + \theta_t + \sum_{\substack{-5 \leq k \leq 7 \\ k \neq -1}} \beta_k \times 1(k \text{ years after police unit in } i) + \varepsilon_{it} \tag{1.7}$$

Figure 1.12 plots the estimates of the event study model (1.7). Similar to what we found for the crime rate, the number of firms in the pacified and would-be pacified favelas moves in parallel before the implementation of the Pacifying Police Unit. Pacification generates a large growth in the number of firms in the first two years: 11 percentage points in the pacification year and more than 12 percentage points in the following year. After that, the cumulative growth in the number of firms stabilizes at approximately 35 percentage points and is statistically significant at the 5% level up to six years after the implementation year.

Is this growth in business activity truly persistent or does it revert with the 2015 fiscal crisis? There is no evidence of reversion in the event study, but in order to make this point clearer, I regress the log number of firms on the treatment dummy plus the treatment dummy interacted with a fiscal crisis dummy, which takes a value of one after 2015. This model is similar to the model used for crime rates:

$$\log(1 + \#Firms_{it}) = \gamma_i + \theta_t + \beta \times 1(\text{UPP in } i \text{ at } t) + \delta \times 1(\text{UPP in } i \text{ at } t) \times 1(t \geq 2015) + \varepsilon_{it} \tag{1.8}$$

In this model, a reversion during the fiscal crisis would be associated with a negative coefficient δ. The results in Table 1.7 confirm that the growth in the number of firms is persistent and does not revert during the fiscal crisis, given that the coefficient of the interaction variable is both economically and statistically nonsignificant.

This result, together with the previous findings about the effect of pacification on the crime rate, is consistent with an equilibrium shift in a world with multiple equilibria for the crime rate and business activity. The initial shock to property crimes brings more business into the region, which helps reduce the number of property crimes. In turn, the reduction in the number of property crimes helps expand business even further. After the new equilibrium is reached, the reduction in the Pacifying Police Units' effectiveness (because of the fiscal crisis) does not move the economy back to the old "poverty trap" equilibrium, which is characterized by high property crime and low business activity. Violent crime reverts, which is consistent with the findings of previous studies that homicides are more sensitive to law enforcement than to local business conditions (Levitt 1997; Raphael and Winter-Ember 2001).

The heterogeneous effects of pacification across different economic sectors is also of interest. Figure 1.13 breaks down the effect of pacification in six economic sectors, plotting the estimates of the 5 year effects in the (log) number of firms, with 90% confidence bands. All sectors benefit from pacification, including manufacturing, with a growth of 20 percentage points over 5 years. The sectors most strongly affected by pacification are the retail sector and hotels/restaurants, both increasing 40 percentage points by five years after the pacification. This result illustrates how property crimes can be detrimental for many sectors. The retail sector is highly affected by direct robbery of stores, while manufacturing losses are mainly due to the robbery of trucks transporting goods to wholesale firms.

Robustness Checks

My result assessing the causal effect of pacification on the number of firms might suffer from several endogeneity and interpretation issues. First, pacified favelas might receive additional investments from the government, creating a confounding factor in the regression (1.7) and potentially biasing my previous estimate. Second, my estimate could be interpreted as a result of firms moving towards the pacified favelas and not as a real growth in the stock of firms. Third, the pacification of a given favela might create adverse effects in neighboring regions. Fourth, comparing pacified and would-be pacified favelas does not solve all possible endogeneity issued, as the sorting of the favelas into the program implementation schedule could depend on unobservables correlated with trends on crime and business activity. I address these concerns in this subsection.

I address the possibility of confounding factors arising from the investment of the government by using an alternative econometric specification, equivalent to the model (1.7) but including one additional term in the left hand side to control for the logarithm of the cumulative state and municipal investment since the year of 2008.[18] Results displayed in the figure 1.14(A) are very similar to the estimate obtained in the previous section. Another possible confounding factor could arise from the electrical renovations in the pacified favelas conducted by a partnership between a private utility company and the state government. I

[18]This data was collected from the annual budget of the city and the state of Rio de Janeiro.

address this issue by estimating the model (1.7) excluding all favelas that received renovations as a consequence of this partnership from my sample.[19] Results displayed in the figure 1.14(B) are consistent with my baseline estimates.

As in several other studies with causal regressions based on geographical units, I am not able to rule out the possibility of a null aggregate effect. In my design, this could happen if pacification induces firm shifting but not firm creation. In addition, pacification could create negative spillovers to non-pacified favelas if pacification induces a flow of criminals to nearby regions, adversely affecting the business activity there. I address these two possibilities – firm shifting and negative spillovers – by analyzing the impact of pacification on the number of firms in favelas up to 5 miles from the newly pacified favela, applying an event study similar to that of (1.7). The resulting estimates in Figure 1.15 show that the spillover effect on the number of firms is statistically and economically insignificant.

Now, I consider the possibility of endogeneities affecting my main estimates for the effect of pacification on firm creation. The underlying identification strategy used to measure the effect of pacification in business activity is the assumption of parallel trends of the number of firms in the favelas that were pacified before the fiscal crisis and in the favelas that would have been pacified if the fiscal crisis had not happened. This assumption could be violated if, for example, the program implementation schedule depended on unobservable variables correlated with trends on the number of firms. If this is the case, then the creation of firms in the pacified and would-be pacified favelas would be different, even in the absence of any real effect caused by the police units. However, in this case, there should be no significant difference between the *last* pacified favelas and the *first* favelas that would have been treated if the fiscal crisis had not happened. With this intuition in mind, I employ a regression discontinuity that explores the cancellation of the unit with number 39 and so forth. This process is based on a list of six police units (covering 14 favelas) that would have been implemented in 2015, if only the fiscal crisis had not happened. My dependent variable will be the two years growth in the number of firms after the planned pacification date, which is the largest possible horizon, given that my sample spans the period only up to 2017:

$$y_i = \log(\#Firms_i[\text{pacification date} + 2 \text{ years}]) - \log(\#Firms_i[\text{pacification date}]) \quad (1.9)$$

Let $n_i \in \{1, 2, \dots 44\}$ denote the implementation rank of Pacifying Police Unit covering favela i. Then, the regression discontinuity model based on cubic splines is:

$$y_i = a + \beta 1(n_i < 38.5) + \sum_{k=1}^{3} [c_k(n_i - 38.5)^k + d_k(n_i - 38.5)^k \times 1(n_i < 38.5)] + \mathbf{X}_i\gamma + \varepsilon_i \quad (1.10)$$

where $\mathbf{X}_i$ is a set of controls, and 38.5 is the cancellation threshold. Table 1.8 displays the estimates of the causal effect of pacification based on this regression discontinuity model,

[19] According to the annual reports of social and environmental responsibility of the utility firm Light S.A., the favelas receiving renovations are: Santa Marta, Chapu Mangueira, Babilnia, Cidade de Deus, Jardim Batan, Ladeira dos Tabajaras, Morro dos Cabritos, Casa Branca, Borel, Pavão-Pavãozinho, Cantagalo, Morro da Providncia, Formiga, Andara, and Salgueiro.

indicating that pacification increases the number of firms by a significant 36 percentage points over two years. This is larger than the point estimate obtained in the previous section (25 percentage points) but still within the limits of the statistical errors. Figure 1.16 graphically plots the average growth in each Pacifying Police Unit bin after controlling for gang fixed effects and graphically illustrates this result. Clearly, the points on the left side are above the points after the cancellation threshold. The variance in the growth in the number of firms is quite small, so even with a reduced sample, the discontinuity analysis is informative.

Impact of Pacification on Household Income

An important concern is the possibility that the growth in the number of firms following pacification was a consequence of formalization induced by the increased presence of law enforcement officials rather than a real improvement in local business activity. In this section, I provide evidence that pacification created a real transformation of the local economic landscape by looking at how it affected local household income.

Data on the average household income in a given favela come from the Brazilian Census, which is available for only 2000 and 2010. As a consequence, only favelas pacified before August of 2010, when the data for the 2010 Census were collected, will be part of the treatment group in my analysis. More specifically, I estimate the following difference-in-differences model:

$$\log(Income_{2010,i}) - \log(Income_{2000,i}) = a + \beta \times 1(\text{UPP in } i \text{ before Aug 2010}) + X_i^\top b + \varepsilon_i \tag{1.11}$$

Two caveats should be kept in mind regarding the use of this estimation. First, the year 2000 is considerably earlier than the implementation of the first Pacifying Police Unit in 2008, which increases the noise in my measure of change in income around the implementation. Second, given that I observe only two "pictures" of local income, for the data from 2000 and 2010, tests for pretreatment trends are not possible. Nonetheless, there was no evidence of pretreatment trends for any of the variables observed in higher frequency in the previous sections, which makes the identification assumption underlying the difference-in-differences model (1.11) reasonable.

Even if pacification induces more self-employment and causes more firms to hire local labor, then local income is not necessarily affected. Local workers could change jobs because they are motivated by nonfinancial factors, such as time spent on transportation and workload. However, a positive effect on income would likely be a consequence of an improvement in local business conditions. Table 1.9 indicates that this appears to be the case: pacification increases income by approximately 7.2 percentage points. In August 2010, when the Census was collected, the average police unit was 0.7 years old, so the annualized effect is approximately 10 percentage points. For the sake of comparison, the inflation adjusted income in the average favela in my sample *decreased* by 0.9 percentage points in the same period. This gives a clear indication of the strength of the economic transformation induced by the reduction in the local crime rate.

Welfare Analysis

Does the additional law enforcement inducing pacification pay for itself? Thus far, I have shown that the Pacifying Police Units had several tangible benefits by reducing crime, increasing the number of firms, and increasing local income. However, this comes at a cost to the government: estimates of the operational costs, investments and payroll sums up to approximately 1.4 billion reais – or 611 million dollars, using the average currency exchange rate. Does the social benefit justify this expenditure? In this section, I answer this question by estimating the economic value added by the Pacifying Police Unit program.

First, I calculate the economic value generated by the reduction in the number of murders using the value of statistical life estimated by Special Secretariat for Strategic Affairs of Brazil (2018). Next, I calculate the value added of the new firms, summing the estimates of compensation for labor and capital. For this calculation, I make the following conservative assumptions: (i) all workers receive the current minimum wage, (ii) the compensation for capital is equal to the total payroll, and (iii) each firm employs two workers, which is the median number of workers according to a survey of formal and informal firms in Brazil (*Pesquisa de Economia Informal Urbana*).

More precisely, the value added by the reduction in the number of murders is calculated as follows. If i represents the favela, N_f represents the total number of pacified favelas, $\{\beta_t\}_t$ represents the coefficient estimates in the event study for log murders, VSL represents the value of a statistical life in Brazil, and r_t is the real annual interest rate for a 1-year government bond, then:

$$ValueLifes = \sum_{i=1}^{N_f} \sum_{t=StartYear_i+1}^{StartYear_i+3} \frac{VSL \times (e^{\beta_t} - 1) \times (PreTreatment\#Murders_i + 1)}{(1 + r_t)^{t-2008}} \tag{1.12}$$

Note that I include the effect between only years one and three in the sum below, given that the sensitivity of the number of murders is statistically insignificant (at the 10% level) for other horizons.

The value added by the firms created because of pacification is defined in a similar way. The conservative estimate is also calculated for only the horizons under which the coefficients of the event study are statistically significant at the 10% level:

$$ValueFirms = \sum_{i=1}^{N_f} \sum_{t=StartYear_i}^{StartYear_i+6} \frac{4 \times w_t^{min} \times (e^{\beta_t} - 1) \times (PreTreatment\#Firms_i + 1)}{(1 + r_t)^{t-2008}} \tag{1.13}$$

where w_t^{min} is the minimum annual wage at year t and $\{\beta_t\}_t$ represents the coefficient estimates in the event study for the log number of firms. I multiply the total annual payroll by 4 because of the assumption that capital and labor are both measured as the annual compensation of two employees. Additionally, I calculate the sum (1.13) using two different methodologies. In the first methodology, which I refer to as *conservative*, I consider only the terms β_t in the event study that are statistically significant, so the second sum spans from

$t = StartYear_i$ to $t = StartYear_i + 6$. The second methodology, which I refer as *perpetuity*, assumes that the effect of the shock on the number of firms is permanent, so the second sum spans from $t = StartYear_i$ to $t = \infty$, assuming that the wages, interest rates and treatment effects stay constant after the sixth year of implementation of the police unit.

The value added by the 1,133 firms created because of pacification ranges from 86 million dollars (conservative calculation) to 297 million dollars (perpetuity assumption), according to Table 1.10. This number is not far from the value gained because of the reduction in the number of homicides: 1,397 fewer homicides provided gains of 262 million dollars. The fact that gains stemming from entrepreneurship are comparable to the gains from fewer homicides is surprising, given that the reduction in the number of homicides is usually the main objective of law enforcement interventions. The present value of the total cost of the Pacifying Police Program, which includes operational costs, investments and payroll, is estimated to be 333 million dollars. Therefore, the net benefit of the program is estimated as -97 million of dollars under the conservative methodology and at +114 million under the perpetuity assumption. In conclusion, the program pays for itself only if we consider that the increase in the number of firms is persistent, which is consistent with the hypothesis of an equilibrium shift.

1.3 A Model for the Crime Rate and Business Activity

In the previous section, I provided evidence that a temporary law enforcement shock can reduce the number of property crimes and increase business activity permanently, which is consistent with the idea of an equilibrium shift. Before the law enforcement intervention, favelas were in a "poverty trap" equilibrium, in which high property crime rates and low business activity feed each other. The implementation of a police station in a given favela provides an initial shock to the crime rate, attracting more firms, which increases household income and reduces the incentives for crimes even more. During the fiscal crisis, when the effectiveness of the police stations was severely compromised, the new "prosperous" equilibrium had already been reached, so the number of property crimes did not revert, and business activity did not decline back to the pre-treatment levels. In this section, I formalize this argument by building a model featuring the interaction of crime rates and business activity.

Model Description

For each $t > 0$, let L_t and C_t denote the percentage of the population engaged in productive labor and in crime at t, respectively. All agents work or commit crimes, so $L_t + C_t = 1$. Each firm hires a single worker (receiving wage w_t at t) to produce a single unit of a generic good at each t. Without loss of generality, I normalize the price of the output to 1. There are no barriers to entry or exit, implying that the number of firms will adjust to the existing labor supply L_t. I assume that the revenue of each firm is reduced by a crime "tax" C_t. Given the absence of barriers to entry, firms will have zero profits at any time t, and therefore

$$profits_t = \underbrace{1}_{\text{firm revenue}} \times \underbrace{(1 - C_t)}_{\text{crime tax}} - \underbrace{w_t}_{\text{costs}} = 0 \tag{1.14}$$

$$\Rightarrow w_t = 1 - C_t \tag{1.15}$$

At each $t > 0$, a new generation distributed over the continuous interval $[0, 1]$ is born. I assume that a newborn agent $j \in [0, 1]$ faces an imprisonment disutility j if caught committing a crime. When born, he decides whether he will be a formal worker or a criminal. I assume that this decision is permanent. This premise has empirical support: past criminals have poor employment prospects (Freeman 1999). On the other hand, agents joining labor markets acquire human capital, which makes switching to crime less likely.

If the agent born at t decides to engage in crime, two different outcomes could occur. First, he could be arrested at any point between t and $t + 1$. I assume that the probability of not being arrested is a survival function with hazard rate λ_s, which is an exogenous parameter representing the effectiveness of law enforcement at time s. This implies that the probability of imprisonment of a criminal living between t and $t + 1$ is:

$$\pi_t \equiv 1 - e^{-\int_t^{t+1} \lambda_s ds} \tag{1.16}$$

If not arrested, the criminal will receive a payoff R_c in the last period of his lifespan. I assume (for tractability) that the crime payoff R_c is a constant independent of the current business activity.

If the agent born at t decides to join the formal labor market, he would receive the total wage during his lifespan:

$$\bar{w}_t \equiv \int_t^{t+1} w_s ds \tag{1.17}$$

This implies that the agent $j \in [0, 1]$ born at t will engage in crime if and only if the expected payoff of crime is greater than the labor wage, assuming risk neutrality and a discount rate of zero:

$$(1 - \pi_t)R_c - \pi_t j \geq \bar{w}_t \tag{1.18}$$

Solution

Equation (1.18) implies that a newly born agent $j \in [0, 1]$ will engage in crime if and only if

$$j \leq \frac{(1 - \pi_t)R_c - \bar{w}_t}{\pi_t} \tag{1.19}$$

Let c^t denote the percentage of the agents born at t engaging in crime. Then c^t can be easily calculated from the equation (1.19):

$$c^t = \int_0^1 1\left(j \leq \frac{(1 - \pi_t)R_c - \bar{w}_t}{\pi_t}\right) dj = \min\left\{1, \max\left\{0, \frac{(1 - \pi_t)R_c - \bar{w}_t}{\pi_t}\right\}\right\} \tag{1.20}$$

Equation (1.20) shows that the decision of the current generation to engage in crime depends on the current wages. The overall percentage of criminals is the sum of the criminals from different cohorts:

$$C_t = \int_{t-1}^{t} c^s ds = \int_{t-1}^{t} \min\left\{1, \max\left\{0, \frac{(1-\pi_s)R_c - \bar{w}_s}{\pi_s}\right\}\right\} ds \tag{1.21}$$

The equilibrium of the model is jointly determined by equations (1.15) and (1.21). Note that equation (1.15) establishes that crime affects wages via increased firms' costs. Equation (1.21) shows that wages in turn affects crime via individuals' career choices. This raises the possibility of an feedback loop between the crime rate and wages, which can lead to multiple steady state equilibria. This is described in the proposition below.

Proposition 1: Assume that the law enforcement effectiveness parameter $\lambda_t = \lambda$ is constant, and let $\pi = 1 - e^{-\lambda}$. There are three possibilities for the set of long-run equilibria:

1. If the payoff of property crime is small when compared to the imprisonment probability, $R_c \leq \pi/(1-\pi)$, then there is a single long-run equilibrium with no crime, $C_t \to 0$, high wages, and a large number of firms.
2. If the payoff of property crime is large when compared to the imprisonment probability, $R_c \geq 1/(1-\pi)$, then there is a single long-run equilibrium with high crime, $C_t \to 1$, and no firms.
3. If the payoff of property crime is moderate, $\pi/(1-\pi) < R_c < 1/(1-\pi)$, then there are three long-run equilibria: no crime, $C_t \to 0$, high crime, $C_t \to 1$, and moderate crime, $C_t \to R_c - \pi/(1-\pi)$.

See appendix for the proof.

Effect of the Temporary Law Enforcement Shock

In this subsection, I restrict my analysis to the more interesting case of multiple long-run equilibria. As in the case of the Pacifying Police Units, I analyze the effects of a temporary increase in law enforcement in my model. More specifically, λ_t will increase from λ to $\bar{\lambda}$ between t_0 and $t_0 + T$, reverting to λ after that. At t_0, the economy is in the poverty trap equilibrium, so crime rates are maximal ($C_{t_0} = 1$) and there are no firms operating locally. The following proposition characterizes the effect of this intervention.

Proposition 2: Let $\pi = 1 - e^{-\lambda}$. Assume that $\pi/(1-\pi) < R_c < 1/(1-\pi)$ and that $C_t = 1$ for any $t < t_0$. The law enforcement effectiveness parameter λ_t is given by λ if $t \in (-\infty, t_0) \cup (t_0 + T, \infty)$ and by $\bar{\lambda}$ if $t \in [t_0, t_0 + T]$, where T is the *length* of the intervention and $\bar{\lambda}$ is the *intensity* of the intervention. Then, there is a minimum intensity of intervention $\bar{\lambda}_{min}$ such that for any $\bar{\lambda} \geq \bar{\lambda}_{min}$, there is a threshold $\tilde{T}$ for the length of the intervention such that if $T > \tilde{T}$, then the economy will converge to the prosperous equilibrium with no crime, high wages, and a large number of firms.

See the appendix for a formal proof of this proposition. Figure 1.17 illustrates this result. For a selected set of parameters and a fixed intensity $\bar{\lambda}$, a law enforcement intervention with a duration of 0.25 is not sufficient to shift the economy to the prosperous equilibrium, so the percentage of criminals eventually reverts to 1. At the other extreme, when the duration of the intervention is one, not surprisingly, the economy shifts from the poverty trap to the prosperous equilibrium. This happens because any agent born during the intervention would choose not to commit crime. After one period, all old criminals are retired, so wages are maximal and there are no incentives to commit crimes, even after the end of the police intervention. However, as the example of duration of 0.55 illustrates, the prosperous equilibrium can also be reached with an intervention of much smaller length, but with a slower speed of convergence, taking approximately twice as long to approach the region of no crime. In this model, the number of firms is the symmetric opposite of the crime rate, so business activity increases permanently.

This result not only confirms the intuition that the dynamics of business activity and of property crime following the implementation of the Pacifying Police Units are consistent with shifting away from the "poverty trap" equilibrium, but it also highlights the possible role played by firms' costs related to property crimes and by the life-time career choices of the local workers. In the model, new generations react to the police shock by choosing productive labor rather than crime as their professional activity, while the old generations retire, which permanently reduces the crime rate. This increases firm revenues and wages, which reduces the incentives of the new generations to commit crime even further. When the additional police leave the region, the low crime rate is enough to ensure that wages remain high and that incentives for crime are low.

1.4 Conclusion

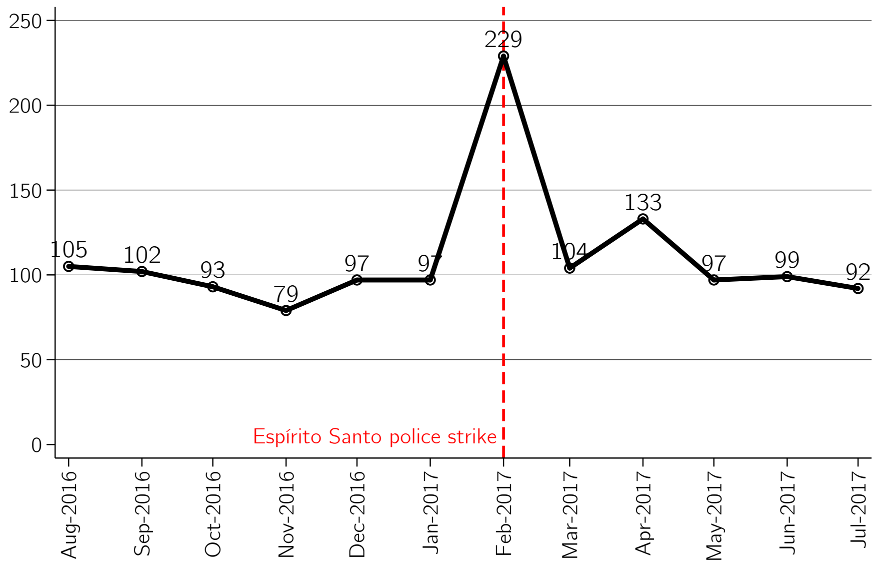
The recent epidemic of crime in several developing countries (especially in Latin America) and in several regions of developed countries (especially in Chicago and England) have renewed the worldwide debate about the optimal spending for law enforcement. The effect of policing on crime in the short run is well understood: Levitt (1997) shows that increasing the size of the police force decreases crime rates. However, the indirect effects on local business activity, as well as the effects over the long run, are not properly addressed by the current literature. Quantifying these benefits is fundamental for the precise measurement of the trade-off in the law enforcement spending decision.

In this paper, I study the impact of crime on business, showing that short-lived increases in crime induced by state-level law enforcement strikes have strong effects on the sales of the retail sector. The loss stemming from additional murders during a strike is similar to the value lost by the retail sector, highlighting the importance of accounting for business activity in the law enforcement cost-benefit analysis.

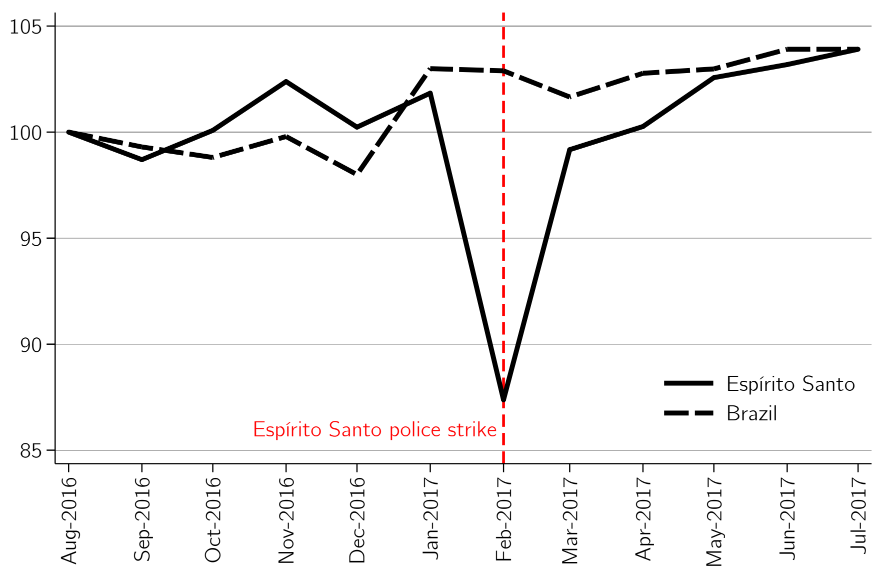
Taken together with the finding of the crime literature that lower business activity leads to more crimes, this implies a feedback loop between crime and business, suggesting the existence of multiple Pareto-ranked equilibria. In this setting, a substantial – yet temporary – law enforcement intervention could shift the economy away from the low-business high-crime undesirable equilibrium. I test the equilibrium shift hypothesis by taking advantage of the introduction of the Pacifying Police Unit program, which aimed at reducing the crime rate in Brazilian favelas. The number of property crimes dropped considerably after the implementation of the police units and stayed at low levels even during the fiscal crisis that dramatically reduced the amount of resources available to the program. The number of firms also increased persistently, suggesting that the initial police intervention induced an equilibrium shift in the pacified favelas. Overall, my results illustrates how law enforcement could possibly be used as a tool to achieve economic development by shifting regions away from the poverty trap.

Figure 1.1: Esprito Santo State Police Strike

(A) Number of Homicides in Esprito Santo

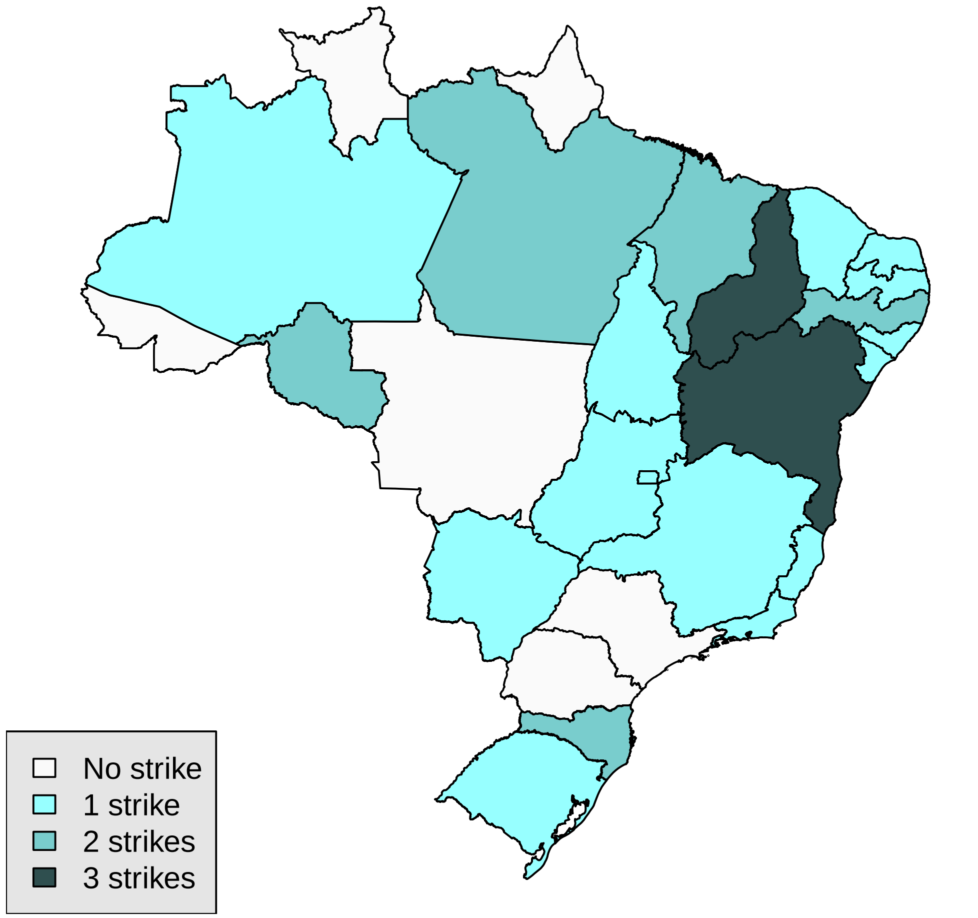


(B) Seasonally Adjusted Retail Sector Sales (August-16 = 100)



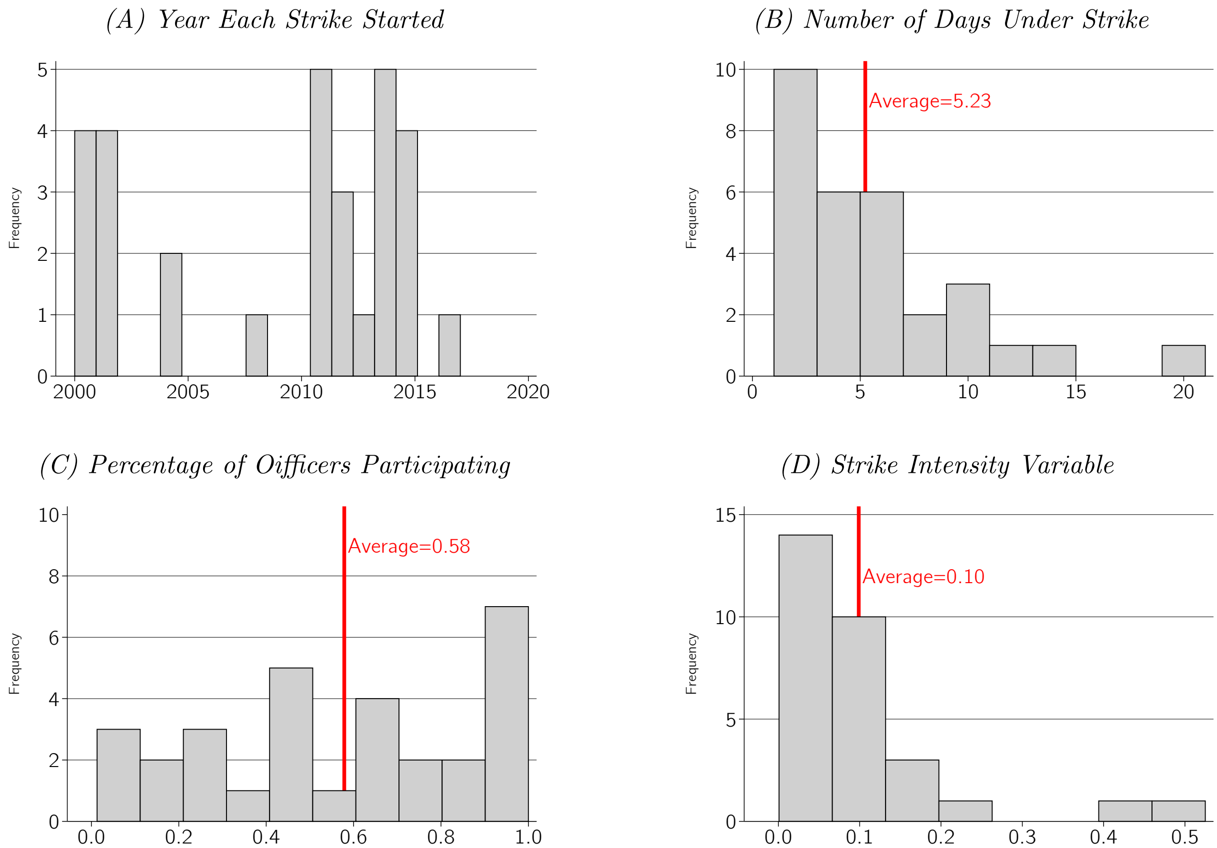
Note: This figure reports the monthly evolution of the number of homicides and on the retail sector sales around the Esprito Santo law enforcement (military police) strike of February 2017. The data on homicides come from the Esprito Santo State Security Secretariat, while data on retail sales come from the Brazilian Institute of Geography and Statistics (IBGE).

Figure 1.2: Law Enforcement Strike Map



Note: In this map, I highlight the Brazilian states where the military police went on strike at some time between 2001 and 2017. Data were collected through manual searches conducted through the main Brazilian newspapers and through information requests to the Secretariat for Public Security and the Military Police unions of each state.

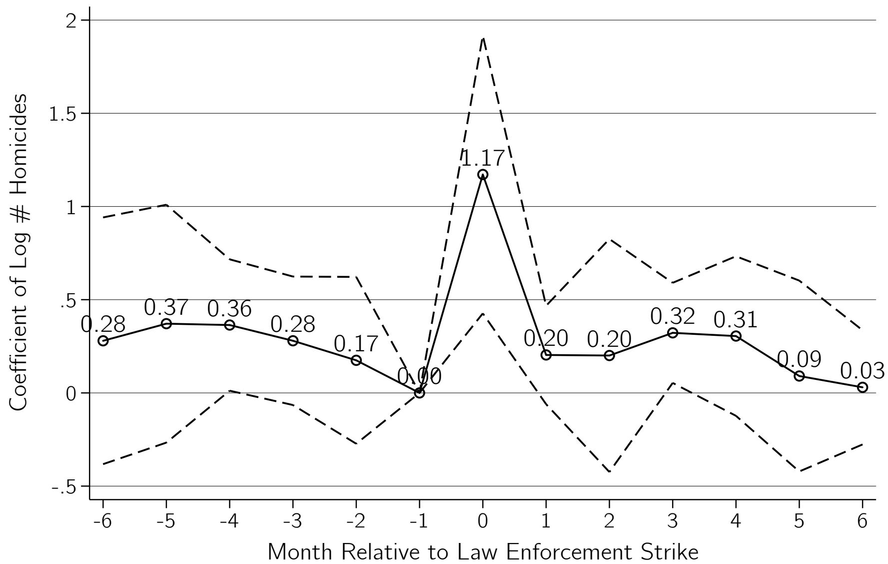
Figure 1.3: Law Enforcement Strikes' Characteristics



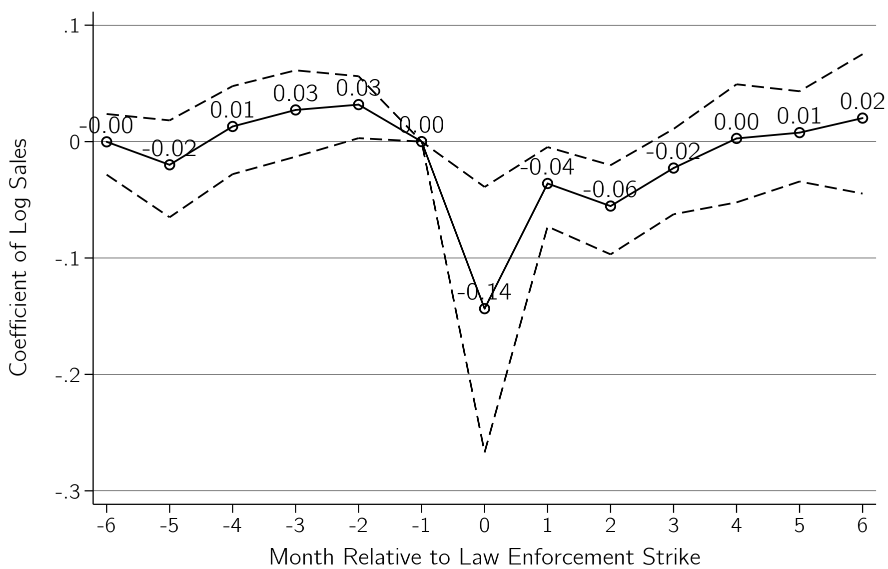
Note: The top left figure is the histogram of the year of each strike in my sample. The top right figure is the histogram of the duration of the strikes, in number of days. The bottom left figure is the histogram of the percentage of the police force that joined the strike. The bottom right figure is the distribution of the $StrikeIntensity_{it}$ variable, which is defined by equation (1.1). Data were collected through manual searches of the main Brazilian newspapers.

Figure 1.4: The Effect of State-Level Law Enforcement Strikes

(A) Log Number of Homicides

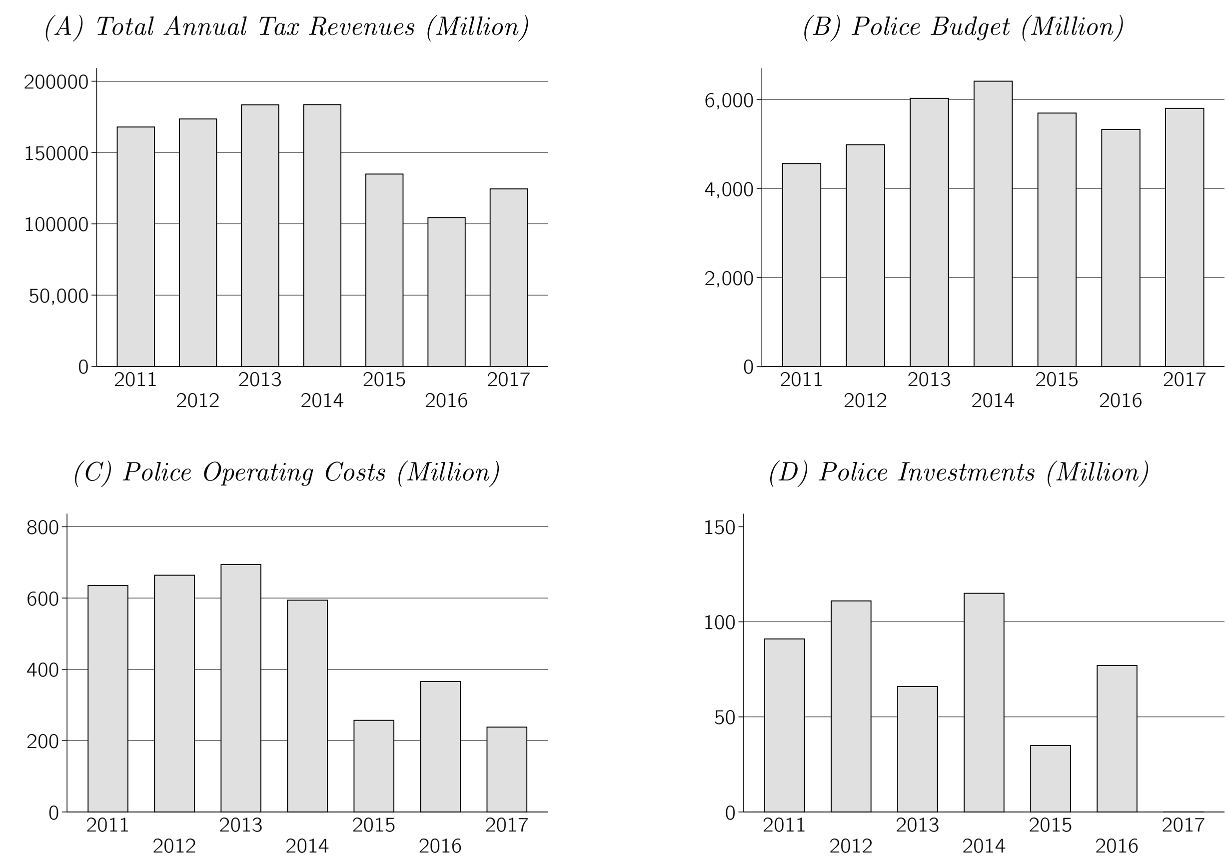


(B) Log Retail Sector Sales



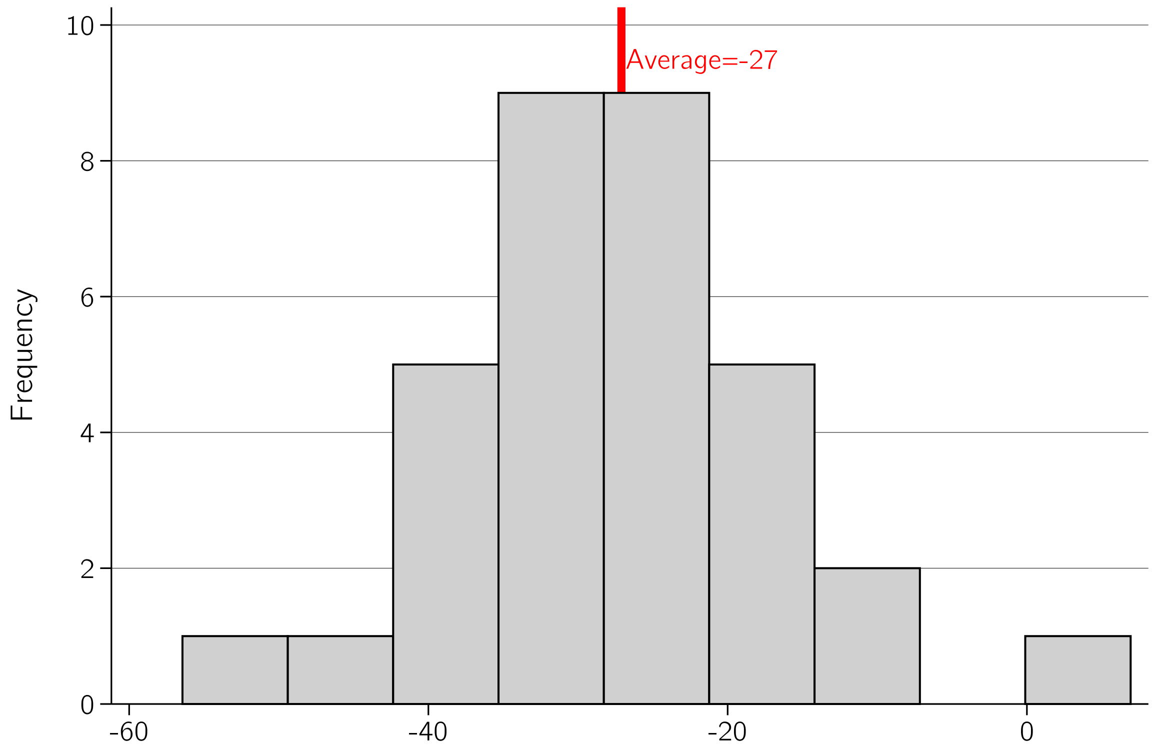
Note: This figure shows the effect of state-level law enforcement strikes on the state-level log homicides and log sales of the retail sector. The y-axis represents the coefficient estimates from a regression of the dependent variable on dummy variables indicating the month of the police strike, controlling for time and state-strike fixed effects. The regression model is described in equation (1.2). The dashed lines represent 95% confidence intervals based on robust standard errors that are clustered at the state level.

Figure 1.5: Rio de Janeiro State Revenues and Police Budget



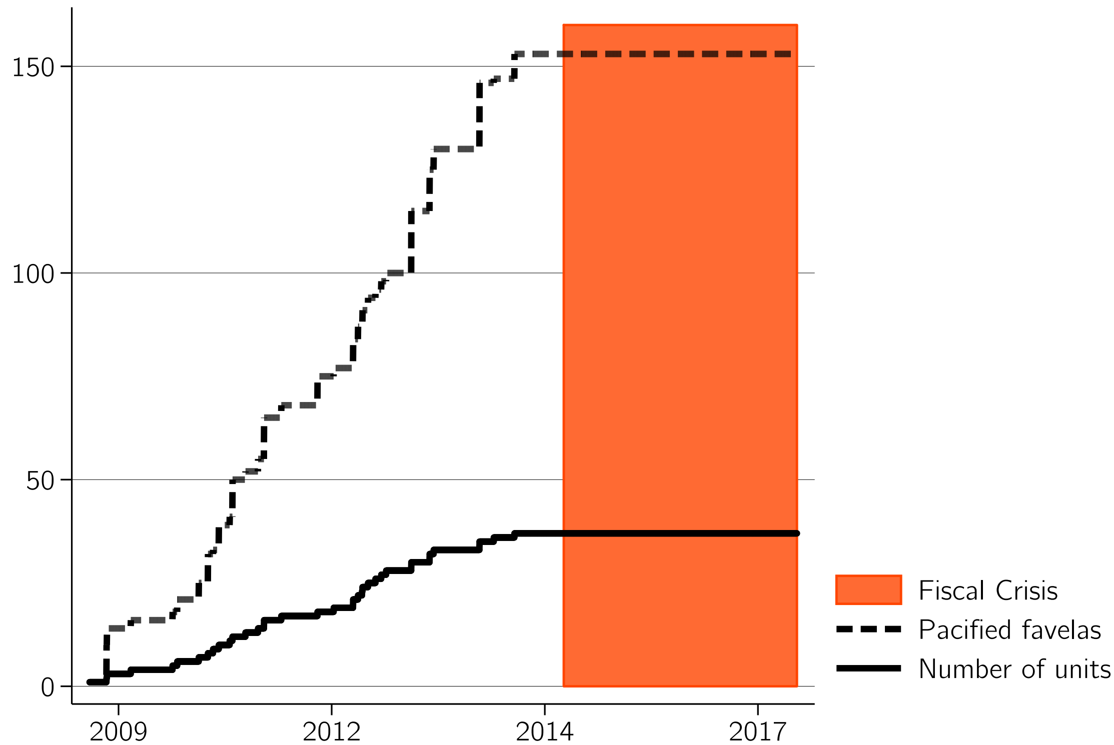
Note: Data come from the Rio de Janeiro State Tax Transparency website. All numbers are converted to December 2017 million reais.

Figure 1.6: % Variation in the Number of Police Officers per 100 People



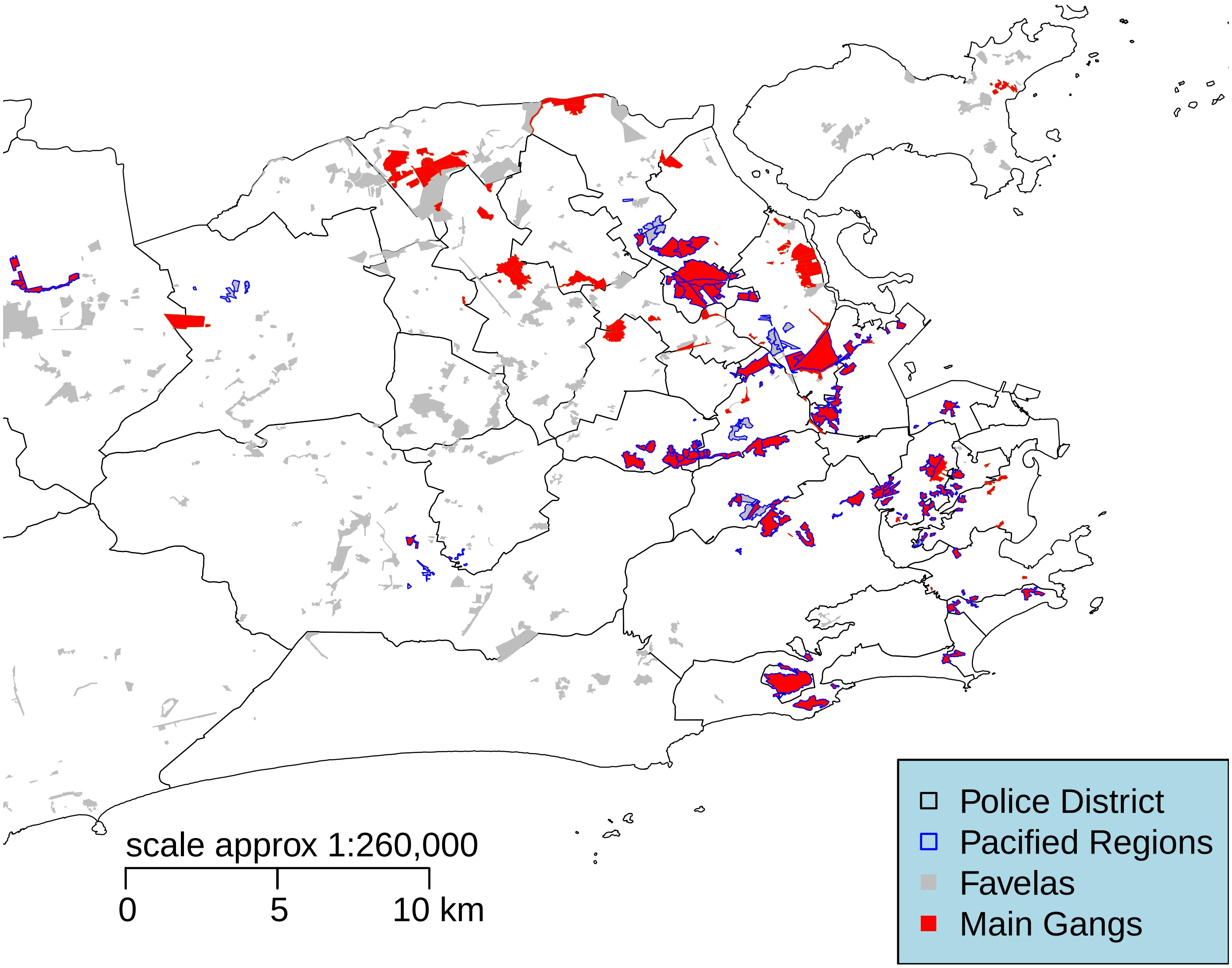
Note: This figure is the histogram of the percentage variation in the number of police officers per 100 people between inception and July 2018 across different Pacifying Police Units. The data come from an information request made by the author to the Security Secretariat of the Rio de Janeiro State.

Figure 1.7: Number of Pacifying Police Units (UPPs)



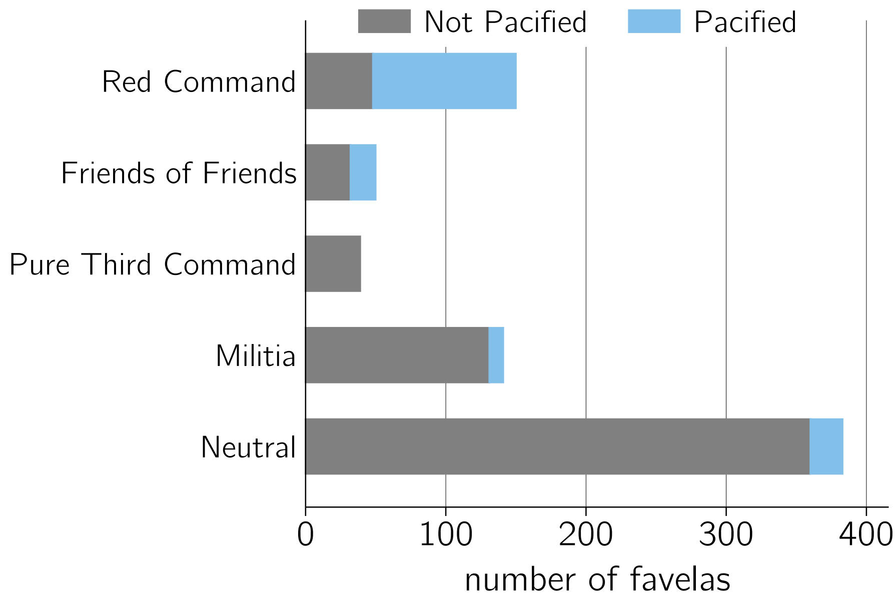
Note: This figure shows the number of pacifying police units (UPPs) in operation between 2008 and 2014.

Figure 1.8: Location of Favelas in the City of Rio de Janeiro



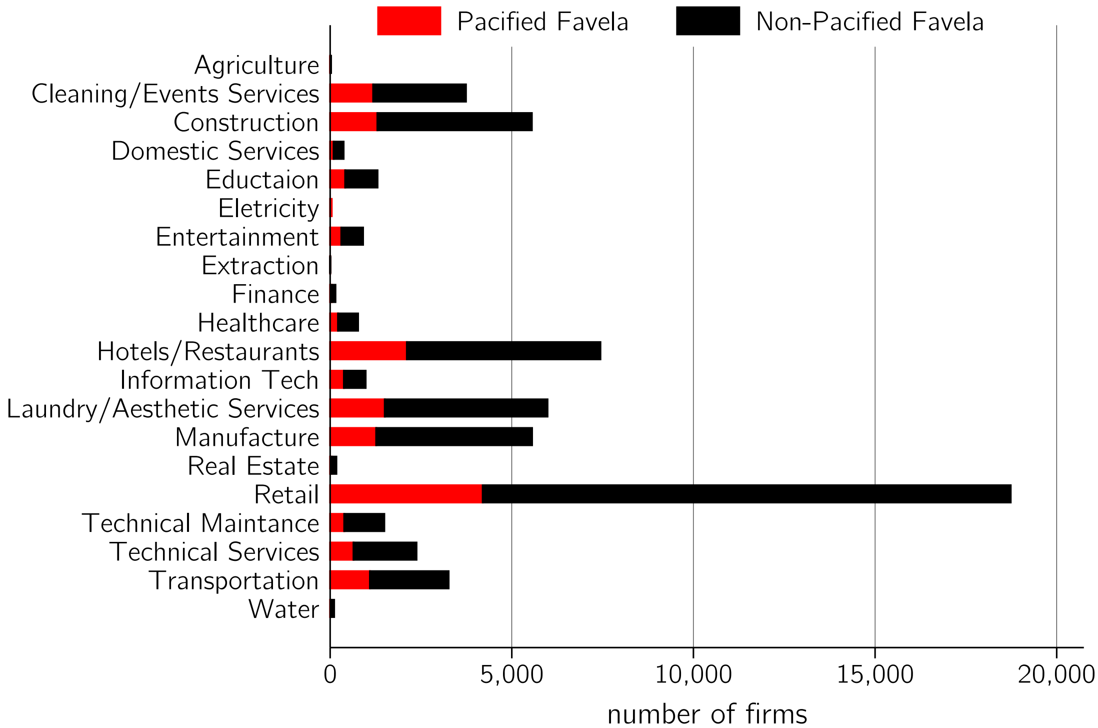
Note: This figure reports the location of the favelas in the city of Rio de Janeiro. The black lines are the borders of the Police Districts for which crime statistics are available. The grey regions are favelas.The blue lines are the favelas selected for the UPP program. THe favelas in red were ruled by either the Friends of Friends or by the Red Command gangs at the baseline.

Figure 1.9: Favelas' Territorial Command



Note: This figure reports the breakdown of the favelas in the city of Rio de Janeiro into five groups of territorial command in 2008. *Red Command*, *Friends of Friends* and *Pure Third Command* are drug-dealing gangs. *Militias* are paramilitary groups. Neutral favelas are not ruled by any specific gang or criminal group. This 2008 classification is based on the favela territorial command census collected by the team of the federal congressman Fernando Gabeira.

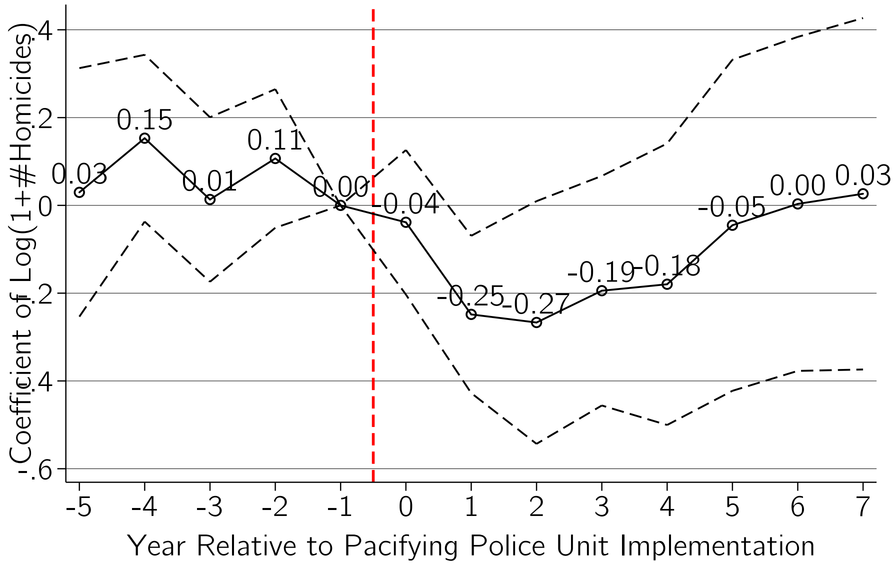
Figure 1.10: Favelas' Firm Sector Breakdown



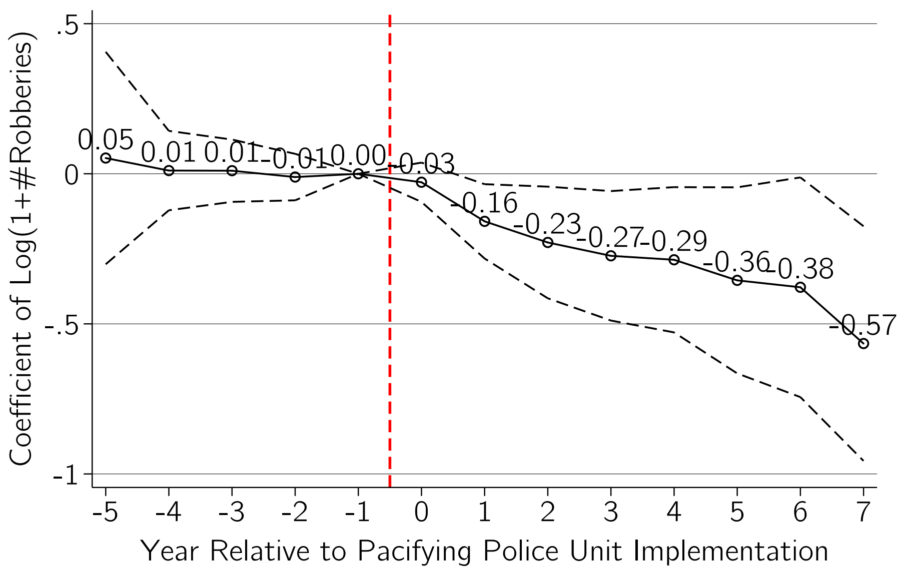
Note: This figure reports the sector breakdown of all formal firms (i.e., those registered with the national tax authority, the Brazilian Federal Revenue) operating in any favela in the period between 2003 and 2017. The red bars represent firms that started operating before the implementation of the Pacifying Police Unit in the corresponding favela, the black bars represent firms that started operating after the implementation of the Pacifying Police Unit in the corresponding favela, and the grey bars represent firms operating in favelas that were never pacified by the UPP program.

Figure 1.11: Impact of Pacification on Crime

(A) Log Number of Homicides

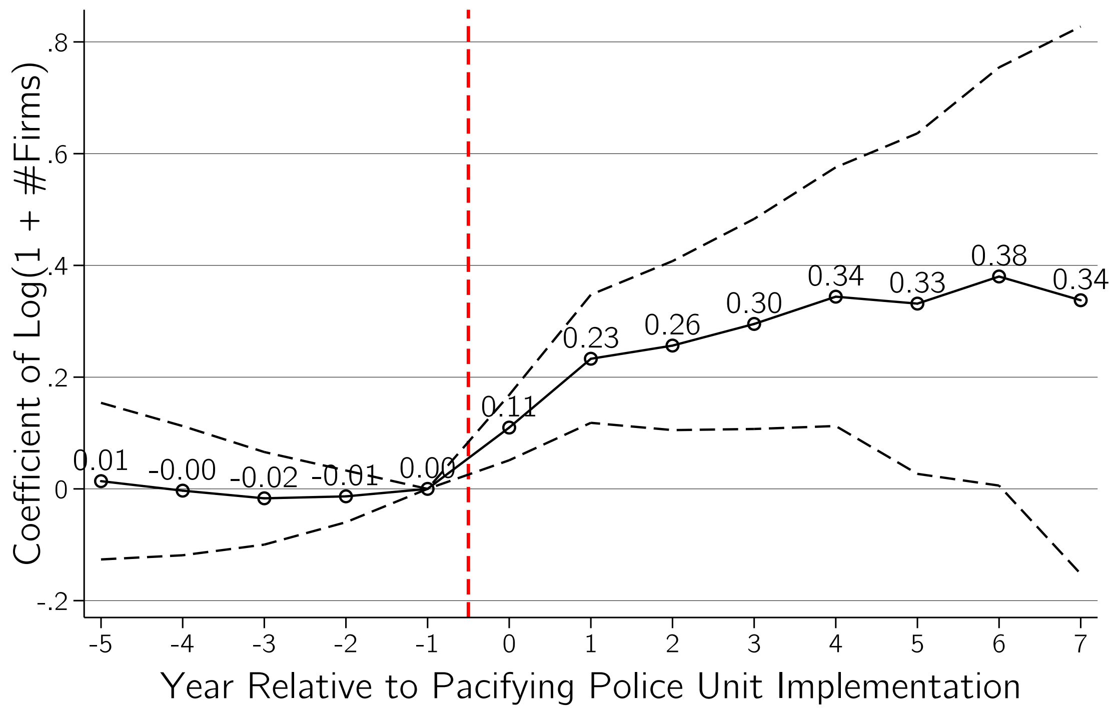


(B) Log Number of Robberies



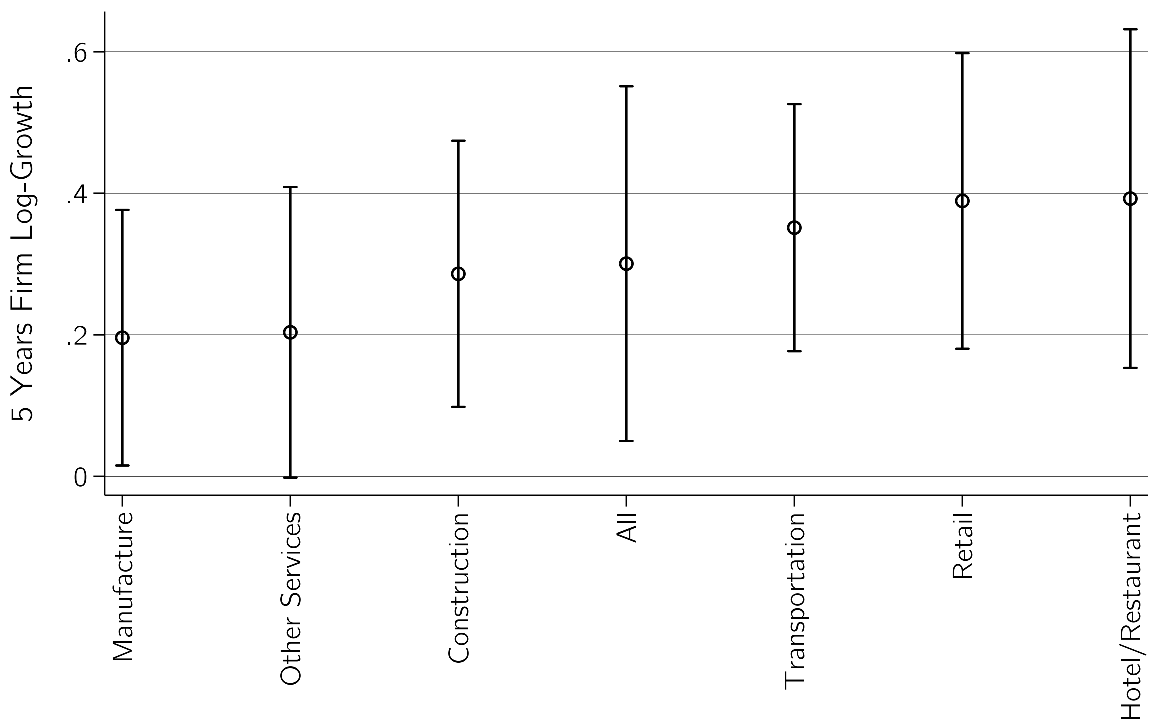
Note: This figure shows the effect of pacification at the police district level and considers the log number of homicides and log number of robberies. The y-axis represents the coefficient estimates for a regression of the dependent variable on dummy variables indicating the year relative to the implementation of the first police unit in a given police district in the city of Rio de Janeiro. The regression model is described in equation (1.5). The dashed lines are 95% confidence intervals based on robust standard errors clustered at the police district level.

Figure 1.12: Effect of Pacifying Police Units on the Number of Firms



Note: This figure shows the effect of pacification in the favela on the number of firms. The y-axis represents the coefficient estimates from a regression of the dependent variable on dummy variables indicating the year relative to the implementation of the police unit in a given favela. The regression model is described in equation (1.7). The dashed lines are 95% confidence intervals based on robust standard errors clustered at favela level.

Figure 1.13: Firm Creation Induced by the UPPs in Each Sector



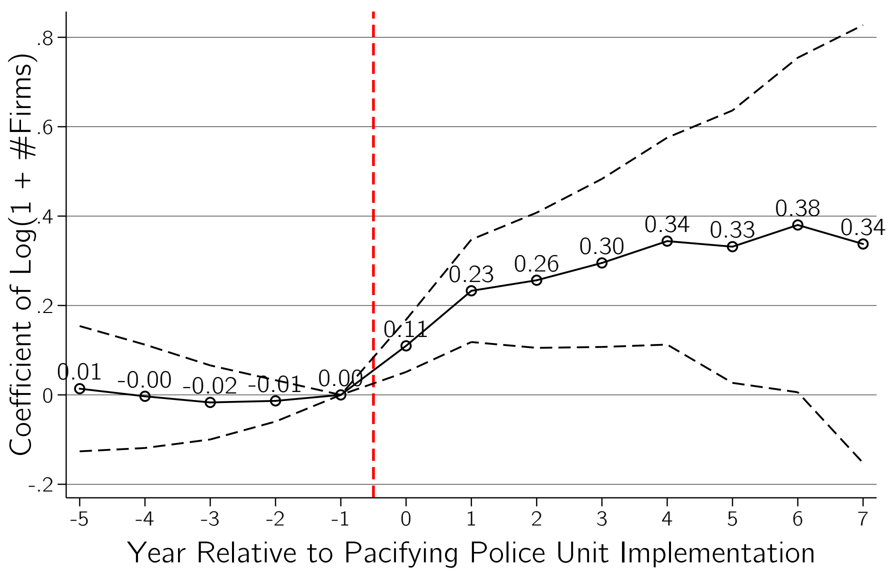
Note: This figure reports the 5-year growth in the log number of firms in each sector caused by the Pacifying Police Units. Is s represents sectors, i represents favelas, and t represents years, then the regression model (with standard errors clustered on 763 favelas) is:

$$\log(1 + \#Firms_{sit}) = \gamma_{si} + \theta_{st} + \sum_{k \neq -1} \beta_{sk} \times 1(k \text{ years after UPP in } i) + \varepsilon_{sit}$$

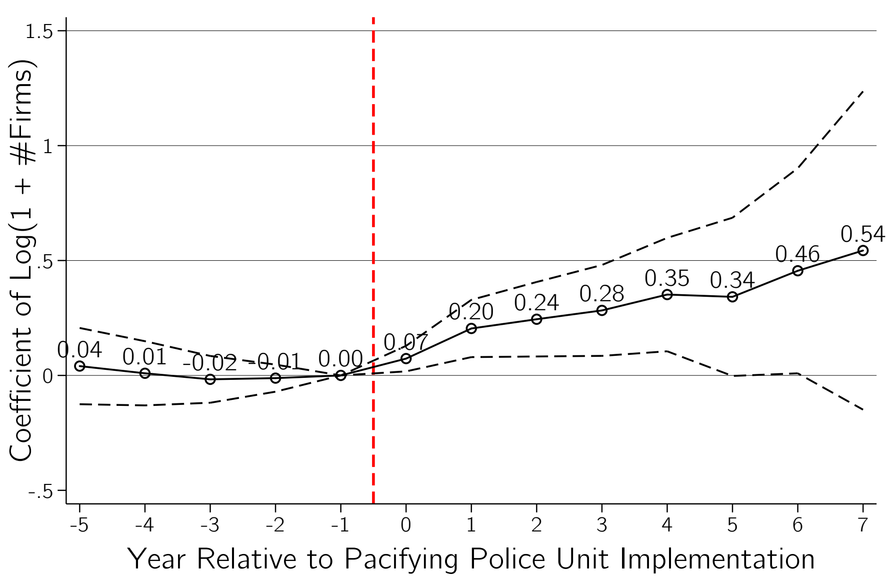
This figure reports the coefficents β_{s5} for each sector s.

Figure 1.14: Alternative Specifications

(A) Controlling for State Investment

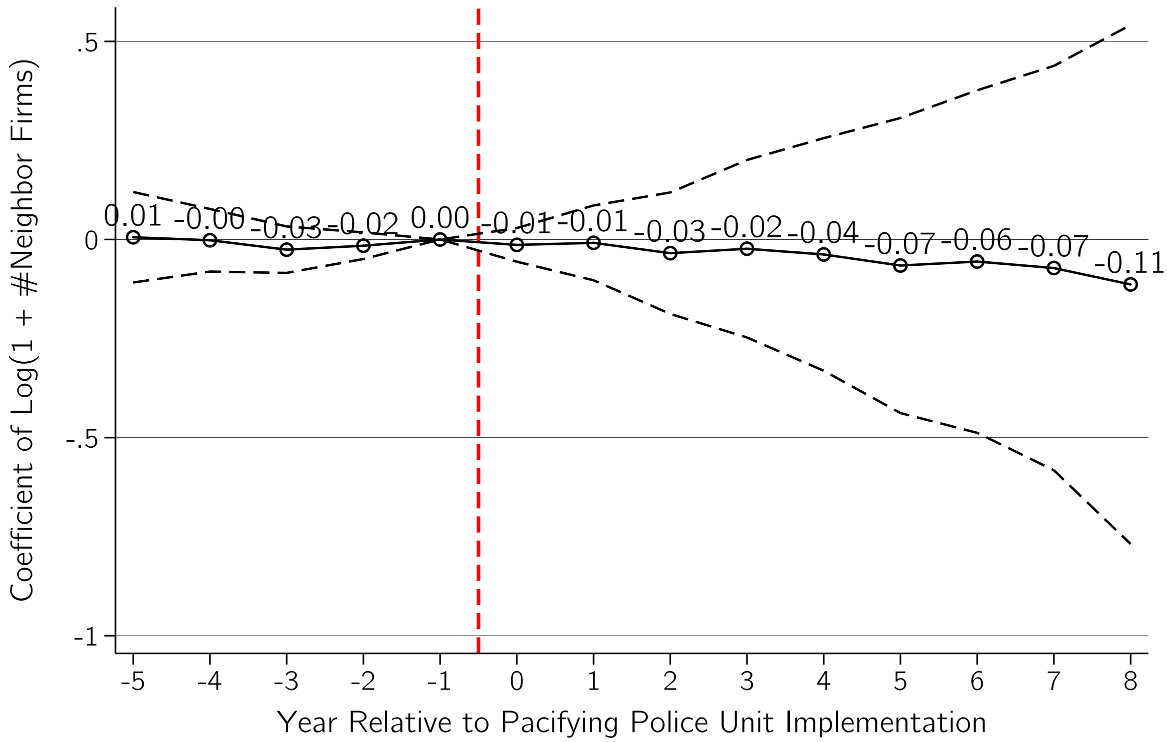


(B) Excluding Favelas Receiving Electrical Renovations from Sample



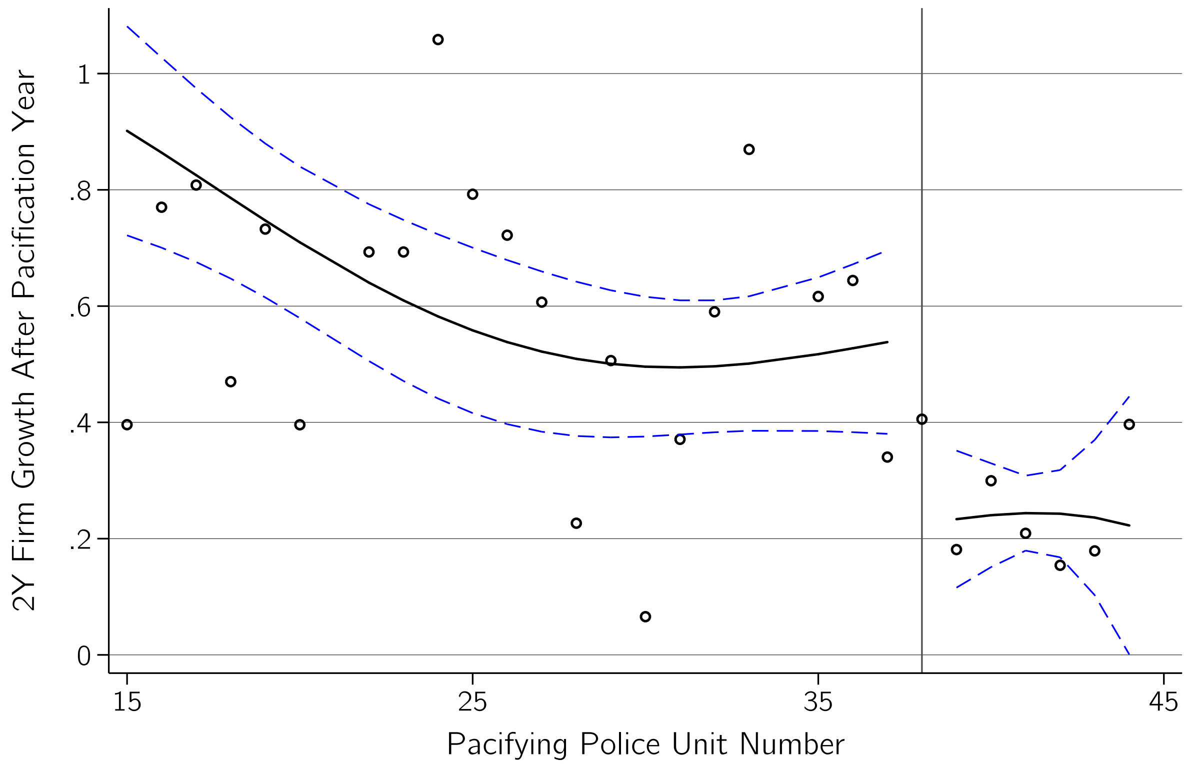
Note: This figure shows the effect of pacification in the favela on the number of firms using alternative specifications. On panel (A) I estimate the model (1.7) controlling for the logarithm of the cumulative state and municipal investments (in each favela) since the year of 2008. On panel (B) I exclude all favelas that received electrical renovations since 2008 from the sample.

Figure 1.15: Effect of Pacification on the Number of Firms in Neighboring favelas



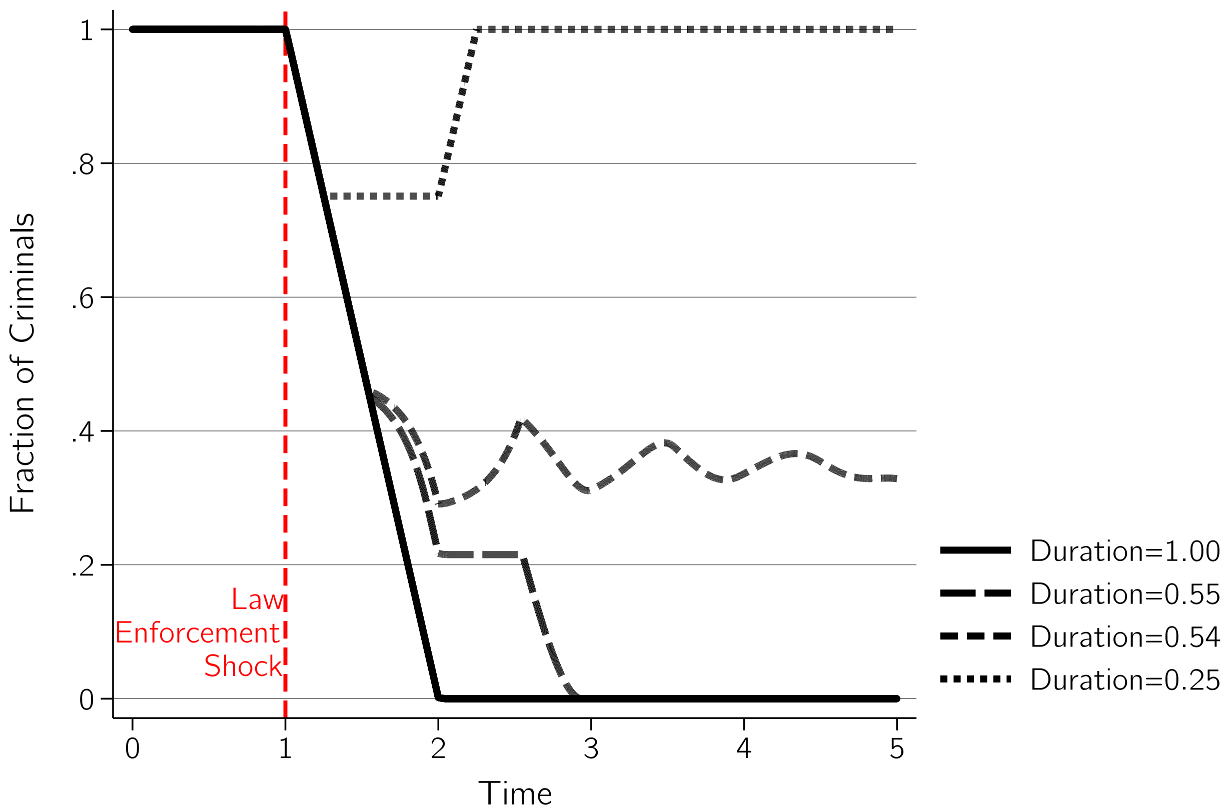
Note: This figure shows the effect of pacification on the number of firms on favelas up to 5 miles away from the pacified favela. The y-axis represents the coefficient estimates from a regression of the dependent variable on dummy variables indicating the year relative to the implementation of the police unit in a given favela. The dashed lines are 95% confidence intervals based on robust standard errors, clustered at favela level.

Figure 1.16: Business Activity in Regions with Implemented and Canceled Units



Note: This figure shows the two years growth in the number of firms in the region corresponding to the implemented or canceled Units of Pacifying Police, starting with the actual or planned year of implementation, along with the fitted values from a third degree polynomial fit on each side of the discontinuity. Points to the left of the vertical line correspond to implemented Pacifying Police Units, while points to the right correspond to canceled Pacifying Police Units. The dashed lines represent 95% confidence intervals.

Figure 1.17: Simulated Effect of a Temporary Increase in Law Enforcement



Note: This figure reports the simulated effect of a temporary increase in law enforcement using the model described in section 1.3.The following parameters were used: $\pi = 0.2$, $\bar{\pi} = 0.94$, and $R_c = 0.2$.

Table 1.1: Descriptive Statistics (2000-2017)

	mean	sd	sd between	sd within
Panel A: Strikes				
Number of days under strike	5.23	4.47	n.a.	n.a.
Percentage of the month under strike	0.17	0.15	n.a.	n.a.
Percentage of cops participating inthe strike	0.58	0.32	n.a.	n.a.
Strike percentage variable	0.10	0.12	n.a.	n.a.
Observations	30			
Panel B: Outcome Variables				
Monthly homicides	162.97	181.30	163.82	83.79
Log of monthly homicides	4.54	1.12	1.07	0.40
Δ Log of monthly homicides	0.00	0.23	0.00	0.23
Monthly retail sales (billions of *reais*)	8.30	13.89	13.47	4.27
Log of monthly retail sales	1.22	1.43	1.41	0.34
Δ Log of monthly retail sales	0.00	0.03	0.00	0.03
Observations	5805			
Number of states	27			

Note: Data on the monthly number of homicides in each state come from the Mortality Information System, a national database of obituaries managed by the federal health department. This database provides data only up to 2016, so data from 2017 come from the Security Secretariat of each Brazilian state whenever available. Data on the monthly sales of the retail sector in each state come from the Brazilian Institute of Statistics and Geography. Data on the strikes come from the main Brazilian newspapers.

Table 1.2: Losses Under 30 Law Enforcement Strikes

	Reais (million)	Dollars (million)
Loss of revenues	3,134.23	1392.99
Loss of value added by retail sector	634.16	281.85
Statistical value of life	0.55	0.24
Losses due to 1,229 additional deaths	676.46	300.65
Cost of hiring substitute police	150.91	67.07
Net benefit of hiring substitute police	1,159.71	515.43

Note: This table reports the estimated losses from the 30 Brazilian strikes between 2000 and 2008. On the first line, I report the revenue loss, defined as the total revenue for the month before the strike multiplied by the strike percentage and by the coefficient $(e^{-0.14} - 1)$. The second line reports the losses in terms of revenues net of costs of goods sold based on the same methodology. The fourth line reports the additional number of homicides caused by the strikes multiplied by the statistical value of life in Brazil, which was obtained from Special Secretariat for Strategic Affairs of Brazil (2018). The fifth line is an estimate of the cost to hire substitute police during the strike period, calculated as the number of strike days multiplied by the daily police payroll for the same state after the strike. The sixth line is the sum of the losses in terms of the value added by the retail sector plus the losses from murders minus the cost of hiring substitute police.

Table 1.3: Placebo Test: Effect of Teachers' Strikes

VARIABLES	(1) Log # Homicides	(2) Log Sales
Teachers' Strike	-0.0362	-0.00800
	[0.0442]	[0.0143]
Observations	5,805	5,805
R-squared	0.357	0.897
Number of states	27	27
State FE	YES	YES
Month FE	YES	YES

Note: This table reports the result of a regression of the dependent variables on a dummy for state-level teachers' strikes, controlling for state and month fixed effects.

Table 1.4: Descriptive Statistics (2003-2017)

	mean	sd	sd between	sd within
Panel A: Police District Level Database				
Homicide rate	49.65	82.10	62.93	53.67
Robbery rate	3601.25	13389.48	12812.72	4388.73
Car robbery rate	354.30	406.96	330.02	243.83
Retail robbery rate	154.30	694.33	624.70	318.91
Truck robbery rate	129.40	578.67	495.72	308.77
House robbery rate	13.37	21.74	18.13	12.34
Bank robbery rate	6.81	52.76	36.88	38.17
Observations	555			
Number of police districts	37			
Panel B: Favela Level Database				
Business density (firms per 1,000 adults)	35.00	109.95	88.72	65.02
2000 total population	1564.92	3510.91	n.a.	n.a.
2010 total population	1827.44	4096.67	n.a.	n.a.
2000 average household income (reais)	322.87	103.79	n.a.	n.a.
2010 average household income (reais)	711.49	237.20	n.a.	n.a.
2000 % of household without income	11.75	8.26	n.a.	n.a.
2010 % of household without income	6.08	5.11	n.a.	n.a.
2000 % literate adults	89.31	5.73	n.a.	n.a.
2010 % literate adults	92.25	4.41	n.a.	n.a.
2000 % of households without running water	8.76	21.54	n.a.	n.a.
2010 % of households without running water	6.96	19.63	n.a.	n.a.
Observations	11445			
Number of favelas	763			

Note: This table reports the descriptive statistics of the data used in section 1.2. Panel A contains information about all police districts in Rio de Janeiro. These data come from the Public Security Institute (ISP), and the variables are measured as the number of occurrences per 100,000 people. Panel B reports the summary statistics of the variables at the favela level. Data on the number of firms come from the Brazilian Federal Revenue. Demographic data come from the Brazilian Census and are collected by the Brazilian Institute of Geography and Statistics.

Table 1.5: Pacified and Nonpacified Regions at the Baseline

	Treatment	Control	Difference	p-value
Panel A: Police District Level				
2007 homicide rate	60.797	49.361	-11.436	.332
2007 robbery rate	1257.036	1503.195	246.159	.7333
2007 car robbery Rate	338.509	351.133	12.625	.8768
2007 retail robbery Rate	31.749	42.463	10.715	.6709
2007 truck robbery rate	49.193	51.19	1.998	.964
2007 house robbery rate	8.649	10.166	1.517	.5068
2007 bank robbery rate	.547	.51	-.038	.9785
Observations	22	15	37	
Panel B: Favela Level				
2007 business density (firms per 1,000 adults)	7.087	6.098	-.989	.6048
2000 average household income (reais)	364.183	359.393	-4.79	.4955
2000 % of household without income	13.727	14.212	.484	.3834
2000 Years of Schooling	5.059	4.99	-.069	.2516
2000 % literate adults	86.384	85.709	-.675	.1005
Observations	153	80	233	

Note: this table reports the statistical tests for differences between the pacified and nonpacified favelas in terms of the observable variables measured at the baseline.

Table 1.6: Test for Reversion of Crime After the Fiscal Crisis

VARIABLES	(1) Log # Homicides	(2) Log # Homicides	(3) Log # Homicides	(4) Log # Robberies	(5) Log # Robberies	(6) Log # Robberies
1(UPP)	-0.309***	-0.429***	-0.427***	-0.275**	-0.259**	-0.239*
	[0.106]	[0.102]	[0.103]	[0.133]	[0.120]	[0.120]
1(UPP) x 1(Year≥2015)		0.473***	0.474***		-0.0641	-0.0530
		[0.104]	[0.105]		[0.113]	[0.117]
Cummulative Investment (Log)			-0.000607			-0.00629
			[0.00529]			[0.00726]
Observations	555	555	555	555	555	555
R-squared	0.330	0.346	0.346	0.170	0.171	0.172
Number of districts	37	37	37	37	37	37
District FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Note: This table reports the estimates of the regression (1.8). If i denotes the police district, t denotes the year, and $y_{it} \in \{\log(1+\text{homicides}_{it}), \log(1+\text{robberies}_{it})\}$ is the crime outcome variable at t of district i, then the regression model used is

$$y_{it} = \gamma_i + \theta_t + \beta \times 1(\text{UPP in } i \text{ at } t) + \delta \times 1(\text{UPP in } i \text{ at } t) \times 1(t \geq 2005) + \varepsilon_{it}.$$

Robust standard errors clustered on the police district level are reported in parentheses. *** represents $p < 0.01$, ** represents $p < 0.05$, and * represents $p < 0.1$.

Table 1.7: Test for Reversion of Business Activity After the Fiscal Crisis

VARIABLES	(1) Log#Firms	(2) Log#Firms	(3) Log#Firms
1(UPP)	0.223**	0.231***	0.239***
	[0.100]	[0.0879]	[0.0886]
1(UPP) x 1(Year≥2015)		-0.0362	-0.0318
		[0.0738]	[0.0739]
Cummulative Investment (Log)			-0.0151
			[0.0109]
Observations	4,194	4,194	4,194
R-squared	0.604	0.604	0.605
Number of favelas	233	233	233
Favela FE	YES	YES	YES
Year FE	YES	YES	YES

Note: This table reports the estimates of the following regression model. If i denotes the favela, t denotes the year, and y_{igt} is the log number of firms at t, if the favela i, then the used regression model is

$$y_{it} = \gamma_i + \theta_t + \beta \times 1(\text{UPP in } i \text{ at } t) + \delta \times 1(\text{UPP in } i \text{ at } t) \times 1(t \geq 2005) + \varepsilon_{it}.$$

Robust standard errors clustered on the police district level are reported in parentheses. *** represents $p < 0.01$, ** represents $p < 0.05$, and * represents $p < 0.1$.

Table 1.8: Effect of Pacification on Business: Regression Discontinuity Evidence

VARIABLES	(1) Δ Log # Firms	(2) Δ Log # Firms	(3) Δ Log # Firms
Pacification	0.363**	0.368*	0.347*
	[0.132]	[0.174]	[0.179]
Red Command		0.127	0.136
		[0.175]	[0.181]
Friends of Driends		0.0744	0.0865
		[0.0816]	[0.0731]
Pure Third Command		0.471**	0.479**
		[0.163]	[0.170]
Militia		0.0955	0.0942
		[0.199]	[0.198]
Log State Investment			-0.0628
			[0.0813]
Observations	167	167	167
R-squared	0.161	0.167	0.169

Note: This table reports the results of regression discontinuity model (1.10), which is based on cubic polynomials fitted on both sides of the discontinuity. in the log number of firms from the pacification actual or planned date to two years after that time. *Red Command*, *Friends of Friends*, *Pure Third Command* and *Militia* are territorial command dummies. *Log State Investment* is the total state investment between the pacification date and 2 years after that time. * denotes $p< 0.1$, ** denotes $p< 0.05$, and *** denotes $p< 0.01$.

Table 1.9: Effect of Pacification on Household Income

VARIABLES	(1) Δ Log Income	(2) Δ Log Income	(3) Δ Log Income
1(UPP before Aug 2010)	0.0697**	0.0732*	0.0723*
	[0.0353]	[0.0407]	[0.0407]
Red Command		-0.00148	-0.00241
		[0.0319]	[0.0322]
Friends of Friends		-0.0286	-0.0294
		[0.0475]	[0.0477]
Pure Third Command		-0.0265	-0.0265
		[0.0364]	[0.0365]
Militia		0.0281	0.0281
		[0.0314]	[0.0314]
Log State Investment			0.00188
			[0.00247]
Observations	483	483	483
R-squared	0.005	0.009	0.010

Note: The left hand side variable is the log variation in the household income of each favela between 2000 and October 2010, when the last decennial Brazilian Census was collected. The first line reports the coefficient of a dummy for favelas pacified before October 2010. *Log Investment* is the sum of all state and municipal investment in a given favela reported in the city and state annual budgets between 2000 and October 2010. The other variables are territorial control dummies. The reported standard errors are robust. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 1.10: Welfare Analysis (2008-2017)

	2008 present value	
	Reais (million)	Dollars (million)
Value added by 1,133 new firms (conservative)	198.08	86.43
Value added by 1,133 new firms (perpetuity)	681.72	297.48
Statistical Value of Life	0.55	0.24
Gains from 1,397 fewer homicides	599.98	261.81
Estimated cost of the program	1021.25	445.64
Conservative net benefit	-223.20	-97.40
Net benefit assuming perpetuity	260.44	113.65

Note: The total cost of the Pacifying Police Unit program is estimated as the sum of (i) the estimated total payroll, which is estimated from the total number of UPP cops and from the average salary of cops in Rio de Janeiro, and of (ii) the total non-payroll costs, extracted from the State Government annual budget. The value added by firms is estimated through equation (1.13), while the value added by the reduced number of murders is estimated through equation (1.12).

Chapter 2

Outraged by Compensation: Implications for Public Pension Performance[1]

2.1 Introduction

Public pensions and sovereign funds hold $21.5 trillion in assets (Official Monetary and Financial Institutions Forum, 2018). When these public funds fail to manage these assets effectively, public sector workers and taxpayers bear the costs to support promised pension payouts to retirees. This paper follows a longstanding literature in examining the potential for politicization of the governance structure of public funds to erode performance following Romano (1993). We complement the prior literature, which has examined distortions arising from politicians' extraction of private-benefits from pension funds' asset management[2], by focusing on the human capital channel. When faced with the threat of outrage over investment manager compensation, politicized pension boards worry about their private-costs from setting high but competitively-appropriate compensation. We model how this political contracting friction affects asset allocation and performance, and then estimate its importance empirically using a global sample of pension plans.

There are good reasons for pension beneficiaries and taxpayers to worry, at least in theory, about costly consequences arising from fears of outrage over the compensation of public pension investment managers. The threat of outrage causes trustees that oversee the public funds to hire lower-skilled managers and offer sub-optimal incentive contracts. This

[1]Based on joint work with Adair Morse and Alexander Dyck.

[2]Hochberg and Rauh (2013) and Bradley, Pantzalis, and Yuan (2016), present evidence of pension fund overinvestment in local assets, leading to lower returns. Andonov, Hochberg, and Rauh (2018) document that politicians on pension fund boards leads to weaker performance in private equities. Theoretical and empirical understanding of the importance of underfunding, and resulting risking-up pressures, for public pensions is found in Novy-Marx and Rauh (2009), Ang, Chen, and Sundaresan (2018), van Binsbergen and Brandt (2015) and Andonov, Bauer, and Cremers (2017).

talent hiring friction affects the fund's asset allocation choices and reduces performance. Public sector funds are particularly prone to outrage concerns because their trustees either are selected by public sector employees or politicians or are themselves politicians. As a result, trustees have career concerns and are sensitive to information emerging in the public domain. What becomes more troubling is that outrage may increase with local income inequality, particularly the difference in the pay of investment managers relative to local workers. Thus, the outrage friction is an additional loss that main street communities face from inequality.

To illustrate the outrage constraint in public pension funds, consider the dilemma of the Oregon State Treasurer in his service as the chair of the state pension fund. The Oregonian newspaper reports:

> *Unspoken, but also politically inconvenient is the compensation to attract talent from the private sector. The state's existing investment officers are some of the best paid public employees, making an average of $200,000 a year. But Treasury officials quietly complain that staff is underpaid by industry standards, and bristle about having to explain and get approval from the Legislature to release performance-based pay each year. As Treasurer Read pleads: "If we have the talent, we will be able to make the decisions better."*

Attempts by Treasurer Reed to hire better-paid investment professionals were rebuffed, with concerns about compensation exceeding members' wages and public pay scales – i.e., outrage.

Appendix Table 1 provides a sampling of similar anecdotes. What is remarkable about the anecdotes is how similar tensions arise across many different types of pension systems and many different geographies of pension funds.

To identify the importance of this human capital channel for public pension fund performance, we first introduce an agency model of portfolio choice. Public pension trustees must hire and compensate an investment manager, who constructs the portfolio over three assets a mean-variance efficient risky asset, a political risky asset that is non-frontier in returns, and fixed income. Boards choose the skill level (ability to capture the risk premium) of the investment manager. Boards then set the manager's compensation contract to attract the desired skill level and incentivize the optimal risk-taking in the portfolio.

The model incorporates the three agency frictions that arise from political influence on the composition of the board of trustees. First, we introduce an 'outrage pay constraint' on skill that binds for some public pension funds. If a pension fund is in a low wage area or has trustees from occupations that are sensitive to wage comparisons (teachers, municipal workers, etc.), the trustees hire managers below a skill threshold to avoid compensation breaching outrage. We also incorporate the previously-documented effects of private benefit extraction and risking-up pressures of unfunded liabilities. Private benefit incentives emerge from political motives (local economy-building and direct vote-chasing) to tilt investments locally, as documented by Bradley et al. (2016), Bernstein, Lerner, and Schoar (2013), Hochberg and Rauh (2013), Brown, Pollet, and Weisbenner (2015), and Dyck and Morse (2011). In addition, private-benefit-taking can emerge from pay-to-play schemes generating campaign contributions or direct side payments (Andonov et al. (2018)). Underfunding affects the risk preferences of boards (Andonov et al. (2017)), as modeled in swinging-for-the-fences or gambling-for- resurrection models of Ang et al. (2018) and van Binsbergen and Brandt (2015).

The model produces comparative statics relating board agency to intermediate outcomes (investment manager skill and the riskiness of asset allocations) and then to the ultimate outcomes of portfolio performance. Of particular interest are the predictions arising from tightening the outrage pay constraint. If the outrage constraint binds, the public pension fund hires lower skill managers. Because managers lack skill to capture the risk premium, they choose to tilt the portfolio towards fixed income and away from risky assets. The fund exhibits lower returns for two reasons: lower performance in each asset class because of poorer skills, and lower overall risk exposure because of the higher portfolio weights in fixed income securities.

To test the theoretical predictions, we use a global sample of 111 public pension funds that account for \$5.4 trillion in assets at the end of our sample period and that cover the U.S., Canada, Oceania, and Europe for 1995-2014. The average (median) fund has \$45 (\$14) billion in AUM. We hand collect data on compensation and biographical information as to the occupations of trustees.

Our empirical methodology mimics our model by setting up a system of two equations, estimated by GMM. In the first equation, compensation is a function of outrage, politicization, underfunding, along with fund characteristics and time fixed effects. In the second equation, performance is a function of outrage-predicted compensation along with other board, fund and time characteristics. With a structurally-motivated, linear system of two equations, we can draw inference concerning the pass-through effect of outrage on compensation.

Our exogeneity condition is that the outrage determinants do not affect within-asset class performance except through the mechanism of managerial contracting. By outrage determinants, we mean the variables which predict that a fund faces beneficiaries who may be more likely to express outrage if investment manager compensation were to be high. We found our ideas of outrage one inequality aversion, whereby disutility depends on the extent to which a wage offered exceeds an individuals' own wage (Fehr and Schmidt, 1999). Thus, we focus on two, related, types of outrage determinants reference wages and occupations of trustees. We measure reference wages with the income of local citizens and the wages of those work in the relevant public sector covered by the pension. Trustee occupation variables are the percentages of municipal workers, teachers and budget civil servants.

Our compensation equation reveals relevant relationships between outrage determinants and compensation. We find that investment manager compensation is about \$90,000 lower in pension funds with a one standard deviation higher level of municipal workers or budget civil servants. Likewise, compensation is \$60,000 lower for pensions whose workers or local citizens have 10% lower reference wages.

We then test for a relationship between outrage-predicted compensation and performance, using our structurally-motivated system to measure the pass-through from outrage to outrage-predicted compensation to returns. We find that a one standard deviation higher fraction of municipal workers or budget civil servants, passing through compensation, results in 6.5 basis points lower portfolio net returns, benchmarked at the asset class level. Across outrage determinant variables, a standard deviation change in making the pension less outrage-prone yields a tight distribution of returns implications (5 bps to 8.5 bps), lending credibility to our identifying a canonical relationship.

The performance result is driven by a strong link between outrage-predicted log compensation and within-asset class returns in two asset classes alternatives (20.9 basis points) and public equities (6.9 basis points). Consistent with our theory and anecdotal evidence, these asset classes are where investment managers may exhibit more experience either in realizing the risk premium or avoiding the paying of high delegation fees for performance. Rich data allow us to look at the percent of delegation in pensions' portfolios. We find that outrage-predicted compensation strongly associates with the degree to which a fund delegates within its asset class. An increase in reference wages by 10% or an increase in outrage-determining occupations on the board result in higher delegation of 4 to 8 raw percentage points.

We document that these results are not driven by funds that are insulated from the outrage threat having greater realized net risk. Realized tracking-error is uncorrelated with outrage-predicted compensation, in the portfolio as a whole or within asset classes.

For an average fund, our estimates suggest that if that fund were to find a way to attenuate the factors constraining management pay (e.g. political board participation) by one standard deviation, it would benefit by producing additional annual benefits of \$22 to \$38 million in annual value-add (using the method of Berk and van Binsbergen (2015a)). Reducing all political appointees or going further to an independent skills-based board produce even higher estimated returns of \$90 to \$114 million in annual value-added.

Finally, consistent with the prior literature, we find that distortions arising from politicians' payoffs to local investment and distortions arising from underfunding also impact asset allocation and returns. Importantly, including them in the model and in our regressions does not eliminate the importance of the human capital channel. Consistent with Andonov et al. (2017) we find that underfunding leads to increased asset allocation to alternatives. Consistent with Andonov et al. (2018) and Hochberg and Rauh (2013) we find that politicization has a direct effect on returns in alternatives asset classes. We interpret our results as complementing these papers, showing an important and neglected human capital channel whereby politics can also undermine returns.

In exploring the impact of politicized governance for public fund outcomes, our paper contributes to a large literature. Romano (1993), for example, hypothesized and found that politicization affects fund performance focusing on a sample of 50 public funds in the 1980s. While our results are broadly consistent, our focus on the human capital channel leads to different policy conclusions. Romano identified the key friction being the lack of accountability of appointed/ex officio trustees, and advocated solutions around increasing the proportion of beneficiary-elected trustees. Focusing on human capital, the key policy implication is to insulate trustees from outrage concerns. Increasing the proportion of beneficiary-elected trustees could exaggerate exposure to outrage concerns, absent other steps. Moving to an independent skills-based board, which could be appointed by beneficiaries and government, has greater potential to limit outrage pressures and improve human capital outcomes. Modifying pension agreements to share risk of underfunding between beneficiaries and government will contribute to such reforms, as both sides will be motivated to avoid the costs of outrage.

The rest of the chapter is organized as follows. In section 2, we fix ideas by introducing a theoretical model of portfolio choice with political agency costs and management contracting. Section 3 lays out our empirical methodology, and section 4 describes our data. In section 5 we present our empirical results. In section 6 we conclude and consider alternative pension policy remedies.

2.2 Model of Portfolio Choice with Political Agency Costs

Imagine a setup in which beneficiaries of a pension fund would optimally invest in a mean-variance efficient portfolio over a risky asset and fixed income. The board of trustees for this pension fund achieves this objective by making manager-contracting choices to maximize beneficiaries' utility subject to manager participation and incentives. In our setting, because the pensions are public pension funds, being in the political domain can affect trustees' incentives and decisions. Although trustees have a fiduciary duty to act in the best interests of their beneficiaries, political private costs and benefits from their funds' choices create incentives to deviate from a strict interpretation of this duty. We call the resulting distortions political agency costs.

Our model and empirical analysis consider three political agency costs. The first emerges from outrage, the inability of politicized boards to pay optimally for investment manager skill because of political costs emerging from workers, retirees, and other voters in the community. The second emerges from politicized boards' preference for investing in political assets. Political assets are defined as investments which generate private benefits to a political board member, either in the form of local-tilted assets (which generate positive media attention, reputation, and ultimately votes and legacies) or in the form of pay-to-play allocations (which involve kickbacks from asset managers to politicians or political campaigns in return for asset allocations). The third political agency cost emerges from the pressure of liabilities that can induce public pension fund boards to risk-up portfolios to meet funding needs (e.g., to pay pensioners) rather than to have to face disclosure of shortfalls.

The focus of our model is on how these political agency costs affect allocations and performance, working through the mechanism of hiring and compensating an investment manager.

Assets and Investment Manager Heterogeneity

A public pension fund board hires and sets a linear compensation contract for an investment manager to allocate the pension's capital among assets. Managers are risk averse and are assumed to have the same risk aversion as the beneficiaries of the pension fund, λ. Managers are heterogeneous in one dimension, their skill in the selection of assets within each asset class (or in the selection of asset manager for delegation within each asset class), represented by the parameter s. Skill levels are transparent, and their supply is perfectly competitive. A manager of type s has an outside option $O(s)$, where $O(\cdot)$ is an increasing function such that skilled managers have higher outside options.

The manager chooses portfolio weights among three assets: fixed income, a mean-variance efficient risky security (MV security) and a political asset. Fixed income pays a riskless return r_f:

$$\text{Fixed Income: } E[R_f] = r_f.$$

The MV security has variance σ^2_{MV} and risk premium ϕ_{MV}:

$$\text{MV security: } E[R_{MV}] = r_f + s\phi_{MV}.$$

The political asset is also risky but has variance σ^2_P and risk premium ϕ_P.

$$\text{Political Asset: } E[R_P] = r_f + s\phi_P.$$

We assume that $\phi_P/\sigma_P < \phi_{MV}/\sigma_{MV}$ so that the MV security dominates the political asset in Sharpe ratio terms.

In both risky securities, managers earn a fraction s of the potential risk premium, in proportion to their skill. Only managers with maximal skill (i.e., $s = 1$) can capture the full risk premium with their asset selections. This assumption is empirically motivated; while some investment managers in public pension funds have significant financial experience from working previously in a finance position in a public pension fund or the private sector, others prior experience is limited to a managerial or civil servant role with no asset management responsibilities.

Differences in s can also be interpreted as delegation costs. If managers delegate portfolio management (or a fraction thereof) to external institutions, they incur intermediation fees, reducing the effective fraction of the risk premium earned by the fund. The skill variable s captures both the managers' skill and the ability to economize on intermediation costs, such as internally managing assets.

Managers form portfolios by selecting the weights on MV-efficient securities, political assets, and fixed income as w_{MV}, w_P, and $(1 - w_{MV} - w_P)$, respectively.[3] For tractability and consistent with Hochberg and Rauh (2013), we assume that the MV security and political assets have a joint normal distribution with correlation , which is large enough to prevent hedging between asset classes.[4]

[3] A pension fund not affected by agency problems would invest in a combination of the MV security and fixed income.

[4] See the appendix for the explicit restriction that prevents the portfolio manager from taking short positions in any asset class.

Utility & Political Agency Costs

Under the assumption of mean-variance preferences, the utility of the board equals that of beneficiaries if no political agency costs are at work:

$$U_{board}^{noagency} = U_{beneficiaries} = E[R - manager\ pay] - \frac{1}{2}\lambda Var[R - manager\ pay] \tag{2.1}$$

where R is the total return of the portfolio; manager pay is the compensation paid to an investment manager; and λ is the risk aversion of beneficiaries. We introduce three political agency costs that cause the board's utility to deviate from that of the beneficiaries.

Outrage Pay Constraints

First, trustees in public pension funds are in a political domain, and this leads them to consider potential political costs arising from their choices. In the typical pension plan, trustees are either beneficiaries or politicians who employ and pay the beneficiaries. Costs arise for trustees if beneficiaries or others in the community who elect politicians become outraged by the compensation of investment managers. A potential basis for outrage for beneficiaries and those in the community is inequality aversion. For example, using distribution experiments, Engelmann and Strobel (2004) find that most people value equality more than efficiency. In practice, these private outrage costs usually take the form of negative media attention and the resulting negative reputation consequences.

If the board were to set compensation sufficiently high such that outrage occurred, it would have to bear some utility cost:

$$U_{board}^{noagency} = U_{beneficiaries} = E[R - manager\ pay] - \frac{1}{2}\lambda Var[R - manager\ pay] - outrage\ cost \tag{2.2}$$

If trustees' utility consequences of outrage are sufficiently large, they would want to preclude the possibility of outrage altogether. The easiest way for trustees to ensure that compensation, which is stochastic, does not go over the outrage threshold is to hire lower quality managers. To model this intuition, we assume that each fund has a threshold on skill, $s^o utrage$. Thus the board's utility reverts to equation (1), but with a constraint:

$$\begin{aligned} \text{maximize} \quad & U_{board} = E[R - manager\ pay] - \frac{1}{2}\lambda Var[R - manager\ pay] \\ \text{subject to} \quad & s \leq s^{outrage}. \end{aligned} \tag{2.3}$$

For some funds, the threshold is large and never binding. This is more likely if the reference wage level of beneficiaries or others in the community is sufficiently high.

Private Benefits from Politicized Investing

Second, allocation choices can create private benefits for political trustees. These private benefits include votes garnered from investing locally and creating employment opportunities for local citizens, or side-payments (e.g. in the form of campaign contributions or direct payouts) from pay-to-play arrangements.[5] We incorporate the political agency cost from private benefits from politicized investing in our model by assuming that the board receives a riskless, private benefit worth dollars for each dollar invested in political assets:

$$U_{board} = E[R - manager\ pay] - \frac{1}{2}\lambda Var[R - manager\ pay] + \kappa w_P. \tag{2.4}$$

Liability-Driven Preference for Risk

Finally, effective board risk aversion, λ_B, can be affected by liability obligations of the pension fund. Ang et al. (2018) model the tensions pensions face due to the constant need to fund payments to retirees. Their main inference is that when funding is low, pension boards have a lower effective risk aversion; i.e., a desire to "swing for the fences." The friction often at work is that boards having to go back to legislatures to request funds to cover a down year of returns face a personal reputational cost. The resulting risk-taking behavior is similar to the gambling for resurrection ideas of van Binsbergen and Brandt (2015). Such increased risk taking in the presence of underfunded liabilities has been found in US public pension funds, for example, by Andonov et al. (2017).

We assume that underfunded status results in a higher risk appetite:

$$\lambda_B = \frac{\lambda}{\theta} \tag{2.5}$$

where θ is an exogenous politically-determined variable that captures the risking-up pressure. The final utility formulation for the board, incorporating all political agency issues, is thus given by:

$$\begin{aligned} \text{maximize} \quad & U_{board} = E[R - manager\ pay] - \frac{1}{2}\lambda_B Var[R - manager\ pay] + \kappa w_P \\ \text{subject to} \quad & s \leq s^{outrage}. \end{aligned} \tag{2.6}$$

Solving for the Optimal Contract and Manger Skill

We solve the model by considering the post-hiring portfolio choice, assuming that a manager with skill s already is hired. The board asserts its preferences for risk and for political

[5] Andonov et al. (2018) find that U.S. pension funds with political boards tend to invest in local and less profitable private equity funds, Dyck and Morse (2011) and Bernstein et al. (2013) show a similar pattern in the investments of sovereign wealth funds. Bradley et al. (2016) show not only a local bias but a bias to invest in politically-connected firms.

investments by offering a compensation contract to the investment manager to induce the preferred portfolio choice. We derive this optimal contract for any skill level s. Next, we calculate the optimal manager skill s chosen by the board, from which we can figure out the resulting asset allocation.

We restrict our model to linear contracts. The manager receives a cash salary c, independent of her performance. In addition, the board gives a share 1a of the realized financial return to the manager to induce risk-taking. The board also asserts its political preferences by giving the manager an additional transfer of b dollars for each dollar invested in political assets. Linear compensation is given by:

$$manager\ pay(R, w_P|c, a, b) = c + (1 - a)R + bw_p \tag{2.7}$$

Like the beneficiaries, we assume that the investment manager has CARA utility with risk aversion λ. Thus, the manager chooses risk and political asset weight (w_{MV}, w_P) solving the following program:

$$\max_{w_{MV}, w_P} U_M = \max_{w_{MV}, w_P} \left\{ E[manager\ pay] - \frac{1}{2}\lambda Var[manager\ pay] \right\} \tag{2.8}$$

The board maximizes the expected monetary payoff penalized by the variance, with penalizing factor $\lambda_B = \lambda/\theta$, which depends on the risking-up pressure θ. The optimization problem is restricted by: (i) the manager's incentive constraint and (ii) the manager's participation constraint, which obligates the board to offer a contract that generates an expected utility for the manager not smaller than her outside option $O(s)$.

The participation constraint is the channel connecting political asset investing to manager contracting. Because political assets are dominated in performance by the MV security, more politicized boards realize smaller utility increments from the skill of managers. Thus, the higher the political benefits κ are, the less willing is the board to pay compensation for skill.

The underlying program, which defines the optimal contract and the indirect utility $V_B(s)$ of the board when hiring the manager with skill s, is given by:

$$\begin{aligned} V_B(s) \quad &= \max_{c,a,T} U_{board} = E[R - manager\ pay] - \frac{1}{2}\lambda_B Var[R - manager\ pay] + \kappa w_P \\ \text{subject to} \quad & \\ & c + (1-a)E[R] + bw_P - \frac{1}{2}\lambda_M (1-a)^2 \mathrm{Var}[R] \geq O(s) \\ & \{w_{MV}, w_P\} = \underset{w_{MV}, w_P}{\mathrm{argmax}} \{U_M | c, a, b\} \end{aligned} \tag{2.9}$$

In the appendix, we show that the optimal contract is given by:

$$a^* = \frac{\lambda}{\lambda + \lambda_B} \tag{2.10}$$

$$b^* = (1 - a^*)\kappa$$

The optimal payment factor a^* reflects the standard sharing rule in which the less risk averse agent receives a larger component of the risky outcome. In the optimal contract, the manager receives the same fraction $1 - a^*$ of the financial return R and of the political return κ. The resulting base salary c^* is the number that makes the participation constraint binding.

Finally, the board will choose the manager skill that satisfies the outrage constraint (if local reference wages are low) and maximizes their ex-ante utility:

$$\max_s V_B(s), \text{ s.t. } s \leq s^{outrage} \tag{2.11}$$

If the outrage constraint is not binding, then marginal disturbances around the optimal s^* are such that the marginal increase on the squared Share ratio is equal to the marginal cost of hiring a slightly better manager. If outrage is binding, the public pension fund will hire a lower skilled manager, foregoing opportunities for an increase in the portfolio Sharpe ratio.

Comparative Statics

The solution to (11) illustrates how funds differ in their performance-cost tradeoff when choosing manager skill. For instance, boards facing high private benefits κ from political investing as well as boards facing an outrage constraint on compensation both prefer to hire managers with lower skill compared to the optimal manager for the beneficiaries. On the other hand, boards facing a personal cost from not having enough returns to cover pension liabilities might optimally choose a higher-skilled manager to benefit from risking-up the portfolio. Table 1 reports these comparative statics, focusing not just on how the agency issues affect manager contracting of skill, but to how ultimately these frictions translate into portfolio choice effects – allocations and performance.

Panel A isolates the effect that the outrage constraint being binding or not has on performance and allocations. The mechanical consequence of a binding outrage constraint is that the board of an outrage-prone pension fund hires a less skilled manager ($\Delta s < 0$). The lower skilled manager realizes lower risky asset returns ($\Delta R_{MV} < 0$, $\Delta R_P < 0$)); thus, the board optimally sets a contract to induce more portfolio weight on fixed income ($1 - \Delta w_{MV} - \Delta w_P > 0$). There is no point in paying compensation for extra risk not rewarded with a capture of extra risk premium. The combination of more investment in fixed income and weaker managerial skill adds up (on both counts) to a portfolio with poorer overall expected performance ($\Delta R < 0$).

Panel B looks at the partial derivatives with respect to changes in the other political agency issues. Boards with greater benefits from investments in political assets ($\partial\kappa$) hire less skilled managers, since the expected return payoff from skill is lower in the portfolio tilted toward the political asset. Lower skill leads to smaller within-asset-class expected returns ($\Delta R_{MV} < 0$, $\Delta R_P < 0$) and less investment in the MV security ($\Delta w_{MV} < 0$). In addition, these boards design contracts to incentivize greater investment in the political asset ($\Delta w_P > 0$), which further reduces overall performance ($\Delta R < 0$).

By contrast, boards with higher liability-driven risk-up pressure (larger θ) hire more skilled managers to take more advantage of the risky asset classes ($\Delta s > 0$, $\Delta w_{MV} + \Delta w_P > 0$), hence increasing within-asset class and overall performance ($\Delta R_{MV} > 0$, $\Delta R > 0$). The extra risk that these boards induce may be rewarded with realization of additional returns, but it is above the level of risk desired by the beneficiaries. As stakeholders and taxpayers, beneficiaries may find themselves bailing out pension liabilities from taxes when bad returns realizations occur.

Although we do not explicitly include the cross partials in Table 1, one final piece of intuition is worth highlighting. When public pension funds have high liability pressures, the effect of an outrage constraint is very damaging: in this situation, public boards incentivize a poorly-skilled investment manager to take on more risk, ending up with a riskier portfolio that underperforms in the risky asset classes.

2.3 Empirical Methodology

Our goal is to estimate how agency affects public pension fund outcomes working through the compensation contract mechanism. Although we are interested in the other political agency issues, we set up our system to focus on the mechanism of outrage, because we can make plausible exogeneity arguments and because the novelty of our paper vis--vis the prior literature is in the introduction of outrage.

We employ a linear sytem of two equations, estimated through GMM. We choose a linear system approach, rather than a structural model approach, for three reasons. First, our dataset of compensation observations is limited in sample size, making inference of more complex non-linear moment optimization problematic. Second, the point of the model is to motivate comparative statics by combining agency with portfolio choice rather than providing a framwork for the exact parameter calibration. Third, because our model is one of outrage working through the mechanism of compensation contracts to distortions in performance, outrage only affects outcomes through the management contract. This restriction lends itself to a linear structural GMM specification, where we can make linear exogeneity assumptions as if we were in the familiar instrument setting. Our linear system of equations, with subscripts i and t respectively refering to the public pension fund and year, is as follows:

$$\begin{aligned} Log(Manager\ Compensation)_{it} &= Outrage_{it}\Phi_1 + \phi_2 Underfunding_{it} \\ &\quad + \phi_3 PoliticalChair_i + X_{it}\Gamma^{eq1} + \varepsilon_{it}^{eq1} \end{aligned}$$

$$\begin{aligned} Performance_{it} &= \beta_1 Log(Man\hat{a}gerComp)_{it} Outrage_{it}\beta_1 + \beta_2 Underfunding_{it} \\ &\quad + \beta_3 PoliticalChair_i + X_{it}\Gamma^{eq2} + \varepsilon_{it}^{eq2} \end{aligned}$$

The equations are naturally dynamic in events; the manager contracting happens first, followed by the realization of returns. In System Equation I, the Outrage variables include (i) trustee occupation variables and (ii) reference wage variables. System Equation I also includes the covariates from System Equation II (the log of lagged public pension fund size and year fixed effects) and the two other political agency variables, Political Chair and Underfunding. System Equation II takes the outrage-predicted compensation as predetermined, included alongside Politicial Chair and Underfunding, as well as controls of lagged fund size and year fixed effects. We estimate this system using GMM and cluster standard errors at the fund level.

We are interested in interpreting outrage working through the mechanism of compensation on performance. The exogeneity condition for a causal interpretation is that outrage variables are exogeneous to performance conditional on compensation. We contend that this condition is plausible because the outrage variables, described in the data section, either reflect board composition or local income levels that should be unrelated to investment performance, except through any effect on management quality.

We do not make the same exogeneity assumption when we consider Political Chair and Underfunding. A policitized chairperson might steer investment choices for political private benefits through pay-to-play arrangements or local favoritism. Likewise, underfunding may not only impact compensation, but also could directly impact portfolio choice by triggering active intervention of the board. Thus, we set up the system so that we can use Outrage, but not the other agency variables, as predetermined causes of some variation in compensation that can later potential explain performance.

2.4 Data

Public Pension Funds Sample

Our sample is from the union of two sets of public pension funds. We source U.S. public pension funds from the Center for Retirement Research (CRR) dataset at Boston College. Globally, we collect all public pension funds with over $10 billion in assets identified in Pensions & Investments in 2011. Because of the need to search manually for the personal characteristics and compensation of trustees and managers, we limited the sample to funds in North America, Oceania, and Europe. Table 2 defines all variables and their sources. We convert all monetary data to 2010 U.S. dollars.

Table 3 reports statistics about our sample of public pension funds. As Panel A reports, the full sample consists of 164 funds and 1,856 fund-year observations. The mean and median pension fund have $45 billion and $14 billion in assets, respectively. Panel B reports our estimation sample that consists of funds with compensation and trustee data. The cross-section remains large, covering 111 public pension funds, but we only have a short panel, with 463 fund-year observations. Our estimation sample reflects larger funds, with a mean and median of $102 and $30 billion in assets respectively.

On the right-hand side of the table we report gross portfolio returns. The mean gross portfolio returns are similar between the full sample with 4.3% gross return (Panel A) and the estimation sample (Panel B) where the gross return is 4.2%. As both Panels show, although our sample favors U.S. pension funds, over a third of the sample is from Canada, Europe, and Oceania. Our results are not just reflective of a U.S. story.

Allocations and Performance Data

In terms of portfolio choice variables, we collect each fund's asset allocations, performance and the fraction of assets managed via delegation over 1995-2011 from a combination of sources: annual reports, funds' current and cached websites, direct requests to the funds, the Boston College CRR dataset and CEM Benchmarking. We analyze portfolio weights and performance in three primary asset classes: (i) alternatives (hedge funds, private equity and real estate), (ii) public equities, and (iii) fixed income. We order these asset classes in decreasing risk. When we make inferences, we assume that alternatives not only have the highest expected risk, but they also provide the greatest opportunities for private benefit-taking by politicians because of their "2-and-20" compensation structure, which affords opportunities for kickbacks and tilting of portfolios towards local investing.

Table 4 reports portfolio summary statistics, starting with allocations in Panel A. We present two sets of portfolio weights those for the sample in which we observe portfolio weights in the corresponding asset class, and those restricted to observing all portfolio weight allocations across the portfolio (used in the weight estimations). The mean distribution of allocations is public equities (0.513), fixed income (0.296), and alternatives (0.191). Table 4 also presents statistics on our delegation variable, defined as the fraction of assets managed by external institutions in each asset class. On average, the fractions of assets managed via delegation are 0.500 for fixed income, 0.734 for equities, and 0.747 for alternatives (excluding hedge funds, which are all outsourced).

Panel B of Table 4 reports performance statistics. At the portfolio level, mean gross and net returns are 4.2% and -0.3% respectively. The benchmark we use is the CEM consulting benchmark. The trustees of a fund select the benchmarks, not the asset manager. These asset class benchmarks, more often than not, are a combination of sub-asset class indices, with appropriate adding-up weights, to reflect the desired portfolio risk.

As another measure of performance, we use the closeness of the investment manager performance relative to benchmark performance, i.e., the realized tracking error. We estimate in-sample, fund-level tracking error, as the standard deviation of the error term in a no constant model where we regress each fund's annual realized return on its benchmark. We produce one measure of tracking error per fund, with a cross-sectional mean tracking error of 0.030 across 110 funds. Not surprisingly, tracking error is highest in alternatives, then equities, and finally in fixed income.

Investment Manager Compensation and Skill Data

We hand-collect compensation data for investment managers. For funds with mandated disclosure, we successfully search for compensation in annual reports and public filings. For the other funds, we search for each named manager and public pension fund in newspaper databases. Newspapers are sometimes able to access compensation information based on freedom of information requests. As we search, we look for the highest paid investment executive, which could be either the CEO or CIO, depending on the fund. The resulting sample covers 111 public pension funds with a total panel of 463 observations, including all geographies spanned by our sample. We report summary statistics on compensation in our dataset in Table 5. The median total compensation of the investment executives is $537,197 USD, with a mean of $807,416. A quarter of the fund managers make salary of $292,328 or less. These are large numbers, but recall that observability limits our sample to large funds, and these managers control pension funds of $102 billion on average.

Our model refers to manager skill, which induces higher compensation. Although we do not have a quantitative measure of skill, we hand gather the prior professions of all investment managers. Table 6, Panel B, reports the breakdown of the immediately prior job these managers held. For almost two thirds of the fund managers their immediate prior experience was in finance, with 4.9% of managers working as a senior investment manager at another pension fund, 31.1% in the private sector in a financial capacity, and 30% as a bureaucrat with financial responsibility. But notably for the other third of investment managers, their prior experience was either as a civil servant with no financial expertise or as a non-financial executive in a pension fund (16.4% + 18% = 34.4%).

Figure 1 depicts box plots of the distribution of compensation by prior profession categories. The [red] dashed vertical lines present the quartile cutoffs from the full sample. The non-finance professionals (non-budget civil servants and pension executives) together account for 34.4% of the sample and clearly earn lower compensation. The mean compensation of non-budget civil servants is only $244,372. Even two standard deviations higher compensation for these individuals does not put them in the realm of the median (or mean) compensation for everyone else. The non-finance pension executives fare a little better, with a mean of $459,576. However, the box plot well portrays that the skew in this category is large; most investment managers with non-finance pension experience have quite modest salaries. The lack of compensation for these public servants reflects strongly the dialogues presented as outrage examples in Appendix Table 1. For example, a recruiter quoted in the New Mexico State Investment example (7 in Appendix Table 1) states: “Pay scales in public plans tend to reflect the pay scales for the state bureaucracy.” In the Missouri State Employees Retirement example (5), the state senator in charge of appropriations calls the idea of bonuses to investment managers “unconscionable” in lieu of payments to services for the disabled, college scholarships, etc.

Outrage Variables

With inequality aversion, disutility depends on the extent to which a wage offered exceeds an individuals' own wage (see Fehr and Schmidt (1999)). Our outrage variables to capture such reference points are of two types – (i) beneficiary-trustee reference wage variables and (ii) trustee occupation variables. The first reference wage variable is the wages of the working beneficiaries. We collect information on the average wages of working beneficiaries either directly from the annual report or as a calculation from data on the employee contributions and the reported average rates of contributions (also predominantly from funds' annual reports). As reported in Table 5, the average wage of working beneficiaries is \$47,811.

The second reference wage we collect is the average household income in the municipality (or MSA) where the fund is located. For each fund we look for the finer measure of regional income calculated by the agency responsible for collecting and compiling income statistics in each country. We presume board members are also likely to be drawn from the same region, and would be sensitive to this average wage. The average local household income (Table 5) is \$55,434. Both measures have a tight and quite symmetric distribution.

Trustee occupation variables emerge from, first, sourcing the names of the trustees from the websites and, then, looking up biographical information from c.v.'s on the funds' websites or other web information sources (e.g., Linkedin). Data availability force us to use a single cross-section of data (2011) for trustee biographies. We were concerned about this limitation. However, empirically, the average fund is in the data for three years, making the board information for one year likely to be relevant for the entire sample period.

Table 6, panel B provides a tabulation and descriptions of professional titles. We split the table into broad categories of Civil Servants and Non-Civil Servants, each representing about half of the mean distribution of trustees. We further break Civil Servants into politicians, budget civil servants and other civil servants. Politicians (those representing the government at large or elected as a politician) are somewhat rare as non-chair trustees, accounting for 6.4% of board seats. Budget civil servants (most commonly, treasurer, revenue commissioner, controller, auditor, and finance directors) hold 34.4% of seats. These civil servants are particularly sensitive to pay levels in public service as they set compensation across multiple government agencies. Other civil servants (clerks, commissioners, public university academics, and legal government officials) hold 13.7% and are not generally involved in pay determination in their regular employment. Among non-civil servants, teachers represent 14.7% of the mean distribution. Next are municipal workers (7.7%), who are fire workers, librarians, workers at city hospitals, and other such public municipal service occupations that are not internal to the running of the government administration per se. Finally, the largest non-civil servant category is professionals (23.1%), who are financial sector professionals as well professionals from medicine, media, NGOs, or other private firms.

We use three board occupation categories to capture outrage Municipal Workers, Teachers, and Budget Civil Servants. A trustee is more likely to perceive costs from outrage, and thus be more likely to want to implement outrage pay constraints on investment manager compensation, if she herself has a history as a local worker (variables: Municipal Workers and Teachers), or if she is involved in the finances of the local government directly (Budget Civil Servants). The exogeniety condition asserts that these trustees do not influence performance except through their role in manager contracting.

One concern would be if our use of trustee occupations as outrage variables were correlated with politicians on the board, which prior research has found has a causal effect on portfolio performance (Andonov et al. (2018)). We do not think this to be the case because Municipal Workers, Teachers, and Budget Civil Servants do not account for all of the non-politician variation in the other civil servants and not-a-civil-servant categories. Nevertheless, to make sure that the main estimations are not driven by an effect of "one-minus Adonov et al (2017)", we include an appendix of estimations defining the outrage variables based on professional designation as a fraction of the non-politician board members.

Political Chair and Underfunding Variables

Using the data we collected on the process by which each member of the board is appointed or elected, we construct a dummy variable called Political Chair, which takes a value of one if the chair is appointed by an executive of government (e.g., governor, mayor, finance minister, king, etc.) or a ministry of government. Fifty-one percent of boards have a Political Chair.

Finally, we measure the extent of underfunding pressures by creating an index of two variables. We have data on the funded ratio (the level of assets-to-liabilities), but not for all funds. The other measure of liability strain comes from Rauh (2008), who finds that funds with a higher age profile of pension beneficiaries have more liability concerns. Thus, we construct the average age of pension beneficiaries, using data on the average age of workers and retirees with the fraction of members being retired. Then we construct the Underfunded Index as the negative of the standardized funded ratio plus the standardized age variable. The underfunded index has correlations of 0.81 with age and of -0.79 with the funded ratio.

2.5 Results

Do Outrage and Political Agency Issues Affect Compensation?

In Figures 2a and 2b, we plot the mean and median compensation across terciles of the outrage variables (Municipal Workers, Teachers, and Budget Civil Servants in 2a and the reference wage variables in 2b). The plots show, with some variation, that compensation is lower with higher outrage occupations percentages and lower reference incomes. The exceptions are in the Teachers' and Worker Wages' plots, which show non-monotonic patterns across the terciles of the x-variable. These patterns may be due to not controlling for pension fund size in the plots; thus, we turn to multivariate results. Table 7 reports the relationship between compensation and political agency variables. As a baseline, in column (1), we regress log compensation on lagged fund size and year fixed effects. Lagged fund size significantly associates with compensation, but with a limited explanatory power (an R-squared of 0.0365).

In columns (2) to (4), we iteratively add in the outrage and other political agency variables. Column (2) adds the trustee composition outrage variables in addition to the baseline controls. All three trustee composition outrage variables – Municipal Workers, Teachers, and Budget Civil Servants – negatively associate with compensation, but only Teachers and Budget Civil Servants are statistically significant. Notably, the R-squared increases sharply to 0.115. Column (3) instead includes the reference wages outrage variables and the baseline controls. We find a positive and significant relationship of both Regional Income and Worker Wages with compensation. The elasticity of manager compensation to reference income is between 0.6 and 0.9. Furthermore, the R-squared in column (3) increases materially relative to column (1) to 0.106.

In column (4), we explore the relationship between other political agency issues (Political Chair and Underfunded Index) and compensation. We find a strong negative association between Political Chair and compensation and an insignificant impact of underfunding. The partial R-squared of political chair is weaker than outrage, but this in no way contradicts the prior literature (Andonov, Hochberg and Rauh, 2017; Andonov, Bauer and Cremers; 2017), as political influence may work directly in the investment choices.

Finally, in column (5), we include all sets of variables and find that most of the results in the prior columns are independent of each other; the R-squared continues to increase (to 0.153) and most variables remain robustly significant. Controlling for the other effects also adds precision in the estimation, making Municipal Workers and Underfunded significant.

In Panel B, we evaluate the economic impact of larger agency concerns for all statistically significant political agency variables, using the column (5) estimates. A board of trustees has on average 11 trustees. A one standard deviation higher fraction of Municipal Workers or Budget Civil Servants implies a higher fraction of these occupations by 0.087 to 0.144 (i.e., 1 to 1.6 trustee members). For the reference income variables, we study the elasticity effect of a 10 percent change, which maps to about \$5,000 higher income for the reference wage group.

We find that a one standard deviation change in the fraction of either Municipal Workers, Budget Civil Servants, or Political Chair implies approximately a \$90,000 lower manager compensation (averaging over \$76,033, \$107,627, and \$94,209). A 10% change decrease in either Regional Income or Worker Wages implies approximately \$60,000 less manager compensation. These effects are 7 to 13 percent of the mean compensation in the sample. Underfunding has a positive but lower effect on compensation, consistent with our model that the trustees will want the manager to risk-up, thereby making it desirable to hire a manager who can better capture risk premia. All of these effects are consistent with our model comparative statics and our intuition of how agency affects manager contracting.

Do Outrage Pay Constraints Affect Performance?

Return Results

Section 3 laid out our empirical methodology as a two-equation, linear sytem of equations to estimate how agency affects public pension fund outcomes working through the compensation contract mechanism. Table 8 reports results from estimating the system using GMM. The first column (Equation I) presents the test of outrage on compensation akin to Table 7, but with additional Equation II control variables . The results are very similar to those in Table 7.

In columns (1) to (4) we estimate the effect of outrage on returns through Log Compensation. We refer to this variable as Outrage-Predicted Log Compensation. The outcome variable is net returns for: the entire portfolio (column 1), alternatives (column 2), public equities (column 3) and fixed income (column 4).

We find that log compensation explained by outrage has a positive and significant effect on portfolio net returns (column 1). The coefficient is a positive 0.00635. Because this return is already net of the benchmark, this portfolio return sensitivity to Outrage-Predicted Log Compensation is likely due to within-asset class outrage-performance sensitivity rather than to the asset class allocations. In columns (2) to (4) we look explicitly at within-asset-class sensitivities. We replace portfolio net returns with net returns in alternatives (2), public equities (3) and fixed income (4). Results for the risky asset classes are very consistent with the model. Outrage-Predicted Log Compensation positively and significantly predicts net returns in alternatives (coefficient of 0.0209) and equities (0.00689). We find no effect for fixed income. We also note that Political Chair has a negative and significant impact in alternatives, consistent with the results in Andonov, Hochberg and Rauh (2017). We discuss the Political Chair result in greater length and related evidence in section 5.5 below.

Economic Impact of Outrage-constrained Compensation

To understand the economic impact of outrage on portfolio returns, we first consider a one standard deviation change in either Municipal Workers, Budget Civil Servants or Regional Outrage (iteratively). Table 9 presents each of these pass-through, economic magnitude results. A one standard deviation lower fraction of Municipal Workers, working through \$76,033 more compensation, results in 0.06% (6 basis points (bps)) greater net returns. Likewise, a one standard deviation lower fraction of Budget Civil Servants, implying \$107,627 more compensation, results in 8.5 bps greater returns. A 10% increase in Regional Income predicts \$63,221 greater compensation and 5 bps greater returns. In summary, all else equal, if a pension fund could unwind just one determinant of outrage by one standard deviation (mimicking about a 10% change in local income or 1-2 trustee member change in the board composition) and hire accordingly a more experienced manager for approximately \$90,000 more in pay, the pension would reap a benefit in returns of 6.5 basis points per year (averaged over 5, 8.5 and 6 bps). Panel B of Table 9 speaks to the dollar-value materiality of these effects. Evaluated at the average estimation sample fund of \$102 billion in AUM, a one standard deviation reduction in one of the outrage determinants would improve returns for that pension of $51-86 million per year. On the right side of the table we also evaluate the economic magnitude for a representative sample of pension plans. Recall that our estimation sample focuses on large pension plans for reasons of data observability. A one standard deviation reduction in one of the outrage determinants would improve returns for an average representative sample pension by $22-38 million per year. The smallest fund evaluated is the 25th percentile fund of the representative sample. For this small fund, the benefit from unwinding one factor determining outrage is $3.8 to $6.4 million in additional AUM per year. We also consider the impact of two policy proposals to improve governance of pension plans to reduce the threat of outrage. First, we simulate the impact of reducing political appointees on the board, an idea raised by Romano (1993), by eliminating all Budget Civil Servants. Then we consider a move towards a skills-based board that also removes Municipal Workers from the board. Reducing political appointees in this way would increase expected compensation by $257,000, and a skills-based board would increase compensation by $324,000. The respective increases in returns would be 20.2 bps and 25.5 bps per year, translating into an annual increase in incremental value-added for the average representative sample pension of $206- $260 million. As we mention in the introduction, the importance of unwinding outrage might be especially important because funds that are most constrained by outrage are lower income area funds where local economy spillovers from poor pension asset performance might be most severe.

Robustness

Higher net returns do not necessarily reflect a higher Sharpe ratio if the net return performance arises from taking on increased risk. We address this possibility in two ways. First, in Appendix Table AT2 we rerun Table 8 tests while including the weights allocated to the sub-asset class categories. These additional levels of investment focus are: (in alternatives) real estate, hedge funds, private equity, and infrastructure; (in equities), domestic versus non-domestic public equities; and (in fixed income), cash versus bonds. We take bonds as the omitted category. We find that that inclusion of the portfolio weights in our system does not change our outrage results.

In a more direct test, in Table 10, we study the effect of outrage on realized tracking error. Tracking error can result from higher risk strategies, but also from lower skill. Yet, in the presence of our results from Table 8, our concern would be that any lower returns due to outrage could be attributed to lower risk strategies, which would be associated with lower within-asset class tracking error.

In this cross-sectional setup, fund-level tracking error is the dependent variable in Equation II, and everything else in the system of equations is the same as before, including the compensation equation. With the fewer number of observations, we drop the agency variables without power in the Table 8 estimations. Columns (1) to (4) show that Outrage-Predicted Log Compensation has no statistically significant relationship with realized tracking error for any of the asset classes. This counters the concern that our findings from Table 8 result from increased within asset-class risk.

A final robustness concern with our main Table 8 result is the exogeneity of the trustee occupations. One might be concerns that we are picking up the inverse of the politicization result of Hochberg and Rauh (2013) if the lack of politicized board members mechanically implies more teachers, workers, and budget civil servants. Thus, also in Appendix Table AT2, we take our outrage trustee occupation counts and divide by the denominator of the total number of non-political trustees, calculating a fraction relative to non-political trustees. As Table AT2 shows, our results are, if anything, stronger.

Delegation Results

One possible mechanism driving the underperformance of outrage-constrained pension funds is that plans with less skilled managers delegate larger fractions of their portfolios to external institutions, resulting in greater payment of intermediation fees. We investigate this possibility in Table 11 by estimating our two-equation system but this time using the fraction of assets managed via delegation (in each asset class) as the outcome variable of Equation II . We use the same asset classes defined in the previous sections, with the only difference being that for alternatives we do not include hedge funds, as they are delegation institutions. Given that our delegation fraction is a number between 0 and 1, we estimate our model using a Tobit specification on the second stage.

We find a negative and significant coefficient on log compensation explained by outrage in all Equation II columns. This shows that funds that are able to avoid outrage constraints on compensation are more likely to reduce their use of delegation and manage assets in-house. The economic impact is meaningful: a reduction of one standard deviation in Municipal Workers or Budget Civil Servants results in 6% or 8%, respectively, fewer delegation percentage points. An increase of 10% in Regional Outrage is followed by a reduction of around 4% in the fraction of assets managed via delegation. The effect of outrage constraints on delegation is largest for asset classes with more risk (i.e., more skill required) and for asset classes with higher mean delegation per Table 5.

Do Outrage Pay Constraints affect Asset Allocations?

Table 11 explores the possibility that outrage pay constraints affect funds' asset allocation according to predictions from our theory; in particular, higher outrage may lead to lower risk-taking, working through reduced compensation and reduced managerial skills. Because asset class weights are jointly determined, we report two sets of standard errors a fund-clustered standard error and a robust standard error under the seemingly-unrelated-regression assumption (SUR).

The results indicate that funds with compensation not constrained by outrage would exhibit higher exposures to the riskiest asset class (alternatives, column 1) in lieu of public equities (public equities). Inside our model, such an effect may arise with the hiring of a skilled manager that can extract a larger fraction of the premia in riskier asset classes. What is perhaps a bit inconsistent with our theory is that eschewing of public equities for alternatives also implies that the pension increases exposure to fixed income (column 3).

Other Political Agency Costs

The results in 5.2 speak to the impact of the human capital channel on allocation and performance in public pension funds. Political influence may come through other channels of distortions arising from politicians payoffs to local investment and distortions arising from underfunding.

Our key variable for exploring any distortions from politicians payoffs is Political Chair. Pay-to-play arrangements of political funds may cause public pension funds to invest in political assets (e.g. local assets) to provide private political benefits for the board chair. The key variable to predict risking-up of portfolios due to pressures from liability obligations is Underfunded Index. These variables are introduced in Tables 7-11 in the compensation regressions (Equation I) as well as in the outcome regressions (Equation II). We include the variables in both equations because we believe these political variables will fail the exogeneity condition, with Political Chair and Underfunding also being directly correlated with outcomes.

In Table 7, we find that Political Chair significantly explains variation in compensation. In particular, a 0.586 greater likelihood of Political Chair implies a $94,209 lower manager compensation. Using the language of our model, a large reward for political investments L leads to a manager with low skill (s). In Table 8, when we include outrage-predicted log compensation, we find a negative and significant direct impact of Political Chair on returns in the portfolio as a whole, being driven by lower returns in alternatives and equities. Table 9 reports that a higher likelihood of having a Political Chair implies 0.212% lower returns. In Table 10, we find that when we include outrage-predicted log compensation, there is no direct impact of Political Chair on realized tracking error at the portfolio level or for alternatives, but Political Chair is associated with lower realized tracking error for public equities and fixed income. Table 11 reports that the only significant direct impact of Political Chair on asset allocation is to weakly increase the allocation to fixed income.

These results suggest that politicization through the board chair reduces risk in areas where politicians are less likely to derive private benefits (outside of alternatives). In alternatives, which includes private equity where prior research has suggested and demonstrated pay-to-play is likely and reduces performance, we find additional negative effects of politicization through the board chair on performance. As a whole, the results support a pay-to-play interpretation consistent with the research of Andonov, Hochberg and Rauh (2017).

Finally, we turn our attention to the role of Underfunded Index. The only significant impact is (weakly) on asset class weights. Consistent with prior papers, notably, Andonov, Bauer and Cremers (2017), we find that Underfunded Index predicts lower allocations to fixed income.

2.6 Conclusion

The paper introduces a model in which trustees of public pension funds incorporate the threat of private costs arising from outrage over high compensation. This concern leads to an equilibrium with trustees hiring investment managers with lower skills, which in turn creates distortions in portfolio allocation and weaker performance in the risky asset classes. We test these predictions using a hand-collected global panel data set that includes information on investment manager compensation and structural features of boards and trustees that predict outrage. We find that outrage pay constraints on compensation induced by public pension funds governance structures impact fund performance and hence beneficiary welfare. For an average fund, our estimates suggest that if that fund were to relax outrage, with a cost of approximately $90,000, it would benefit by producing additional benefits of $22 to $38 million in annual value-add. More extensive reforms to reduce political appointees or adopt an independent skills-based board, would be associated with higher compensation costs and a higher annual increase in incremental value-added of $90- $114 million. The result is particularly important in areas where finance salaries are much larger than the average income of local residents. Such areas may be more readily prone to outrage, but also are areas where the local wage earners have little slack to support faltering pension systems. Of course it is natural to ask if it possible for funds to change outrage constraints. Funds cannot change the possibility that their disclosures may garner media attention to high finance salaries, resulting in public outrage. They also cannot also change the fact that beneficiaries and governments have a strong interest in trustee choice, as they are required to pay into the plan and suffer in case the fund is unable to fulfill its promises. Yet, although it is beyond our scope to consider all possible ways to insulate the board from outrage pressures, we have a few ideas. First, the most obvious step is to educate both beneficiaries and governments of the costs of exposure to outrage for plan performance that we document in this paper. Second, a lasting impact could emerge from a refocus of trustee appointment procedures toward an independent skills-based board. In an independent skills-based board, the government and beneficiaries continue to select trustees, but they require that the board members, either individually or as a group, have certain skills. Those skills are often financial skills, rather than political skills, or representing the industry, effectively insulating these board members from outrage pressures. Likewise, pensions might move from a representative board chair to one based on skills and chosen by the other board members. Third, modifying or clarifying risk and profit-sharing arrangements so that beneficiaries expected benefits become more closely tied to the performance of the fund could increase salience to the importance of quality investment management. One possible reform that we do not advocate is reducing the transparency of compensation arrangements. While this is a crude way to insulate board members from outrage pressures, it is likely to be imperfect. Board members likely fear that compensation arrangements will be eventually released or leaked, leading to much of the same behavior. Therefore, as long as the board members remain exposed to outrage concerns, the same problems will emerge. A policy of transparency combined with a skills-based board should insulate trustees from outrage concerns, much as is the case with boards of traded companies. The added advantage of transparency is that it also reduces the likelihood of pay-to-play arrangements and other political frictions.

Figure 2.1: Compensation of Investment Manager by Prior Profession

Graphed are the distribution of investment manager compensation for each category of prior professions of the managers. The box plot displays the mean (box center line) as well as the first (box edges) and second (stem edges) standard deviations. The dashed (red) line indicates the overall sample 25th , 50th, and 75th percentiles. The distribution of the sample is as follows (also reported in Table 6, along with the more detailed titles of the professions under the categories): Pension Finance (4.9%), Pension Non-Finance (18.0%), Private Professional (31.1%), Civil Servant Finance (29.5%), and Civil Servant Non-Finance (16.4%).

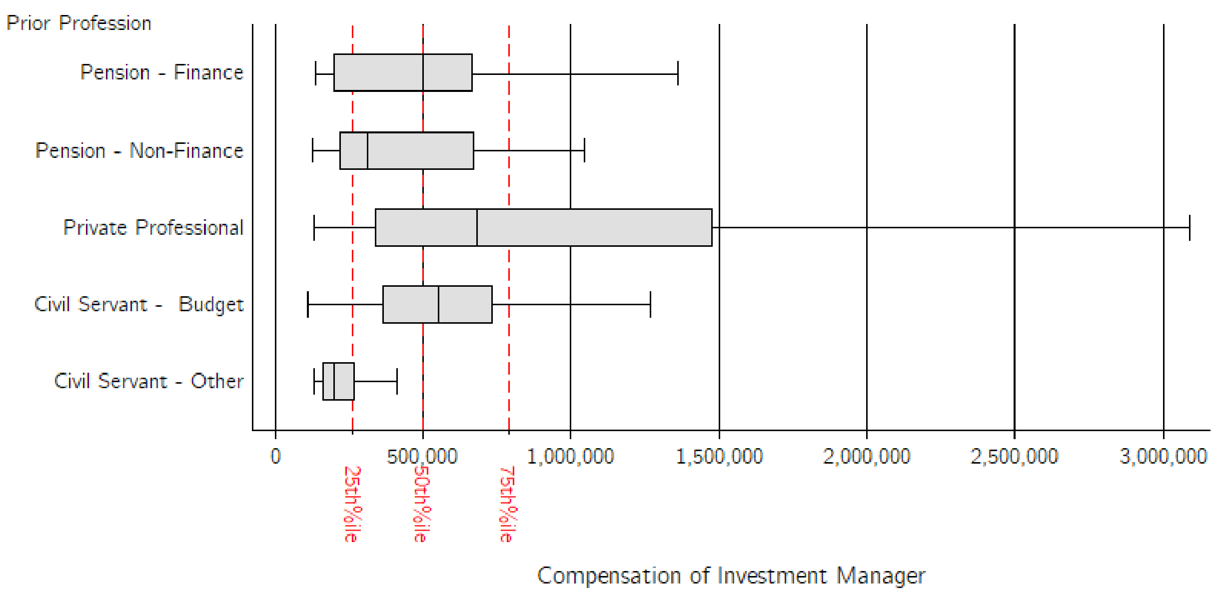


Figure 2.2: Compensation by Tertitle of Trustee Occupation-Outrage Variables

Plotted are the mean (blue/darker bars) and median (green/lighter bars) manager compensation by tertitles of the trustee occupation variables which are our proxies for outrage. The variables (from left to right plotted) are the percentage Municipal Workers, the percentage Teachers and the percentage Budget Civil Servants.

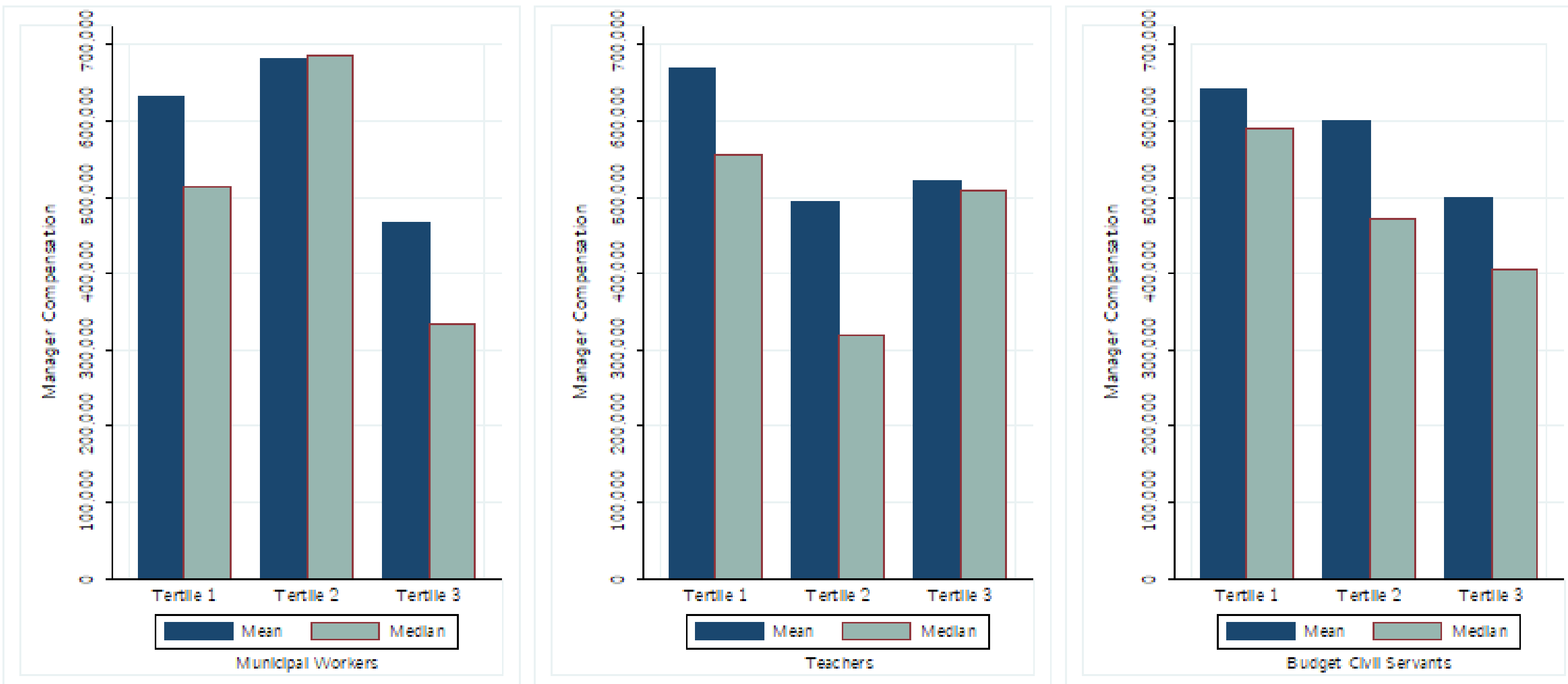


Figure 2.3: Compensation by Tertitle of Reference Wage-Outrage Variables

Plotted are the mean (blue/darker bars) and median (green/lighter bars) manager compensation by tertitles of reference income variables which are our proxies for outrage. The variables (from left to right plotted) are the Regional Income and the pension Worker Wages.

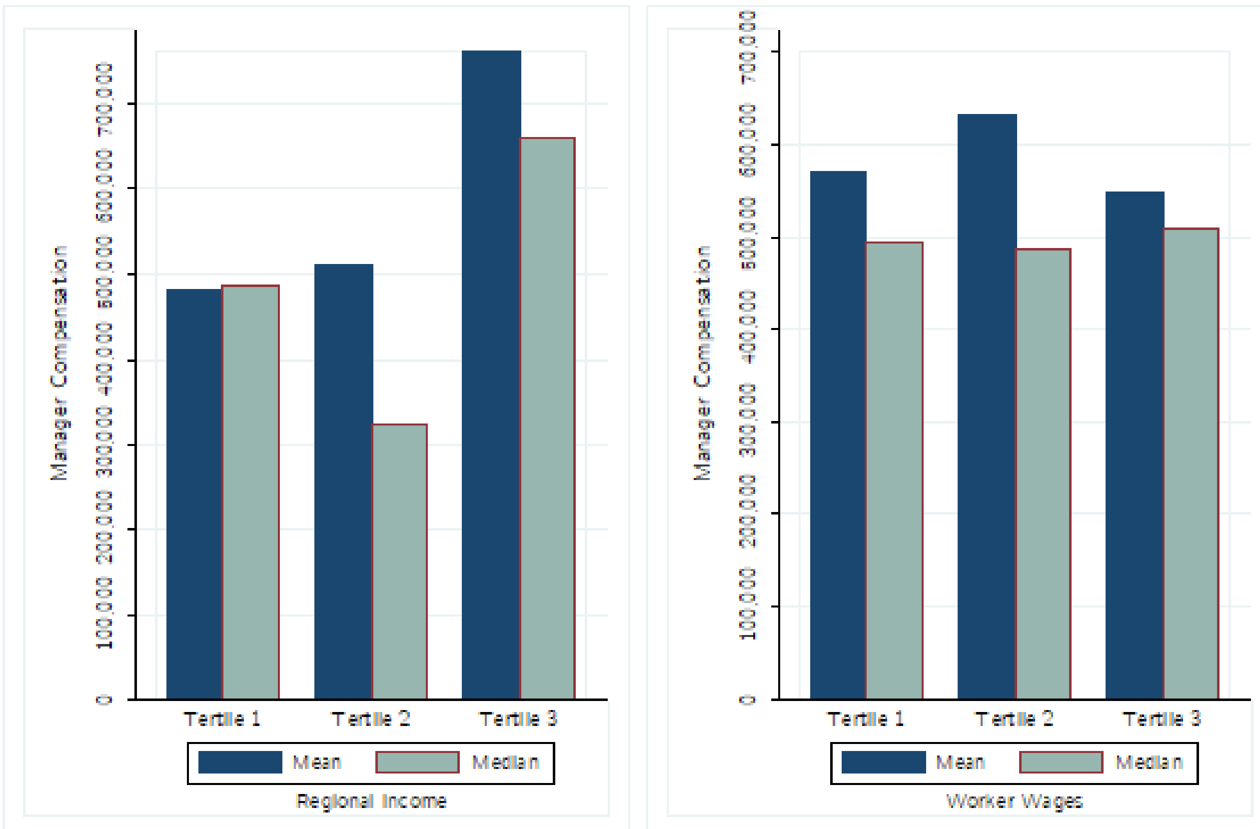


Table 2.1: Comparative Statics: Political Agency Variables Role

Panel A: Effect of a Binding Outrage Constraint

Variable	Model Notation	Predicted Change to Row Variable with: Δ Outrage
Manager skill	Δs	< 0
Allocations		
Weight on MV security	$\Delta(w_{MV})$	< 0
Weight on political asset	$\Delta(w_P)$	< 0
Weight on fixed income	$\Delta(1 - w_P - w_{MV})$	> 0
Weight on all risky	$\Delta(w_P + w_{MV})$	< 0
Performance		
E[return on MV security]	$\Delta(R_{MV})$	< 0
E[return on political asset]	$\Delta(R_P)$	< 0
E[portfolio return]	$\Delta(R)$	< 0

Panel B: Effect of Other Political Agency Costs

Variable	Model Notation	Partial Derivative with Respect to: $\partial\kappa$	$\partial\theta$
Manager skill	s	< 0	> 0
Allocations			
Weight on MV security	$\partial(w_{MV})$	< 0	> 0
Weight on political asset	$\partial(w_P)$	> 0	?
Weight on fixed income	$\partial(1 - wP - w_{MV})$	?	< 0
Weight on all risky	$\partial(w_P + w_{MV})$	?	> 0
Performance			
E[return on MV security]	$\partial(R_{MV})$	< 0	> 0
E[return on political asset]	$\partial(R_P)$	< 0	> 0
E[portfolio return]	$\partial(R)$	< 0	> 0

Note: This table lays out model predictions. showing the comparative statics of how manager skill, portfolio choice, and returns change in the model with changes in political agency variables. The political agency issue of outrage is considered in Panel A. Because outrage is a binding-or-not constraint, the comparaitive statics reflect a discret change from not binding to binding. In panel B, the political agency issues of private benefits of political assets and the underfunding are considered. In Panel B, the comparative staics show the partical derivatives of a change in either manager skill, allocations and performance with respect to a change in agency – private benefits of political asset investing (κ) and the board preference for risk, driven by pension liabilities (θ). The right column relates the prediction to the table of reference for empirical results.

Table 2.2: Variable Definitions

Variable	Definition	Source
	Compensation, Portfolio Choice, and Performance Variables	
Investment Manager Compensation	The maximum compensation of the fund's investment managers, including CEO and CIO.	Hand-collected from annual reports, public filings, newspapers, and Freedom of Information requests.
Portfolio Allocation	Portfolio weights in each of three asset class – alternatives (real estate, private equity, hedge funds, infrastrcuture), public equity, and fixed income. Expressed as a percentage of the total.	Center for Retirement Research (CRR), CEM Benchmarking and annual reports.
Return	Realized returns in each asset class and for the overall portfolio.	Center for Retirement Research (CRR), CEM Benchmarking and annual reports.
Benchmark Return	CEM requires benchmarks to be chosen by pension trustees rather than being self-selected by asset managers. Most funds report multiple indices and weights. A visual inspection of this information indicates the benchmarks capture dimensions of risk differences across and within asset classes.	CEM Benchmarking
Tracking Error	A single observation by fund for each asset class and the portfolio, calculated as the time-series average of the squared residuals from a regression of the pension fund returns on the benchmark returns, with no constant.	Center for Retirement Research (CRR), CEM Benchmarking and annual reports.
Portfolio Delegation	Fraction of assets managed via delegation in each asset class.	CEM Benchmarking.
	Political Agency Variables	
Municipal Workers	The fraction of trustees that are workers providing basic services to city residents, usually through city government.	From annual reports. Professional designation based on biographies and web sources such as LinkedIn.
Teacher	The fraction of trustees that are workers providing basic services to teachers or education administrators.	From annual reports. Professional designation based on biographies and web sources such as LinkedIn.
Budget Civil Servant	The fraction of trustees that are civil servant in finance service to the government.	From annual reports. Professional designation based on biographies and web sources such as LinkedIn.
Regional Income	Logarithm of the local household income within the smallest region available (MSAs for the US).	Regional income reported by National statistical offices (Census Bureau in the US).
Worker Wage	Logarithm of the average wage of the constituents of the pension fund.	Hand-collected from annual reports. If not reported, we estimate based on working employee contributions and reported contribution rates as a percentage of salary.
Political Board	A dummy equal to one if the chair is appointed by government executives or ministries or serves in the role ex officion because of his or her executive government position.	Collected from pension fund charters and annual reports.

Note: This Table reports the definitions and the data sources for the main variables used in this paper.

Table 2.3: Pension Fund Profile Statistics

Panel A: Full Sample

		Assets under Management ($billion)					Gross Portfolio Returns				
	Number of funds	Fund-Year Observations	Mean	25th Percentile	Median	75th Percentile	Fund-Year Observations	Mean	25th Percentile	Median	75th Percentile
Canada	16	210	37.02	11.45	17.04	59.90	210	0.0548	0.0012	0.0672	0.1160
Europe	39	333	122.70	8.45	17.76	71.33	302	0.0173	0.0004	0.0018	0.0268
Oceania	17	163	15.11	6.61	12.84	19.13	160	0.0312	0.0001	0.0018	0.0960
United States	92	1150	27.65	6.88	12.81	32.03	1130	0.0498	0.0004	0.0323	0.1235
Total	164	1856	44.66	7.59	13.70	35.55	1802	0.0433	0.0004	0.0195	0.1098

Panel B: Sample with Compensation & Trustee Data

		Assets under Management ($billion)					Gross Portfolio Returns				
	Number of funds	Fund-Year Observations	Mean	25th Percentile	Median	75th Percentile	Fund-Year Observations	Mean	25th Percentile	Median	75th Percentile
Canada	10	97	49.68	13.14	33.78	81.30	97	0.0589	0.0009	0.0857	0.1267
Europe	17	115	283.42	19.41	70.39	322.17	115	0.0245	0.0009	0.0114	0.0360
Oceania	11	55	21.48	11.41	16.81	27.74	55	0.0385	-0.0010	0.0212	0.1053
United States	73	196	44.07	11.72	26.00	59.21	196	0.0451	-0.0290	0.0518	0.1331
Total	111	463	102.01	11.99	29.51	72.59	463	0.0421	-0.0005	0.0330	0.1126

Note: This Table reports the assets under managementand portfolio returns statistics by region of the penion fund. Panel A presents these statistics for the full sample of funds in our sample, and Panel B, for the pension funds for which we have manager compensation of trustee profile data.

Table 2.4: Performance and Allocation Statistics

	Count	Mean	Standard Deviation	25th Percentile	Median	75th Percentile
Panel A: Allocations						
Weights: Full Sample						
Alternatives	251	0.2290	0.1750	0.1250	0.1970	0.2730
Public Equities	304	0.5980	0.1840	0.4850	0.5710	0.6690
Fixed Income	253	0.3230	0.1210	0.2500	0.3050	0.3680
Weights: Sample restricted to having data on all weights						
Alternatives	204	0.1910	0.0960	0.1170	0.1860	0.2520
Public Equities	204	0.5130	0.1060	0.4420	0.5250	0.5830
Fixed Income	204	0.2960	0.0750	0.2430	0.2970	0.3500
Delegation Fraction						
Alternatives	214	0.7470	0.3270	0.4840	0.9900	1.0000
Public Equities	190	0.7340	0.3600	0.3860	1.0000	1.0000
Fixed Income	180	0.5000	0.4680	0.0000	0.4880	1.0000
Panel B: Performance						
Gross Returns						
Alternatives	355	0.0610	0.1190	0.0020	0.0750	0.1350
Public Equities	367	0.0530	0.2060	-0.1070	0.1170	0.2060
Fixed Income	337	0.0610	0.0490	0.0340	0.0550	0.0800
Portfolio	463	0.0420	0.0960	0.0000	0.0330	0.1130
Net Returns						
Alternatives	251	-0.0080	0.1010	-0.0530	-0.0040	0.0460
Equities	304	0.0050	0.0200	-0.0040	0.0030	0.0130
Fixed Income	253	0.0050	0.0310	-0.0030	0.0030	0.0160
Portfolio	351	-0.0030	0.0540	-0.0110	0.0010	0.0140
Tracking Error Realized						
Alternatives	70	0.0689	0.0728	0.0275	0.0552	0.0834
Equities	96	0.0384	0.0536	0.0120	0.0191	0.0345
Fixed Income	92	0.0212	0.0160	0.0087	0.0181	0.0285
Portfolio	110	0.0299	0.0232	0.0137	0.0238	0.0449

Note: this Table reports summary statistics of the portfolio weights and performance, at the portfolio level and by asset classes. Asset classes are: (i) alternatives, defined as hedge funds, real estate, private equity, and infrastructure, (ii) public equities, and (iii) fixed income. In Panel A, we present the weights in the main estimation sample (that with compensation and trustee data) plus the sample where we observe all weights such that the weights sum to unity. Also in Panel A are the fractions of each asset class delegated to outside management. Panel B reports performance in three metrics – gross returns, net returns over the CEM benchmark, and realized tracking error. The realized tracking error is calculated in the data relative to the benchmark return; thus there is only one observation per pension fund.

Table 2.5: Compensation, Trustee Occupation, Wages and Other Agency Statistics

Panel A: Statistics	Count	Mean	Standard Deviation	25th Percentile	Median	75th Percentile
Manager Compensation						
Manager Compensation ($)	463	807416	1018136	292328	537197	819979
Log Manager Compensation	463.00	13.20	0.83	12.59	13.19	13.62
Outrage: Reference Wages						
Worker Wages	463	47811	15197	38763	45345	55066
Log Worker Wages	463.00	10.73	0.30	10.57	10.72	10.92
Regional Income	463	55434	17955	40873	50127	68228
Log Regional Income	436.00	10.86	0.32	10.62	10.78	11.11
Outrage: Non-Political Trustee Occucations						
Municipal Workers (% Trustees)	463.000	0.053	0.087	0.000	0.000	0.100
Teachers (% Trustees)	463.000	0.109	0.167	0.000	0.077	0.133
Budget Civil Servants (% Trustees)	463.000	0.102	0.144	0.000	0.083	0.154
Other Agency Variables						
Political Chair	463.000	0.514	0.586	0.000	0.364	1.000
Underfunded Index	463.000	0.171	1.303	-0.144	0.000	0.203

Panel B: Correlations	Compen-sation	Municipal Workers	Budg. Civil Servants	Teachers	Worker Wages	Local Income	Political Chair
Municipal Workers	-0.092						
Budget Civil Servants	-0.150	-0.198					
Teachers	-0.226	-0.111	-0.114				
Worker Wages	0.061	0.106	0.112	0.015			
Local Income	0.364	0.022	-0.176	-0.193	0.250		
Political Chair	-0.120	-0.133	-0.012	0.023	0.000	-0.080	
Underfunded Index	-0.101	0.075	0.082	-0.069	0.052	0.027	-0.030

Note: Panel A reports the summary statistics, and Panel B reports the correlations of the main variables characterizing the governance of pension funds in our sample. Manager Compensation is defined as the highest paid executive (CEO or CIO) for the public fund. Municipal Workers is the percent of the board whose career is in the municipal labor force, defined as police, fire department, hospitals, libraries, and other non-civil servant positions. Budget Civil Servant is the percent of the board whose background is in public sector financial positions (e.g., city controllers, auditors, etc.). Teachers is the percent of the pension board who are teachers. Political Chair is a dummy taking value 1 if the chair is appointed by the executives or ministers of the government. Underfunded Index is an index constructed by taking the mean across the standardized value of (1- the funded ratio) and age following Rauh (2008). The two outrage income measures – Worker Wages and Local Income – are, receptively the average wages of workers and the municipal income.

Table 2.6: Professions of Investment Managers and Trustees

Panel A: Investment Managers' Professions

Occupation	**Description**	**Professions Represented**	%
	Prior Pension Executives		
Pension - Investment Executive	Investment manager from another pension fund	Director of Investment, CEO, CIO	4.9%
Pension - Other Executive	Other executive position in another pension fund	Assistant General Counsel, Assistant Executive Director, Deputy Executive Director, Chief of Staff, COO	18.0%
	Prior Private Firm Finance Professionals or Executives		
Private Firm Professional	Financial position from privately firm	CEO, CIO, Director, Managing Partner, Accountant, Actuary, Auditor, Consultant, CRO	31.1%
	Civil Servants		
Civil Servant (Budget)	Civil servant with financial experience	Treasurer, Auditor, Accountant, Controller, Budget Officer, Finance Director, Public Institution Professor	29.5%
Civil Servant (Non-Budget)	Civil servant without financial experience	City Council CEO, City Manager, Executive Director, Department of Correction Administrator, Deputy Chief of Staff, Director, Executive Commissioner, Natural Resource Advisor, Teacher, Senator	16.4%

Panel B: Trustees' Professions

Occupation	Description	Professions Represented	%
	Civil Servants		
Politician	Includes any representative or elected official of municipal, state or federal government	Senator, House Representative, Mayor, Governor, Lieutenant Governor, Secretary of State, Attorney General, Assembly Speaker, State Representative, Secretary, Minister, Borough President, City Manager, Assistant Deputy Minister, Deputy Governor, Premier Deputy Chief of Staff, Deputy Minister, , City Council, County Commissioner, Deputy City Manager, Deputy General Counsel,	6.4%
Budget Civil Servant	Civil servant with financial experience	Treasurer, Auditor, Accountant, Controller, Budget Officer, State Finance Director	34.4%
Other Civil Servant	Civil servant without financial experience	Judge, Prosecutor, Clerk, Commissioner, Assistant Commissioner, Professor, Dean	13.7%
	Non-Civil Servants		
Teacher	Teachers	Teachers	14.7%
Municipal Worker	Workers providing services to city residents, union labor	Police Officer, Fire Officer, Jail Worker, Railway , Steel , Construction, Electrician, Mail Employee, Librarian, Miner, Bus Driver, Chimney Sweep, Food Worker, Manufacturing Worker, Telecommunications	7.7%
Professionals	Local private sector professionals and NGO executives	Financial Sector Expert, Doctor, Nurse, Dentist, Private Firm CEO, CIO, Chairman, Pharmacist, Journalist, Media Professional, Architect, NGO Chairman, Owner of Private Firm	23.1%

Note: This table reports the immediate prior profession of investment managers (Panel A) and the current professions of trustees (Panel B). The data are collapsed to the cross section of public funds. All data are hand collected.

Table 2.7: Effect of Outrage and Political Agency on Manager Compensation

Panel A: Estimates

	Dependent Variable: Log Compensation				
	(1)	(2)	(3)	(4)	(5)
Municipal Workers		-0.604			-1.082*
		[0.546]			[0.611]
Teachers		-0.619**			-0.405
		[0.293]			[0.324]
Budget Civil Servants		-1.401***			-0.925**
		[0.284]			[0.374]
Log Regional Income			0.923***		0.783***
			[0.181]		[0.193]
Log Worker Wages			0.618**		0.690**
			[0.285]		[0.293]
Political Chair				-0.212**	-0.199**
				[0.0985]	[0.0971]
Underfunding Index (lag)				0.0204	0.0418*
				[0.0288]	[0.0238]
Log Size (lag)	0.284***	0.279***	0.154*	0.286***	0.164*
	[0.0842]	[0.0830]	[0.0828]	[0.0843]	[0.0836]
Year Fixed Effects	Y	Y	Y	Y	Y
Observations	453	453	426	453	426
Number of Funds	110	110	110	110	110
R-Squared	0.0365	0.115	0.106	0.0498	0.153

Panel B: Economic Magnitude

Change Evaluated		$ Impact on Compensation	Percentage Change
1 standard deviation change =	0.087 higher fraction of Municipal Workers	-76033	-9%
1 standard deviation change =	0.144 higher fraction of Budget Civil Servants	-107627	-13%
10% percentage change =	4781 higher Regional Income ($)	63221	8%
10% percentage change =	5543 higher Worker Wages ($)	55712	7%
1 standard deviation change =	0.586 greater likelihood of Political Chair	-94209	-12%
1 standard deviation change =	1.303 higher Underfunding Index	43982	5%

Note: The dependent variable is the log compensation of the investment manager. Municipal Workers is the percent of the board whose career is in the municipal labor force, defined as police, fire department, hospitals, libraries, and other non-civil servant positions. Teachers is the percent of the pension board who are teachers. Budget Civil Servant is the percent of the board whose background is in public sector financial positions (e.g., city controllers, auditors, etc.). Political Chair is a dummy taking value 1 if the chair is appointed by the executives or ministers of the governmentor is ex officio designated as chair as an executives or ministers of the government. Underfunded Index is an index constructed by taking the mean across the standardized value of (1- the funded ratio) and age following Rauh (2008). Worker Wages and Regional Income are the outrage reference wages, equal to mean pension workers wages and meadian local area incomes. Log Size is the log of the fund AUM. All money variables are in 2010 USD. Panel A present Estimation is OLS with year fixed effects. Panel B presents the economic magnitude with the change induced as noted, choosing a half standard deviation in situations where the cross-sectional changes would be large for a time series application by a pension fund. Standard errors are clustered at the fund level. *** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1.

Table 2.8: Effect of Outrage on Returns

		(1)	(2)	(3)	(4)
	Equation I:	Equation II: Net Returns			
	Log Compensation	Portfolio	Alternatives	Public Equities	Fixed Income
Outrage-Predicted		0.00635**	0.0209*	0.00689*	-0.00441
Log Compensation		[0.00291]	[0.0111]	[0.00400]	[0.00370]
Municipal Workers	-0.997**				
	[0.470]				
Teachers	-0.217				
	[0.252]				
Budget Civil Servants	-1.163***				
	[0.279]				
Log Regional Income	1.034***				
	[0.154]				
Log Worker Wages	-0.142				
	[0.156]				
Political Chair	-0.0978	-0.00362**	-0.0155**	-0.00353*	-0.000123
	[0.0705]	[0.00143]	[0.00777]	[0.00187]	[0.00219]
Underfunding (lag)	0.023	0.000736	-0.00117	-0.000458	0.00297
	[0.0409]	[0.00133]	[0.00544]	[0.00179]	[0.00199]
Log Size (lag)	0.304***	-0.00314***	0.000651	-0.00409**	-0.000433
	[0.0393]	[0.00117]	[0.00552]	[0.00161]	[0.00178]
Observations	303	303	243	285	243
Number of Funds	89	89	71	86	80
Cragg-Donald F-stat	20.31				
F-Stat p-value	0				

Note: Reported in columns (1)-(4) are estimates from a GMM system of two equations. The dependent variable in numbered columns (denoted above Equation II) is the net return over the asset-class benchmark. The far left column presents (Equation I) estimates of the effect of political agency on compensation for the sample in column (1). A similar estimate (unreported) is used for each of columns (2)-(4). In columns (1)-(4), the log compensation variable is the outrage-predicted compensation, from Equation I (the left column). Municipal Worker, Teachers, and Budget Civil Servant are the trustee composition outrage variables. Worker Wages and Regional Income are the outrage reference wages, equal to mean pension workers wages and meadian local area income. Political Chair is equal to one for funds whose chair is appointed by the government. Underfunded Index is an index constructed by taking the mean across the standardized value of (1- the funded ratio) and age following Rauh (2008). Log Size is the log of the lagged fund AUM. Standard errors are clustered at the fund level. The number of funds per estimation is indicated below the number of observations. The Cragg-Donald F-statistic, and the p-value of ints significance, is included as a test of weak identification. *** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1.

Table 2.9: Effect of Outrage on Returns

Panel A: Return Implications

Equation I Change Evaluated	Working through Equation II Effect		Resulting Change in Returns
1 sd ↓ in Municipal Workers ⇒	\$76033	change in Compensation ⇒	0.060%
1 sd ↓ in Budget Servants ⇒	\$107627	change in Compensation ⇒	0.085%
10% ↑ in Regional Income ⇒	\$63221	change in Compensation ⇒	0.050%
	1 s.d. ↑ in Political Chair ⇒		-0.212%

Panel B: Assets under Management Implications

	Annual AUM Change (\$Million) for Different Size Pension Funds: Estimation Sample (Table 3, Panel B)				Representative Sample (Table 3, Panel A)			
	Mean	25th Percentile	Median	75th Percentile	Mean	25th Percentile	Median	75th Percentile
1 sd ↓ in Mun. Workers ⇒	\$61.0	\$7.2	\$17.6	\$43.4	\$26.7	\$4.5	\$8.2	\$21.3
1 sd ↓ in Budget Servants ⇒	\$86.3	\$10.1	\$25.0	\$61.4	\$37.8	\$6.4	\$11.6	\$30.1
10% ↑ in Regional Income ⇒	\$50.7	\$6.0	\$14.7	\$36.1	\$22.2	\$3.8	\$6.8	\$17.7

Note: Panel A simulates the economic magnitudes associated with reducing the outrage threat, using estimates from Table 8. In the system, the Equation I outrage variables affect compensation, which in turn affects manager compensation. We present this pass-through effect of a change in the outrage variables in Equation I to the return performance implication of column (1), Table 8. We simulate eliminating Budget Civil Servants by a 2 s.d. change in Budget Civil servants (2 s.d.=0.28, % board Budget Civil Servants=0.34 from Table 6. We estimate a move to skills-based board as the combination of eliminating Budget Civil Servants and eliminating Municipal Workers (1 s.d.=0.087, % board that is Municipal Worker=0.077).The interpretation is not necessarily that the outrage variable can be changed, but the extent to which, the political agency that allows outrage to happen could be unwound. Returns are expressed in annual performance. In the final row, we show the effect in Equation II of a change in political chair on returns, following Andonov et al (2017). Building off the calculations in Panel A, Panel B contains the implied changes in AUM per year for a pension fund evaluated at different points in the pension fund size distribution. The numbers presented can be interpreted as the inference to the following question: how much in assets' dollar returns might a pension fund gain by reducing outrage by the indicated channel? We repeat the exercise for two pension fund samples (the estimation sample and the representative sample) represented in Table 3, Panels A and B. On the left are the pension funds in our estimation sample, which is biased toward larger funds because of our need to have observability in compensation.

Table 2.10: Effect of Outrage on Realized Tracking Error

		(1)	(2)	(3)	(4)
	Equation I:	Equation II: Tracking Error			
	Log Compensation	Portfolio	Alternatives	Public Equities	Fixed Income
Outrage-Predicted		0.00843	-0.0303	0.00179	-0.00626
Log Compensation		[0.00731]	[0.0275]	[0.0216]	[0.00509]
Municipal Workers	-1.029**				
	[0.507]				
Budget Civil Servants	-0.637*				
	[0.353]				
Log Regional Income	0.519**				
	[0.202]				
Political Chair	-0.0884	0.00474	-0.0135	-0.0174***	-0.00457*
	[0.0967]	[0.00380]	[0.0146]	[0.00674]	[0.00245]
Underfunded Index	-0.0227	0.00205	0.000394	0.00804	0.000375
	[0.0500]	[0.00238]	[0.00678]	[0.00544]	[0.00191]
Log Fund Size	0.0217	0.000484	-0.00298	0.0105**	-0.00208
	[0.0551]	[0.00186]	[0.00861]	[0.00455]	[0.00196]
Weights					
Private Equity	0.242	0.023	0.899**		
	[1.612]	[0.0506]	[0.373]		
Real Estate	-0.243	-0.00857	-0.141*		
	[0.864]	[0.0300]	[0.0745]		
Hedge Funds	-0.183	-0.115**	-0.241		
	[2.374]	[0.0516]	[0.184]		
Domestic Equity	-1.182***	-0.0143		-0.169***	
	[0.425]	[0.0183]		[0.0526]	
Foreign Equity	1.621***	-0.0151		0.0181	
	[0.576]	[0.0317]		[0.0918]	
Cash	-5.742**	0.0463			-0.0442
	[2.696]	[0.0901]			[0.105]
Bonds					0.00897
					[0.0146]
Observations, 1 per fund	112	112	70	97	94
R-Squared	0.337	0.009	0.072	0.38	–
Cragg-Donald F-stat	5.292				
F-Stat	0.00145				

Note: Observations in this Table are limited to one observation per fund, collapsed to funds who have at least 3 years of portfolio returns for which tracking errors can be calculated. The dependent variable in numbered columns is the realized tracking error for the fund, calculated by regressing portfolio returns on benchmark returns with no constant for each pension fund. The residuals are squared, and we take the standard deviation of the mean squared error across time. The far left column presents Equation I estimates of the effect of political agency on compensation for the sample in column (1). A similar estimate (unreported) is used for each of columns (2)-(4). Municipal Worker, Teachers, and Budget Civil Servants are the trustee composition outrage variables. Worker Wages and Regional Income are the outrage reference wages, equal to mean pension workers wages and meadian local area income. Political Chair is equal to one for funds whose chair is appointed by the government. Underfunded Index is an index constructed by taking the mean across the standardized value of (1- the funded ratio) and age following Rauh (2008). Log Size is the log of the lagged fund AUM.Weight variables are asset allocation weights, including null weights. *** denotes $p<0.01$, ** denotes $p<0.05$, and * denotes $p<0.1$.

Table 2.11: Effect of Outrage on Asset Class Weights Results

		(1)	(2)	(3)
	Equation I:	Equation II: Asset Class Weights		
	Log Compensation	Alternatives	Public Equities	Fixed Income
Outrage-Predicted		0.0355	-0.0666	0.0375
Log Compensation		[0.0144]**	[0.0179]***	[0.0156]**
		[0.0256]	[0.0314]***	[0.0344]
Municipal Workers	-0.801			
	[0.797]			
Teachers	-0.291			
	[0.371]			
Budget Civil Servants	-0.367			
	[0.514]			
Log Regional Income	1.221***			
	[0.320]			
Log Worker Wages	-0.206			
	[0.300]			
Political Chair	-0.0771	0.0000938	-0.0169	0.0177
	[0.132]	[0.00962]	[0.0117]	[0.00960]*
		[0.0145]	[0.0184]	[0.0146]
Underfunded Index (lag)	-0.00322	0.00904	0.00126	-0.0101
	[0.0675]	[0.00587]	[0.00712]	[0.00584]*
		[0.00843]	[0.00749]	[0.00673]
Log Size (lag)	0.389***	0.0154	0.0037	-0.0220**
	[0.102]	[0.00884]*	[0.0109]	[0.00923]
		[0.0140]	[0.0216]	[0.0202]
Observations		197	197	197
Wald Chi-squared		64.4	50.41	17.9

Note: Reported in columns (1)-(3) are estimates from a GMM system of two equations. The dependent variable in numbered columns (denoted above Equation II) is the asset class weight. The sample is limited to fund-years for which we observe a full (sums to unity) set allocation weights. The far left column presents (Equation I) estimates of the effect of political agency on compensation for the sample in column (1). A similar estimate (unreported) is used for each of columns (2)-(3). In columns (1)-(3), the log compensation variable is the outrage-predicted compensation, from Equation I (the left column). Municipal Worker, Teachers, and Budget Civil Servants are the trustee composition outrage variables. Worker Wages and Regional Income are the outrage reference wages, equal to mean pension workers wages and meadian local area income. Political Chair is equal to one for funds whose chair is appointed by the government. Underfunded Index is an index constructed by taking the mean across the standardized value of (1- the funded ratio) and age following Rauh (2008). Log Size is the log of the lagged fund AUM. Standard errors are clustered at the fund level. Two sets of standard errors are presented beneath the coefficient - standard errors clustered at the fund level (top) and robust standard errors under the seemingly unrelated assumption (bottom), inlcuded because of the joint determination of allocation weights. *** denotes $p<0.01$, ** denotes $p<0.05$, and * denotes $p<0.1$.

Table 2.12: Effect of Outrage on Portfolio Delegation

		(1)	(2)	(3)	(4)
	Equation I:	Equation II: Delegation Fraction (Tobit Estimates)			
	Log Compensation	Portfolio	Alternatives	Public Equities	Fixed Income
Outrage-Predicted		-0.639***	-0.635***	-0.273*	-0.31
Log Compensation		[0.186]	[0.217]	[0.165]	[0.223]
Municipal	-1.334				
	[1.022]				
Teachers	0.4				
	[0.634]				
Budget Civil Servants	0.459				
	[0.808]				
Log Regional Income	1.829***				
	[0.472]				
Log Worker Wages	0.189				
	[0.158]				
Political Chair	-0.11	-0.198**	0.0653	0.0665	0.3
	[0.164]	[0.0962]	[0.0919]	[0.150]	[0.218]
Underfunding Index (lag)	0.0322	-0.0897	0.152	0.241	0.179
	[0.0996]	[0.117]	[0.102]	[0.171]	[0.184]
Log Size (lag)	0.362***	0.0684	0.0503	-0.225*	-0.532***
	[0.104]	[0.105]	[0.0967]	[0.126]	[0.169]
Observations	245		258	245	251
Pseudo R-squared	0.531				
F-Stat	29.17				

Note: Reported in columns (1)-(3) are marginal effects from Tobit-MLE estimates from a system of two equations. The dependent variable in numbered columns (denoted above Equation II) is the delegation fraction to external managers. (For alternatives, we omit hedge funds, which are all delegated.) The far left column presents (Equation I) estimates of the effect of political agency on compensation for the sample in column (1). A similar estimate (unreported) is used for each of columns (2)-(4). In columns (1)-(4), the log compensation variable is the outrage-predicted compensation, from Equation I (the left column). Municipal Worker, Teachers, and Budget Civil Servants are the trustee composition outrage variables. Worker Wages and Regional Income are the outrage reference wages, equal to mean pension workers wages and meadian local area income. Political Chair is equal to one for funds whose chair is appointed by the government. Underfunded Index is an index constructed by taking the mean across the standardized value of (1- the funded ratio) and age following Rauh (2008). Log Size is the log of the lagged fund AUM. Standard errors are clustered at the fund level. *** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1.

Chapter 3

Characteristics of Mutual Fund Portfolios: Where Are the Value Funds?[1]

3.1 Introduction

Since the seminal study by Jensen (1968) the focus of most of the literature on active mutual funds has been on the question about their performance and the related issue about whether fund managers have skill or not. Some recent examples include Fama and French (2010), Berk and van Binsbergen (2015b), Cremers and Petajisto (2009), Pstor, Stambaugh, and Taylor (2015), Kacperczyk, Nieuwerburgh, and Veldkamp (2014), and many more. The composition and characteristics of mutual funds portfolios have largely been ignored.[2] For example, the performance literature focuses on the distribution of Jensen's α's across funds but pays less attention to the distributions of β's of risk factors.

The goal of this paper is to provide a comprehensive analysis of the cross-section of portfolios of active mutual funds through the lens of risk (anomaly) factors.[3] Following Fama and French (1992), the asset pricing literature has identified an ever-growing list of characteristics that are associated with return premia (see Harvey, Liu, and Zhu (2016) for a recent overview). According to the three "classic" size, value and momentum anomalies, small stocks, value stocks, and high momentum stocks earn return premia relative to large, growth and low momentum stocks. To what extent do active fund managers exploit these factor premia? If there are limits to arbitrage, do active funds contribute to the existence of these anomalies or do they overweight underpriced stocks? And, more broadly, what set of strategies is available to retail investors via active funds? The literature on mutual

[1]Based on joint work with Martin Lettau and Sydeny Ludvigson.

[2]One recent exception is Pastor, Stambaugh, and Taylor (2017) who study the relationship between liquidity and fund characteristics, in particular the optimal choice of stocks of different size.

[3]There is an ongoing debate whether the factors are due to risk premia or behavioral biases. We remain agnostic about the underlying source of factor premia.

funds typically takes the universe of funds as given. However, the set of funds in existence is an endogenous object subject to demand and supply. What are the market forces that determine the set of funds that are available to investors?[4] This paper takes a first step in answering these questions by establishing a comprehensive set of stylized facts about the characteristics of portfolios of mutual funds, ETFs and, to a limited degree, hedge funds.

We find that (most) mutual funds do not systematically exploit return premia of well-known risk/ anomaly factors. In fact, for some factors mutual funds target the low-return leg of long/short factor portfolios rather than the high-return leg. This bias is especially strong for book-to-market (BM) ratios. The BM premium is one of the most well-known and robust stylized facts in the asset pricing literature. Yet, the BM ratio of mutual funds, ETFs and hedge funds is tilted towards low BM values rather than high BM ratios. While there are over 1,000 mutual funds with consistently low BM ratios, there are virtually no high-BM funds in our sample. When we analyze fund portfolios in more detail, we find that even funds with an explicit "value" objective hold a larger share of low BM stocks than high-BM stocks in their portfolios. This bias is present in other value/growth measure, such as the earnings-to-price and dividend-to-price ratios as well as the Morningstar value/growth index. While there are over 100 "value" ETFs in our sample, very few have consistently high BM-ratios. Instead, ETFs mostly track indices that are based on the Morningstar value/growth index that is based not only on price-multiples but also on growth rates of fundamentals. Yet, portfolio sorts based on the Morningstar index produce a small and insignificant return spread. We conclude that the universe of active mutual funds and ETFs, with some exceptions, does not include high-BM investments. The BM distribution of our limited sample of hedge funds is close to that of mutual funds. We also finds that the majority of mutual funds hold predominantly very large stocks. The fund-level distributions of other factor characteristics that are associated with return premia, such as momentum, profitability and investment growth are centered around the CRSP-VW index and exhibit little variation across funds. This suggests that funds do not systematically target these characteristics. The body of the paper focuses on the presentation of empirical findings. We return to the implications of the results in the conclusion.

Our analysis focus primarily on holdings of mutual funds instead of factor exposures estimated from regressions of fund returns on factor portfolios. There are several reasons why holdings give a more accurate description of mutual fund strategies than factor loadings. First, factor loadings are estimated and thus subject to estimation error while holdings data is directly observable. Second, loadings might vary over time and estimates with historical data might not reflect high-frequency changes in fund portfolios. Third, regression loadings are more difficult to interpret than characteristics computed from portfolio holdings, as we will show below.

We use data on fund holdings to construct characteristics of active mutual funds in each quarter that a fund is listed. The paper focuses on the "classic" size, value and momentum

[4]Berk and Green (2004) study how demand and supply affect flows performance across funds but they take the set of funds that are available to investors as given.

characteristics, while the appendix includes results on a variety of other characteristics (e.g., investment and profitability). Fund-level characteristics are constructed by appropriately value-weighting the stocks in fund portfolios, e.g., the book-to-market ratio (BM) of a mutual fund is the portfolio-weighted average of the BM ratios of all stocks in the fund's portfolio. Following Daniel (1997), we use quintile scores based on NYSE breakpoints, so that a stock in quintile j has a score of j.[5] The fund score is the value-weighted score of all stocks in its portfolio. Hence, a fund that only invests in stocks in the lowest BM quintile has a BM score of "1" and a fund that only invests in stocks in the highest BM quintile has a BM score of "5". A score of "3" corresponds to a fund that focuses on stocks in the middle BM quintile. In addition to BM, we also compute an alternative measure of value/growth following the methodology of the widely-used Morningstar index.[6]

We then study the distributions of mutual fund characteristics in several ways. We analyze the average univariate distributions of size, value/growth, and momentum as well as joint distribution and time-variation in the distribution. We use two methods to frame fund characteristics. First, we use the components of the Fama-French factors as natural "pseudo-fund" benchmarks. In other words, we treat the "S" and "B" components of SMB, "H" and "L" components of HML, and "U" and "D" of MOM as if there were mutual funds and compute their characteristics in the same way as we do for actual mutual funds. Hence we can investigate to what extent mutual fund portfolios compare to the Fama-French portfolios that have served as benchmarks in the academic literature. Second, we compare the characteristics distribution of mutual funds to that of individual stocks.

We find that, except for a relatively small number of small-cap and mid-cap funds, mutual funds overwhelmingly hold very large stocks, similar in magnitudes to the CRSP-VW index and "B" in SMB. The distribution of book-to-market ratios of mutual funds, shown in Figure 3.1, is more surprising. The figure, described in more detail below, shows the histogram of BM scores for all funds (solid black), "growth" funds (dashed green) and "value" funds (dotted blue).[7] The BM ratio of funds is based on quintile scores where "1" and "5" corresponds to the extreme low and high BM quintiles, respectively. The vertical lines show the BM ratio of the CRSP-VW index and the "L" and "H" components of HML. Hence, funds that mimic "L" or "H" would have a BM scores of 1.27 and 4.6, respectively, while the overall market has a BM score of 2.3. The figure shows that the distribution of mutual funds is heavily tilted towards low BM values. 40% of all mutual funds have a BM score between 1 and 2 and a further 51% between 2 and 3. On the other hand, 9% of funds have a moderately high BM score between 3 and 4, but only 7 out of 2,657 funds in our sample have a BM score above 4. In this sense, high BM "value" funds are missing from the US equity market. To put this differently, an investor can easily find "growth" mutual

[5] Our results are robust to alternative characteristic measures. The online appendix includes a variety of robustness checks.

[6] The Morningstar index (MS) is an average of price-to-fundamental ratios and growth rates of fundamentals, see section 3.2 for details.

[7] We classify mutual funds as "value" or "growth" based on the fund name or on CRSP/Lipper/Wiesenberger style codes, as explained in more detail below.

funds that are similar to the "L" portfolio, but it is virtually impossible to use mutual funds to mimic the "value" portfolio "H". In contrast to the BM distribution of mutual funds, the BM distribution of individual S&P 500 stocks is much more spread out. 46% of S&P 500 stocks have a BM score above 3 and 18% above 4.

Figure 3.1 also shows that even "value" funds are not necessarily high BM funds. The dotted blue histogram shows that the bulk of "value" funds have BM scores between 2 and 3.5. In contrast, the majority of "growth" funds have low BM scores between 1 and 2. Moreover, the BM distribution of hedge funds and ETFs is similar to that of mutual funds (with the caveat that our sample of hedge funds is very limited and not representative). These findings are robust to different measures of "value" and different methodologies for constructing BM rankings.

In contrast to the asymmetric distribution of BM, the distributions of other characteristics are more symmetric and clustered around scores of 3. For example, the mean mutual fund has a momentum score of 3.28, a profitability score of 3.17 and investment score of 3.08. In each case, few funds have scores below 2 and above 4. This suggests that funds do not systematically exploit high momentum, high profitability or high investment strategies. We estimate the relationship of mutual portfolio holdings and characteristics more formally using a Probit-model. The estimation results confirm the patterns in the histograms. One interesting finding is that stocks with higher Morningstar indices are more likely to be held by mutual funds than stocks with high book-to-market ratios.

Next, we study portfolio compositions in more detail and compute the portfolio weights by quintiles for each mutual fund. The average fund holds 40% of its portfolio in stocks with BM scores between 1 and 2 and only 6% in stocks in the highest BM quintile. Not surprisingly, the portfolios of "growth" funds are even more tilted towards low BM stocks. For example, 95% of all "growth" funds hold over a quarter of their portfolios in low-BM stock. But we find that "value" funds hold a larger portion of their portfolio in stocks in the lowest BM quintile (24%) than in stocks in the highest BM quintile (13%). More than half of all "value" funds hold a larger share of low-BM stocks than high-BM stocks, and only 7% hold more than 25% of their portfolio in high-BM stocks. Evidently, "value" funds are not high-BM assets and do not tilt their portfolios towards high-BM value stocks.

We also study the joint distribution of characteristics as well as the time-variation of fund characteristic scores. For instance, we find that unconditionally, there is no link between the BM and MOM scores of mutual funds but momentum of low-BM funds varies significantly over time while momentum of higher BM funds is more stable.

How do portfolio holdings compare to factor loadings, β's, of mutual funds? On first glance, loadings yield a very different picture. For each 15-year window in our sample, we estimate the 4-factor model for all mutual funds that are listed in that subsample window. The median of mutual fund SMB β's distribution is (slightly positive) while the median HML β hovers around zero. Hence, at first glance, these betas suggest that the median mutual funds is slightly tilted towards small and BM-neutral stocks. How can this be reconciled with the results based on fund holdings that mutual funds hold very large and low BM stocks? It turns out that the β distributions are misleading without proper framing. We estimate the SMB and HML β's of the S, B, H and L portfolios from which SMB and HML are constructed. We show that β's estimated in univariate regressions depend on the relative volatilities of the components of the long/short portfolio. The β of the more volatile component is larger (in absolute) value than that of the less volatile component. In multivariate regressions, the magnitudes of β's depend on the covariance structure of all portfolios that make up the long/short factors and are not necessarily centered around zero. In our sample, $\beta_{\mathrm{H,HML}} = 0.72$ and $\beta_{\mathrm{L,HML}} = -0.28$. Hence, an asset with an HML beta of -0.25 is comparable to "L" while an asset with an HML beta of +0.25 is comparable to the "BM-neutral" portfolio (H+L)/2. Once the distribution of mutual fund betas is framed in the context of HML betas of H and L, the bias of mutual funds towards low-BM ratios is confirmed. While there are many funds with HML betas close to L, there are (virtually) no funds with HML betas that are as high H.

While studying mutual fund performance is not the primary goal of this paper, it is worthwhile to ask how mutual fund characteristics relate to their returns. When we compute the average return of stocks by characteristic quintiles, the familiar pattern emerges: Small stocks with high BM and MOM have higher returns than large stocks with low BM and MOM. We also find that the return spread across Morningstar quintiles is much smaller than that of BM quintiles, implying that one of the most popular measures of "value/growth" is not associated with a significant return premium. For all characteristics, the spread across quintiles is much smaller for mutual funds than for stocks. For example, the BM return spread for stocks is 2.82% per quarter while it is only 0.78% for mutual funds. Mutual fund returns across size, momentum and Morningstar quintiles show no pattern at all. Hence, there is no size, value and momentum effect in mutual funds returns and investors in mutual funds are not rewarded for factor premia that exist for individual stocks. These results are confirmed in Fama-MacBeth regressions. Our results are consistent with those in Becker, Ferson, Myers, and Schill (1999), who show that returns of value funds are similar to returns of growth funds. They hypothesize that the lack of value/growth spread in mutual fund returns might be due to value funds not taking extreme positions. While the distribution of mutual holdings confirms this hypothesis, we show that the value/growth return spread is much smaller for mutual funds than for stock even when comparing funds and stocks with similar BM ratios.

The rest of the paper proceeds as follows. Section 3.2 describes the sample and data construction. Results about the characteristics distributions of mutual funds, ETFs and hedge funds are presented in section 3.3. Section 3.4 compares the characteristics distributions derived regression factor loading to those of portfolio holdings. Sections 3.5 and 3.6 provide additional details of mutual funds portfolios. Results about the link between mutual fund characteristics and returns are reported in section 3.7. Section 3.8 concludes.

3.2 Data Construction

The mutual fund and ETF holdings data are from CRSP/Thompson-Reuters. Our sample is 1980Q1 to 2016Q2 and uses standard screens. We group active mutual funds into three categories based on their stated investment objective: "Growth", "Value" and "Other". We analyze ETFs separately. Unlike mutual funds, hedge funds are not required to report their portfolio holdings to the SEC. However, every institutional investment manager, including hedge funds, with at least $100 million in equity assets under management has to disclose their aggregate equity holdings using form 13F. Since only aggregate holdings are reported, it is not possible to obtain holdings data for individual funds for the majority of hedge funds. Instead, we manually identify 13F filings of 114 hedge funds with only a single fund under management.[8] For this subset of hedge funds, the 13F filings of portfolio holdings correspond to individual funds and are thus comparable to the holdings data of individual mutual funds. Given that we can only identify portfolio holdings of hedge funds with only a single individual fund, our HF sample is very limited, not representative and biased towards small hedge funds.

Table 3.1 reports descriptive statistics of the sample. Our sample of active mutual funds consists of 2,638 funds, of which 574 are "value" funds and 1,130 are "growth" funds. Furthermore, the sample includes 955 ETFs and 114 hedge funds. The number of active mutual funds has grown from 185 in 1980Q1 to 1,424 in 2016Q2 with a peak of 1,946 in 2008Q3. The number of "growth" and "value" funds has risen from 96 and 7 in 1980Q1 to 564 and 350 in 201Q2, respectively. The median fund size is $149 mil. over the sample but the size distribution is heavily right-skewed. In 2016Q2, the net asset value of 320 funds exceeded $1 bil. and 30 funds exceeded $10 bil.

The average age of mutual funds is 11.5 years and "growth" funds are slightly older on average than "value" funds. Not surprisingly, ETFs are on average younger than mutual funds. The number of stocks in mutual fund portfolios varies substantially across funds. The median number of stocks is 54 with a minimum of 10 and a maximum of 1,813. ETFs hold on average 99 stocks in their portfolios. Consistent with the literature on mutual fund performance, returns of mutual funds are on average lower than those of the S&P 500 index; however, the median ETFs has a higher return than the S&P index. Median 4-factor alphas are negative, including those of ETFs.

Mutual fund characteristics

Next, we construct characteristics of mutual funds, ETFs and hedge funds. The paper includes results for size (market equity, ME), the book-to-market ratio (BM), momentum (MOM) characteristics as well as the Morningstar value/growth index (MS, defined later in this section). Results for other characteristics, including other price multiples, ROE and asset growth, are reported in the online appendix. We consider a number of different methods.

[8] To identify hedge funds in the 13F filings we follow Agarwal, Fos, , and Jiang (2013a) and Agarwal, Jiang, Tang, and Yang (2013b).

The benchmark case follows Daniel (1997): In each quarter t we sort all stocks into five quintiles based on characteristic C using NYSE breakpoints. Stock i in quintile j is assigned a characteristic score of $C_{i,t} = j, j \in \{1, 2, \ldots, 5\}$. The characteristic score of fund m in quarter t, $C_{m,t}$, is computed as the portfolio-weighted average of the characteristic scores of the stocks in the fund's portfolio:

$$C_{m,t} = \sum_{i \in S_t} w_{m,i,t} \, C_{i,t},$$

where S_t is the set of stocks listed in quarter t and $w_{m,i,t}$ is the weight of stock i in the portfolio of fund m in quarter t.

This procedure has several advantages. First, it is robust to stocks with extreme values of $C_{i,t}$. Second, scores have the same units and are comparable across characteristics. On the other hand, quintile scores depend on the breakpoints. We follow the standard procedure and use NYSE breakpoints. Note that the total market capitalization of the stock quintiles varies across quintiles and, therefore, the value-weighted market portfolio does not necessarily have a characteristic score equal to the midpoint of 3 but will be biased towards the quintiles with larger market caps. For example, the top size quintile accounts for about 73% of the total market cap while the bottom quintile accounts for only 3%. Hence, the size quintile score of the value-weighted CRSP index will be strongly tilted towards the fifth quintile. In contrast, the low BM quintiles account for a larger share of the total market cap that the high BM quintiles. Thus the BM score of the CRSP-VW index is below the midpoint of 3.[9]

As an alternative measure, we compute "market-adjusted" characteristics. For example, in each quarter we compute the "market-adjusted" BM ratio for each stock i as

$$\widetilde{BM}_{i,t} = \frac{BM_{i,t}}{BM_{MKT,t}},$$

where $BM_{MKT,t}$ is the book-to-market ratio of the CRSP-VW index. The market-adjusted BM ratio of a mutual fund is the portfolio-weighted average adjusted BM-ratios of the stocks in its portfolio. For momentum, we compute the difference of momentum of each stock and momentum of the CRSP-VW portfolio:

$$\widetilde{MOM}_{i,t} = MOM_{i,t} - MOM_{MKT,t}.$$

This method has the advantage of not relying on breakpoints. Moreover, the market-adjusted characteristics of the value-weighted CRSP portfolio are equal to one for $\widetilde{BM}_MKT$ and zero for $\widetilde{MOM}_MKT$ by construction. On the other hand, adjusted characteristics can be sensitive to outliers. For example, the distribution of stock-level $\widetilde{MOM}_i$ is right-skewed. The minimum $\widetilde{MOM}_i$ is 0.04, the median is 1.35, and the maximum is 14.47. Hence the

[9]We also consider the case where breakpoints are chosen so that the market cap in each quintile is identical. Results are reported in the online appendix.

mutual fund level $\widetilde{MOM}_m$ constructed as the portfolio-weighted average of the stock-level $\widetilde{MOM}_i$ can be dependent on whether the mutual fund holds one of the few outlier stocks with very high $\widetilde{MOM}_i$. Another drawback is that the units differ across characteristics making a comparison difficult. Most of the results reported in the paper are based on characteristics scores. The online appendix includes results for adjusted characteristics and quintile score based on different breakpoints. Our main results are not affected by the methodology of how mutual funds characteristics are constructed.

While the book-to-market ratio has become the standard metric for value/growth in academic research, there are many alternative measures. One popular measure is the Morningstar value/growth index that is used in Morningstar's "style box". The MS value/growth index is defined as the difference of a multiples (MULT) index and a growth (GR) index. Both components are scaled from 0 to 100 so that the MS index ranges from -100 to 100:

$$\text{MULT} = \frac{1}{2}\frac{E(\text{Earn})}{P} + \frac{1}{2}\,\text{avg}\left(\frac{B}{P}, \frac{S}{P}, \frac{CF}{P}, \frac{D}{P}\right)$$

$$\text{GR} = \frac{1}{2}\Delta E(\text{LT Earn}) + \frac{1}{2}\,\text{avg}\left(\Delta\text{Earn}, \Delta S, \Delta CF, \Delta B\right)$$

$$\text{MS}[-100, 100] = \text{scaled MULT}[0, 100] - \text{scaled GR}[0, 100],$$

where $E(\text{Earn})$ are the expected earnings, $E(\text{LT Earn})$ are expected long-term earnings and P, B, S, CF, D are price, book value, sales, cash flow, and dividend, respectively. MS has two components: *(i)* an average of multiples (MULT) and, *(ii)* an average of expected long-term earnings growth $E(LT\Delta E)$ and growth of current earnings, sales cash flow and book value (GR). Note that the terms with *expected* earnings have a larger weight in MULT and GR than the terms with current fundamentals. The index is constructed so that high MS scores correspond to "value" and low MS scores correspond to "growth" in line with the BM ratio.[10] We construct the MS index for each stock in each quarter, form quintiles and compute the MS score for mutual funds as the portfolio-weighted average of MS scores of the stocks in the fund's portfolio. More details are given in the Appendix.

We also compute the characteristics of the components of the Fama-French portfolios, SMB, HML, and MOM, as benchmarks. For example, HML is defined as HML = 1/2 (SH+BH) - 1/2 (SL+BL), where SL is the small/low-BM portfolio, BL is the big/low-BM portfolio, etc. The component portfolios of HML are based on the intersection of two size and three BM-sorted portfolios (with NYSE breakpoints). We treat each of the component portfolios, SL, BH, ..., as a "passive mutual fund" and construct its characteristics following the same methodology described above for mutual funds. Lastly, we compute the characteristics of the CRSP-VW portfolio as a proxy for the market portfolio.

[10] The Morningstar index used in the style box is defined as scaled $\text{GR}[0, 100] -$ scaled $\text{MULT}[0, 100]$. We adjust the definition so that low/high MS values have the same value/growth interpretation as low/high BM scores.

As an illustration, Figure 3.2 plots the characteristics of one of the oldest and largest mutual funds, "The Investment Company of America Fund" (ticker AIVSX), over time. Panel A shows the characteristic scores while the market-adjusted characteristics are plotted in Panel B. Adjusted MS is divided by 10 to make the scales comparable. ME scores are close to the maximum of five indicating that the fund only invests in the very largest stocks. The BM score ranges from 1.5 to 2.6 while the MS score exhibits an upward trend and is higher than the BM score throughout the sample. BM and MS are both value/growth measures, but they can differ substantially. Recall that MS is a combination of price multiples and growth rates of fundamentals. The plot suggests that the fund targets firms that have high fundamental growth rates rather than firms with high BM ratios. MOM scores vary more over the sample than ME, BM and MS scores. This is not surprising since the persistence of momentum on the stock level is lower than that of the other characteristics and will be the case for most mutual funds.

Market-adjusted characteristics show similar patterns. Recall that the adjusted ME, MS and MOM characteristics of the CRSP-VW portfolio are zero and one for adjusted BM. However, the scales are not comparable since different characteristics have different "units". The plot shows that adjusted ME and MS are (mostly) positive suggesting that the fund invests in very large stocks that have higher MS values than the market. In contrast, adjusted BM is close to one apart from the first 10 years of the sample. Finally, adjusted MOM hovers around zero.

3.3 Mutual Fund Portfolios

Passive benchmarks

Before analyzing characteristics of mutual funds, we start with the characteristics of the CRSP-VW index and the components of the Fama-French SMB, HML and MOM factors as benchmarks for mutual fund characteristics. Table 3.2 reports the average scores as well as average adjusted characteristics of the CRSP-VW index, the components of HML (SL, BL, SH, BH) and the components of MOM (SD, SD, SU, BU). Consider first the characteristic scores of the "market" CRSP-VW index. The average value-weighted size (ME) score is 4.50 while the average book-to-market (BM) score is 2.31. The average Morningstar (MS) score is slightly higher than the average BM score. The average momentum (MOM) score is 3.44. The reason these value-weighted averages are are not equal to the midpoint of 3 is that the total market capitalizations in each quintile are different. As mentioned above, the market cap in the fifth size quintile is much larger than that of the first quintile. Hence, the average ME score of the value-weighted CRSP index is higher than 3. For the same reason, the BM and MS scores are below the midpoint of 3 while the MOM score is above 3. In contrast, the "adjusted" characteristics of the CRSP-VW index are either 1 or 0, by construction.

Next, consider the passive Fama-French portfolio components. The four "small" portfolios have ME scores between 1.86 and 2.07, while the "big" portfolios range between 4.6

and 4.8. This pattern is similar for the other characteristics. The “low BM” portfolios have a BM score of 1.28 and 1.25, respectively, and the scores of the “high BM” portfolios are 4.63 and 4.56. Note that the BM, MS and MOM scores of “small” and “high” portfolios are similar, the BM score of “small/high BM” is 4.63 while the score of the “big/high BM” portfolios is 4.56. Hence, portfolios with high BM scores can be constructed not just from small and potentially illiquid stocks but also from large liquid stocks. Given this similarity, we follow Fama and French and aggregate the “small” and “high” portfolios into a single portfolios, e.g. “small/high BM” and “big/high BM” are combined into “high BM”, etc. These portfolios correspond to the components of the SMB, HML and MOM factors.

Portfolio Characteristics of Mutual Funds

Next, we study the univariate distributions of mutual fund characteristics. Results for multivariate distributions are presented in section 3.6. The histograms of mutual funds scores of size (ME), book-to-market (BM), Morningstar (MS) and momentum (MOM) are shown in Figure 3.3. Each panel shows the histogram for all funds (solid black) in the sample as well as “growth” (dashed green) and “value” (dotted blue) funds. The numbers at the bottom of each histogram represent the percentage of all funds with characteristic scores between 1 and 2, 2 and 3, etc., respectively. The vertical lines show the scores of the CRSP-VW index and the passive Fama-French benchmarks. The percentiles of the distributions are reported in Table 3.3.

The ME histogram in Panel A shows that the size score of 79% of all funds is above 3 implying that most mutual funds invest in very large stocks. The histogram shows that 19% of funds have an ME score between 2 and 3. The vertical lines indicate the characteristics scores of the CRSP-VW index as well as scores of the “small” and “big” portfolios S and B (see Table 3.2). The size score of 33% of all mutual funds is higher than the size score of the H portfolios. In contrast, only 2% of all active mutual funds have a size score comparable that of the “small” S portfolio. Thus, the stocks that make up the composition of the “S” component of SMB are significantly smaller than the stocks held by all but 2% of mutual funds, and it is thus virtually impossible to replicate a portfolio similar to “S” by investing in mutual funds. Clearly, most mutual funds do not exploit the small stock premium. The figure also shows the size distribution of “growth” and “value” funds. The ME distribution is similar for “growth” and “value” funds, although growth funds have somewhat larger ME scores than “value” funds. Pastor et al. (2017) argue that mutual funds tilt towards large stocks because small-stocks a more expensive to trade. In equilibrium, funds optimally choose the tradeoff of trading costs versus potentially higher returns of small stock. Large funds have higher trading costs and therefore hold large stocks.

The BM histogram in Panel B is identical to that in Figure 3.1. As already described in the introduction, the BM distribution is heavily skewed towards low BM scores as 91% of all funds have a BM score below 3, and virtually no funds have a BM score that exceeds 4. The histogram also shows that many funds have a BM score that is close to that of the “Low BM” portfolio but no funds with a BM score that is similar to that of the “High BM” portfolio. Only 7 of the 2,657 funds in the sample are in fact high-BM funds with a score above 4, while 1,063 funds have a BM score below 2. While it is not surprising that the distribution of "growth" funds is more skewed towards low BM scores, it is noteworthy that the BM score of the majority of "value" funds is below 3. The means of BM scores, shown in Table 3.3, are 2.23 for all funds, 1.89 for "growth" funds and 2.74 for "value" funds. Thus, "value" funds with high BM scores are largely missing. This fact is not driven by the lack of large and liquid stocks in the top quintile. Recall from Table 3.2 that the "big/high BM" portfolio is made up entirely of large stocks but has a BM score of 4.56.

One possible explanation is that the BM ratio does not capture the notion of "value" as viewed by fund managers and investors. To explore this possibility further, we compute the Morningstar value/growth index (MS) that underlies the well-known Morningstar style box. As explained in section 3.2, the Morningstar index is an average of price multiples and growth rates of firm fundamentals. The histogram of MS scores in Panel C shows that the distribution is somewhat shifted to the right compared to the BM distribution but still skewed towards low MS scores. 33% of mutual funds have an MS score below 2 while only the MS score of only 1% is above 4.

Finally, Panel D shows the momentum (MOM) histogram. The vast majority of mutual funds have a MOM score between 3 and 4 and are thus somewhat tilted towards higher momentum stocks. However, only 4% of funds have a MOM score above 4 indicating that few funds focus on momentum as a primary strategy. We will see below that the momentum tilt is due to the fact that most funds hold low BM stocks that on average have higher MOM scores than high BM stocks. Moreover, we will also show that there is more time variation in the momentum scores of individual funds than in the size and growth/value scores. Hence the distribution of fund averages is less informative for momentum than for the other more persistent characteristics.

Since mutual funds hold mostly large stocks, it is instructive to compare the characteristics distribution of mutual fund portfolios to that of individual large stocks. Figure 3.4 plots the ME, BM, MS and MOM histograms of individual stocks (dashed black) along with the histograms for mutual funds. We include stocks that were a constituent of the S&P 500 index for at least eight quarters during the sample period. The size distribution in Panel A shows that ME scores of mutual funds are on average higher than those of S&P 500 stocks confirming the previously mentioned result that mutual funds hold mostly very large stocks. Panel B plots the BM score distribution. The BM scores of S&P 500 stocks is more spread out than those of funds. While there are few funds with a BM score above 3, 39% of stocks have BM scores that exceed 3. The average BM score of stocks is 2.62, which is significantly higher than the mean for mutual funds of 2.23 (Table 3.3). Fund managers choose portfolios that are more tilted towards low book-to-market values than the set of stocks that are available to them. Moreover, since there are many large liquid stocks with high BM scores, there is no obvious constraint that might preclude managers from constructing high BM funds. The distribution of Morningstar scores in Panel C shows that fund portfolios have lower MS scores than those of stocks confirming that fund managers choose portfolios that are tilted towards growth than the market overall. Panel D shows that the momentum distribution of mutual funds is slightly shifted to the right relative to that of individual stocks.

Next, we take a closer look at the mutual funds with the lowest and highest BM scores. Table 3.4 shows the 10 funds with the highest BM score and the 10 funds with the lowest score. The scores of the "H" component of HML are included for comparison. In our sample of 2,657 funds, only seven funds have a BM score above 4, and only one fund exceeds the BM score of "H". Only four of the 10 funds have an AUM above $1 bil. Interestingly, three of the four large funds are Dimensional Fund Advisor (DFA) funds that, according to their prospectuses, specifically target stock with high price multiples but, in contrast to the Morningstar notion of "value", do not take fundamental growth into account.[11] Note, however, that the BM scores of the DFA funds are significantly below that of the "H" portfolio. The bottom panel shows the 10 funds with the lowest BM scores. Note that their BM scores are all below that of the low BM benchmark portfolio.

How do mutual fund portfolios compare to those of hedge funds and Exchange Traded Funds (ETFs)? Figure 3.5 shows the BM and MS histograms of ETFs and our limited samples of (mostly small) hedge funds. Panels A and B show the BM and MS distribution of hedge funds. Means and percentiles are reported in the bottom left panel of Table 3.3. The BM distribution of hedge funds is very similar to that of mutual funds with almost identical means of 2.29 and 2.23, respectively, as well as comparable percentiles. While 40% of HFs have a BM score lower than 2, there are no HFs with a BM score above 4. The MS distribution is shifted more towards low MS than the mutual fund distribution.

The BM distribution of ETFs shown in Panel C is shifted towards low BM scores but slightly less so than the distribution of mutual funds. 26% of ETFs have a BM score above 3, compared with 9% of mutual funds, but the BM score of only 3% is above 4, and no ETF approaches the BM score of the "H" portfolio. However, the ETF MS distribution in Panel B differs significantly from that of the distribution of mutual funds shown in Figure 3.3. The distribution of all ETFs is centered around the midpoint of 3 and spread out symmetrically. In other words, there are (almost) as many high MS ETFs as there are low MS ETFs. Furthermore, MS scores of "value" ETFs are much higher than those of "value" mutual funds and, unlike mutual funds, there are many ETFs with an MS score between 4 and 5.

The reason for this difference is that many "value" ETFs track indices that are constructed using a similar methodology as that of the Morningstar index, i.e., indices that are based on multiples as well as growth rates of firm fundamentals. For example, the largest "value" ETF (iShares Russell 1000 Value ETF) tracks the Russell 100 Value Index that follows the Morningstar classification closely, as stated in its documentation: "FTSE Russell uses three variables in the determination of growth and value. For value, book-to-price (B/P) ratio is used, while for growth, two variablesI/B/E/S forecast medium-term growth (2-year) and sales per share historical growth (5-year) are used". The notion of "value" in fund management differs from that in the academic literature, which has focused on pure price multiples as measures of "value". The evidence of the "value puzzle" in the academic

[11]The prospectuses of the DFA funds state: "Securities are considered value stocks primarily because a companys shares have a low price in relation to their book value."

literature is based on sorts on variables such as the book-to-market, earnings-to-price, sales-to-price, etc., but does not include information of fundamental growth rates. We will show in section 3.7 that return spreads based on portfolios sorted according to the Morningstar index are significantly smaller than that for portfolios constructed from book-to-market sorts. The reason is that the GR component of the MS index produces no return premium and the premium of the MULT component is smaller than that for BM. In other words, the widely used value/growth MS index is not associated with a "value premium".

Next, we perform a number of robustness checks that are reported in Figure 3.6. Results for additional robustness checks are reported in the online appendix. We consider the earnings-to-price ratio (EP) as an alternative "value" measure (Panel A), plot the histogram of fund/quarter observation instead of fund averages (Panel B), consider the "adjusted BM" measure instead of BM scores (Panel C), the AUM-weighted histogram (Panel D), the distribution of a larger set of mutual funds that includes index and sector funds (Panel E), and the BM distribution at four different points in time (Panel F). The distribution of EP scores is slightly shifted towards higher scores compared to the BM distribution, but there are virtually no funds with an EP score above 4. The histogram with fund/quarter observation is similar to the histogram with fund-average observations. Unlike the BM scores, the "adjusted BM" ratio does not depend on breakpoints and is scaled to the overall market has a value of one but is more sensitive to outliers. The "adjusted BM" ratios of "H" and "L" are 3.1 and 0.6, respectively, implying that the BM ratio of the "H" portfolio is about three times as high as that of the CRSP-VW index while the BM ratio of the "L" portfolio is 40% lower than that of the CRSP-VW index. The "adjusted BM" histogram confirms the pattern found for BM scores. Few funds have an "adjusted BM" ratio above 2 and there are no funds that have a ratio that is as high as that of "H". Panel D shows the histogram when funds are weighted according to their AUM. The only significant difference compared to the equally-weighted histogram is the higher mass for BM scores around 4, which is due to the large DFA value funds shown in Table 3.4. Our benchmark sample of mutual funds excluded index and sector funds. In Panel E, we plot the BM distribution for the sample that including index and sector funds. The distribution is almost identical to that for the benchmark sample. Finally, we study the BM distribution across time. Panel F shows the BM histograms of mutual funds in the fourth quarters of 1985, 1995, 2005, and 2015.

To assess more formally how stock characteristics affect mutual fund portfolios, we estimate the following Probit model:

$$P(y_{i,j,t}) = \Phi(\mathbf{X}'_{i,t}\boldsymbol{\beta}), \tag{3.1}$$

where $y_{i,t}$ is an indicator variable that is 1 if stock i is held by mutual fund j in quarter t and zero otherwise, $\mathbf{X}_{i,t}$ is a vector of ME, MOM, BM and MS characteristics of stock i in period t and Φ is the cumulative distribution function of the standard normal distribution. We estimate the model for all funds, "growth" funds and "value" funds. Results are reported in Table 3.5. The regression with all mutual funds shows that larger stocks with higher momentum are more likely to be included in mutual funds portfolios but higher BM and MS scores decrease the probability to be included in a fund portfolio. All scores are measured on the same $[1, 5]$ interval, so the magnitudes of the point estimates are comparable. The ME coefficient is by far the largest showing that mutual funds mostly invest in very large stocks. All characteristics coefficients are statistically significant.

If the Probit model is estimated for only "growth" funds, the MS coefficient becomes larger. It is larger by a factor of 5 than in the estimation of all funds and also larger than the BM coefficient. For "value" funds the MS coefficient is positive and significant while the BM coefficient is insignificant. These estimates are consistent with the distributions of mutual fund characteristics shown above. These results suggest that in the mutual fund industry the Morningstar index is more widely used as a measure of "value" than the book-to-market ratio.

Time-series variation of mutual fund characteristics

So far, we focused on time-series averages by fund. Next, we investigate the variations of fund characteristics over time. For each mutual fund, we compute the time-series standard deviation of ME, BEME, MS and MOM characteristics. The mean standard deviations across all mutual funds are as follows: $\overline{\sigma}_{ME} = 0.23, \overline{\sigma}_{BM} = 0.29, \overline{\sigma}_{MS} = 0.32, \overline{\sigma}_{MOM} = 0.42$. ME scores vary the least over time followed by BM and MS. Fund-level momentum has a significantly higher standard deviation than size, book-to-market, and Morningstar characteristics. On the stock level, momentum is less persistent than the other characteristics. If a mutual fund invests in stocks without using information on momentum, momentum in the mutual level will also be less persistent than the other characteristics. However, if a fund consistently targets either high or low momentum stocks, then fund momentum is more persistent than the momentum stocks in its portfolios. In our sample, the distribution of momentum persistence on the mutual fund level is very similar to that on the stock-level suggesting that mutual funds do not target either high or low momentum stocks.

As an illustration of the time series behavior of different characteristics, we plot the characteristics of the largest mutual fund in our sample as well as those for the largest "value fund" in Figure 3.7. The figure also shows plots for the characteristics of passive benchmark portfolios. The figures shows that ME and BM are stable for some funds but vary for others. The time-series variation of these characteristics tends to be on a lower frequency as funds shift their investment objectives. In contrast, the variation in fund-level MOM is of higher frequency. This variation is "passive" in the sense that it is due to changes in momentum of the stocks in a fund portfolio rather than due to portfolio reallocations. As a consequence, a fund can be high-momentum in one quarter and low-momentum in a different quarter.

3.4 Loadings vs. Holdings

In the literature on mutual fund performance, the magnitudes of regression factor loadings (i.e., betas) are less relevant since the factors serve only as controls for diversifiable risk. For our purposes, the question is whether loadings estimated from time series regressions of fund returns on factors such as SMB, HML, and MOM are informative indicators of fund strategies. Next, we argue that While factor loadings are appropriate as a measure of exposure to diversifiable risk, they are not necessarily reliable indicators of the underlying investment strategy of an active mutual fund.

First, risk exposures are estimated using historical data and are thus subject to estimation error. Historical data might also not reflect the current portfolio of an active fund. This is especially true for firm characteristics that change over time, such as momentum. Unless a fund deliberately hedges momentum, the momentum of a fund's portfolio changes as the momentum of the stocks in its portfolio changes over time. In contrast, measuring fund characteristics directly from portfolio holdings yields an accurate assessment of the fund's portfolio at each point in time.

Second, the interpretation of the magnitudes of estimated loadings in factor regressions are not straightforward and can easily be misinterpreted. Consider the univariate setting with two portfolios, P and Q, that are based on sorts on some characteristic. Let $\mathrm{PMQ}_t = \mathrm{P}_t - \mathrm{Q}_t$ be the corresponding long/short portfolios and consider the regressions of P_t and Q_t on PMQ_t:

$$Y_t = \alpha_{\mathrm{Y}} + \beta_{\mathrm{Y,PMQ}}\, \mathrm{PMQ}_t + e_{\mathrm{Y},t}, Y \in \{\mathrm{P},\, \mathrm{Q}\}.$$

Since $\mathrm{PMQ}_t = \mathrm{P}_t - \mathrm{Q}_t$, the P and Q betas have the property

$$\beta_{\mathrm{P,PMQ}} - \beta_{\mathrm{Q,PMQ}} = 1.$$

However, the magnitudes of the two betas depend on the variance-covariance structure of $[\mathrm{P}_t, \mathrm{Q}_t]$:

$$\beta_{\mathrm{Y,PMQ}} = \rho_{\mathrm{Y,PMQ}}\ \frac{\sigma_{\mathrm{Y}}}{\sigma_{\mathrm{PMQ}}}$$
$$|\beta_{\mathrm{P,PMQ}}| > |\beta_{\mathrm{Q,PMQ}}| \iff \sigma_{\mathrm{P}} > \sigma_{\mathrm{Q}}.$$

The last line follows from the fact that $Cov(\mathrm{P}_t, \mathrm{P}_t - \mathrm{Q}_t) = Var(\mathrm{P}_t) - Cov(\mathrm{P}_t, \mathrm{Q}_t), Cov(\mathrm{Q}_t, \mathrm{P}_t - \mathrm{Q}_t) = Var(\mathrm{P}_t) - Cov(\mathrm{P}_t, \mathrm{Q}_t)$. Hence, betas are not necessarily symmetric around 0 and the more volatile portfolio has a larger (in absolute value) beta with respect to the long/short portfolio. The $\mathrm{P}_t - \mathrm{Q}_t$ beta of the "neutral" portfolio $(\mathrm{P}_t + \mathrm{Q}_t)/2$ is positive if $\sigma_{\mathrm{P}} > \sigma_{\mathrm{Q}}$ and negative otherwise. In other words, the magnitudes of betas are more informative about the volatility of the portfolios that make up the long/short portfolios than as a measure of how tilted a portfolio is towards the underlying characteristic.

The dependence of regression loadings on the volatility of the long/short portfolios is borne out in the data. In our sample, univariate HML betas are not centered around zero since $\sigma_{\mathrm{L}} > \sigma_{\mathrm{H}}$ and thus $|\beta_{\mathrm{L,HML}}| > |\beta_{\mathrm{H,HML}}|$. The estimated univariate betas are $\beta_{\mathrm{L,HML}} = -0.75, \beta_{\mathrm{H,HML}} = 0.25$. The HML beta of the "BM-neutral portfolio" (H+L)/2 is -0.25. In contrast, the HML beta of a "growth-tilted" portfolio of 0.75H+0.25L is 0. Hence, a comparison of HML loadings of two portfolios based only on the magnitudes of their HML betas is misleading. Say, the HML betas of two portfolios are -0.2 and 0.2, respectively. The portfolio with an HML beta of 0.2 is much closer to "H" than the portfolio with an HML beta of -0.2 is to "L".

This pattern is even more pronounced for the SMB β's of "S" and "B": $\beta_{\mathrm{S,SMB}} = 1.60, \beta_{\mathrm{B,SMB}} = 0.60$. The *positive* sign of $\beta_{\mathrm{B,SMB}}$ is counterintuitive since SMB=S$-$B but is due to the fact that "S" is much more volatile than "B" and $Cov(\mathrm{B,S}) > Var(\mathrm{B})$. Hence, the SMB beta of any linear combination of "S" and "B" with non-negative weights is strictly positive. Thus univariate SMB betas of large stocks, or mutual funds that hold large stocks, are positive. By themselves, beta coefficients in regressions on long/short factors are generally not informative. Instead, betas need to be interpreted relative to the range spanned by the betas of the components of the long/short factors.

In multivariate regression, the patterns of betas are more complicated and depend on the joint variance-covariance structure of the left- and right-hand side variables. Consider the 4-factor model

$$Y_t = \alpha_{\mathrm{Y}} + \beta_{\mathrm{Y,MKT}}\,\mathrm{MKT}_t + \beta_{\mathrm{Y,SMB}}\,\mathrm{SMB}_t + \beta_{\mathrm{Y,HML}}\,\mathrm{HML}_t + \beta_{\mathrm{Y,MOM}}\,\mathrm{MOM}_t + e_{\mathrm{Y},t},$$

where $Y \in \{\mathrm{S, B, H, L, U, D}\}$. As in the univariate case, the betas have the property

$$\begin{aligned}
\mathrm{SMB} = \mathrm{S} - \mathrm{B} &\implies \beta_{\mathrm{S,SMB}} - \beta_{\mathrm{B,SMB}} = 1,\\
\mathrm{HML} = \mathrm{H} - \mathrm{L} &\implies \beta_{\mathrm{H,HML}} - \beta_{\mathrm{L,HML}} = 1,\\
\mathrm{MOM} = \mathrm{U} - \mathrm{D} &\implies \beta_{\mathrm{U,MOM}} - \beta_{\mathrm{D,MOM}} = 1.
\end{aligned}$$

However, the estimates of the individual β's depend on the variance-covariance of $Z_t = [\mathrm{MKT}_t, \mathrm{S}_t, \mathrm{B}_t, , \mathrm{H}_t, \mathrm{L}_t, \mathrm{U}_t, \mathrm{D}_t]'$. As with univariate betas, it is not necessarily the case that the coefficients are symmetric in the sense that $\beta_{\mathrm{S,SMB}} = -\beta_{\mathrm{B,SMB}}, \beta_{\mathrm{H,HML}} = -\beta_{\mathrm{L,HML}}, \beta_{\mathrm{U,MOM}} = -\beta_{\mathrm{D,MOM}}$.

In general, the relative magnitudes and signs can take any value. The betas of the six component portfolios of SMB, HML and MOM in 4-factor regressions are shown in Table 3.6. The SMB loading of "S" is 0.90, while the loading for "B" is -0.10 and thus much smaller in absolute value. The same pattern is true for the HML loadings of "H" and "L": $\beta_{\mathrm{H,HML}} = 0.72$ and $\beta_{\mathrm{L,HML}} = -0.28$, while $\beta_{\mathrm{U,MOM}} = 0.34$ and $\beta_{\mathrm{D,MOM}} = -0.66$. While the signs of the betas are intuitive in the sense that betas of "long" portfolios S, H and U are positive and betas of "short" portfolios B, L and D are negative, the betas are not symmetric around zero.

The effect of this asymmetry is visible in the magnitudes of SMB and HML loadings of 25 ME-BM sorted portfolios shown in Table 3.7. The SMB betas of all portfolios constructed from size quintiles 1 to 4 are positive and only the five portfolios with the smallest stocks have a negative SMB beta. The magnitudes of the SMB betas are only interpretable in comparison to the "S" and "B" betas of 0.9 and -0.10 (Table 3.6). [12] The pattern of HML betas is similar. Only the portfolios with the lowest BM quintile have negative HML betas. As in the case of SMB betas, HML betas need to be interpreted in conjunction with "H" and "L" betas. Otherwise, asset betas can lead to incorrect inference. Consider for example the "neutral" portfolio formed from stocks in the third ME and BM quintiles. The SMB beta of this portfolio is 0.51, and the HML beta is 0.42 suggesting, incorrectly, that this portfolio is tilted towards large, high BM stocks. However, the betas are close to the midpoints of the "S" and "B" SMB betas and the "H" and "L" HML betas, which is an indication that this portfolio is indeed BM-neutral and ME-neutral.

Consider two mutual funds with $\beta_{1,\mathrm{HML}} = 0.25$ and $\beta_{2,\mathrm{HML}} = -0.25$, respectively. Since the HML betas are equal in absolute value, it might seem that both funds are comparable in terms of their respective value and growth strategies. However, the HML beta of fund 2 is close to the HML beta of "L" of -0.27 while the HML beta of fund 1 is much smaller than the HML beta of "H" of 0.73. Hence, the proper interpretation is that fund 1 is a "moderate" value fund while fund 2 is an "extreme" growth fund.

The third issue with factor exposures estimates is that they are varying over time. Figure 3.8 shows this time-variation of factor betas for the passive benchmarks as well as the distribution of mutual fund SMB and HML betas. The solid lines in Panel A show "S" and "B" SMB 4-factor betas in 10-year rolling samples. In addition to the "S" and "B" loadings, the figure also shows betas of an ME-neutral portfolio of SB=(S+B)/2. Panel B shows the "H", "L" and HL=(H+L)/2 betas. The shaded areas are 95% confidence bands. SMB betas vary slightly over the sample ranging from -0.14 to -0.04 for "S" and 0.87 to 1.01 for "B". The time-variation of HML betas are more pronounced. Panel B shows that $\beta_{\mathrm{H,HML}}$ and $\beta_{\mathrm{L,HML}}$ are both higher towards the end of the sample than in early in the sample. The variation is economically and statistically significant. The estimates for $\beta_{\mathrm{H,HML}}$ range from 0.57 in 1991Q3 to 0.76 in 2012Q2 and from -0.41 in 1991Q4 to -0.21 in 2007Q2 for $\beta_{\mathrm{L,HML}}$. Hence a mutual fund with a β_{HML} of 0.5 is an extreme-value fund similar to "H" in the 1980s but only a moderate-value fund in the 2000s. Similarly, a mutual fund with a β_{HML} of -0.2 is an extreme-growth fund in the 2000s but only a moderate-growth fund in the 1980s.

Next, we estimate 4-factor betas of mutual funds in rolling 10-year regressions. Each estimation window includes all funds with at least 75% available data. Figure 3.9 shows violin plots of the distribution of mutual fund HML and SMB loadings for 10-year windows ending in the second quarters of 1988, 1995, 2002, 2009, and 2016. The figures also show the median and inter-quartile range of each distribution. The solid lines are betas of "S", "B", "H" and "L". The SMB-beta distribution in the top panel is stable over time, which is intuitive since the majority of mutual funds holds almost exclusively large stocks throughout the sample. The median of the distribution varies between 0.04 in 2009 to 0.17 in 1988.

[12]The betas of the 25 ME-BM portfolios can be larger in absolute value than the "S", "B", "H" and "L" betas since the they are based on quintiles while "S", "B", "H" and "L" are constructed from two ME quantiles and BM terciles.

The majority of mutual funds have *positive* SMB betas, from 55% in 2009 to 78% in 1988. Without proper context, this would incorrectly indicate that most mutual funds hold small stocks. SMB-betas of very few mutual funds are as high as the "S" SMB-beta; but many funds have an SMB-beta that is comparable to that of the "B" portfolio. The upper interquartile range is close to the SMB-beta of the ME-neutral SB=(S+B)/2 portfolio and lower interquartile range is close to that of the "B" beta. Hence, properly interpreted, the SMB-beta distribution confirms the pattern found in portfolio holdings that most mutual funds invest in large stocks.

The HML beta distributions are shown in Panel B. First, notice that there is time-variation in the mutual fund HML-β distribution. The distribution shifts up in the middle of the sample compared to the beginning and end of the sample. The median in 1988 is $-0.08, 0.14$ in 2002 and -0.07 in 2016. Thus mutual fund HML-betas follow a similar pattern as the HML-betas of the passive "H", "HL" and "L" portfolios shown in Figure 3.8. The medians of the distributions are around 0 and would, as in the case of SMB-betas, incorrectly indicate that funds are on average BM-neutral. However, there are virtually no mutual funds with an HML-beta close to the "H" HML-beta while most funds have an HML beta that is lower than the beta of the BM-neutral HL=(H+L)/2 portfolio. For example, in 2016 93% of all mutual funds had an HML beta that was lower than the HL beta confirming that very few mutual funds are high-BM funds, whether the degree of "value" is measured based on portfolio holdings or regression loadings.

In summary, magnitudes of SMB and HML loadings are difficult to interpret and vary over time. Estimated portfolio or mutual fund betas should be compared to betas of the individual portfolios that are used to construct long/short factors.

Finally, we compare the distribution of mutual fund loadings to that of hedge funds in Figure 3.10. The distributions of HML and MOM betas of hedge funds are very similar to those of mutual funds while hedge funds have on average slightly higher SMB-betas. However, the hedge fund and mutual funds distributions of market betas are fundamentally different. Most mutual funds have a market beta between 0.5 and 1.5 with a mean close to 0. Market betas if hedge funds are on average lower. The mean market beta is close to 0.6, and about 40% of hedge funds have a market beta of less than 0.5. Since mutual funds are restricted to only hold long positions, it is not surprising that their market betas are around one. Hedge funds can take short positions and thus create portfolios with lower market betas. However, the histograms show that hedge funds exposures to factors other than the market do not differ significantly from those of mutual funds.

3.5 Portfolio Composition by Quintiles

So far, we have focused on average portfolio characteristics of mutual funds. Next, we analyze portfolio compositions in more detail. Specifically, for each fund, we compute the average portfolios shares in each of the characteristic quintiles over the lifetime of the fund. Table 3.8 reports average portfolio shares in the five BM quintiles for the CRSP-VW index and mutual funds, hedge funds and ETFs, as well as for the five largest "value" and "growth" funds in our sample. Figure 3.11 shows the histograms of portfolio shares in BM quintiles across mutual funds, hedge funds and ETFs.

Since the total market capitalization is higher in the lower BM quintiles than in the higher BM quintiles, the average portfolio share of the CRSP-VW index declines from quintiles one to five. The average BM-quintile portfolios shares across all funds, shown in the second row of Panel A, are very close to the shares of the CRSP-VW index. The portfolios of growth funds are heavily concentrated in extreme low BM stocks. The average portfolio share of stocks in quintile one is 53% and 22% in quintile two. Only 12% of portfolios of "growth" funds are invested in higher BM stocks. However, the pattern for "value" funds is very different. The average share of stocks in the lowest BM quintiles of value fund portfolios is 22%, and an additional 23% are invested in stocks in the second BM quintile. On the other hand, only 14% are held in high BM stocks. In other words, on average value funds hold a higher fraction of their portfolios in low BM "growth" stocks than in high BM "value" stocks. Figure 3.11 shows the distribution of portfolio shares across mutual funds (in black). Very few mutual funds hold more than 30% or their portfolios in quintiles 4 and 5 stocks. This pattern is particularly stark for "value" funds. Only 41 of 574 "value" funds hold more than 25% of their portfolio in BM quintile-5 stocks (in comparison, 1,082 out of 1,130 "growth" funds hold more than 25% of their portfolio in quintile-1 stocks). In contrast, 209 "value" funds hold more than 25% of their portfolio in quintile-1 stock and 309 "value" funds hold a larger share of their portfolio in quintile-1 than in quintile-5 stocks. In other words, "value" mutual funds hold significantly more "low-BM growth" stocks than "high-BM value" stocks.

Panel A of Table 3.8 also shows the average BM-quintile portfolios shares of hedge funds and ETFs. The quintile shares in BM hedge fund and ETF portfolios look remarkably similar to those of mutual funds. The histograms in Figure 3.11 show that not only the mean of the shares distributions across BM quintiles are similar for mutual funds, hedge funds, and ETFs but also the overall shape.

Panel B reports the portfolio shares across BM quintiles for the five largest value funds. Four of the five largest value funds hold the highest portfolio share in stocks in the lowest BM quintile. Their portfolio shares decline (mostly) monotonically and the lowest portfolio shares are in stocks in the highest BM quintile (with one exception). The largest "value" fund, the "T. Rowe Price Equity Income" fund (with assets of $21.6 bil. as of July 2018), holds 29% of its portfolio in stocks in the lowest BM quintile and only 13% in stocks in the highest BM quintile. Using the BM as a measure of "value", as is done in most of the academic literature, this fund would be labeled as a "growth" fund rather than a "value" fund. The portfolios of the second to fourth largest "value" funds have similar patterns. The notable exception is the fifth largest "value" fund, the "DFA US Large Cap Value" fund. This fund holds very small fractions of stocks in the lowest two BM quintiles and holds on average 70% in stocks in the two highest BM quintiles. In contrast, portfolios of "growth" funds are more concentrated in low BM stocks. Panel C shows the average portfolio weights for the five largest "growth" funds in our sample. These funds hold at least 65% of their portfolios in BM1 and BM2 stocks and the portfolio shares are declining in BM.

3.6 Joint Distribution of Mutual Fund Characteristics

So far, we have focused on the univariate characteristics distributions. Next, we will study the joint distribution of average BM and MOM scores of mutual funds; additional results are reported in the online appendix. Figure 3.12 shows the 2-dimensional scatter plot with BM scores on the x-axis and MOM scores on the y-axis. The plots also show the scores of the CRSP-VW and S&P 500 indices as well as the components of Fama-French portfolios ("S", "B" for small/big, "H", "L" for high/low BM, "U", "D" for high/low MOM).

Panel A shows the BM/MOM distribution for individual stocks. Each dot represents the average BM and MOM characteristics for an individual stock. Smaller/larger dots correspond to smaller/larger stocks. Average MOM scores for most stocks are between 2.75 and 3.75 while average BM score are more spread out. The scatter plot also shows no strong link between BM and MOM scores for stocks. Panel B shows the same plot for mutual funds with different mutual funds types indicated by different colors. The BM/MOM distribution of mutual funds is different from that of stocks in a number of ways. First, it is more clustered around BM scores between 1.2 and 3 and MOM scores between 3 and 4, as already indicated by the univariate BM and MOM histograms. Second, there is a negative correlation between a fund's BM and MOM scores. Funds with low BM scores have higher MOM scores than those with higher BM scores. Hence "growth" funds (in green) have on average a higher MOM score than "value" funds (in blue). The figure shows that there are no funds with a portfolio that is tilted towards high BM *and* high MOM.

Panel C shows the BM/MOM distribution for fund/quarter observations instead of time series averages. Compared to the fund/averages observations shown in Panel B, the fund/quarter MOM scores are more spread out. The bulk of the observations are the BM scores are between 1 and 3. About 60% of all fund/quarter observations have a BM score between 1 and 3 and a MOM score between 3 and 4. It is instructive to compare the mutual fund distribution of S&P 500 stocks. Since stocks are assigned integer scores between 1 and 5 in each quarter, we add some random noise around the integer values to show the distribution. The stock/quarter distribution is superimposed in red. Since the breakpoints in the constructions of portfolios are reset each year, the distribution of stocks is almost uniform. Hence the BM/MOM scores of portfolios of mutual funds are more concentrated than scores of individual (large) stocks. Panel D shows the BM/MOM distribution for hedge funds and ETFs. Both distributions are very similar to the distribution of mutual funds.

The scatter plots in Figure 3.12 show the joint BM/MOM distribution over the entire sample but there is significant time variation in the joint distribution. Figure 3.13 plots the joint distributions in 1999Q1 and 2001Q2. In 1999Q1 mutual funds with low BM scores have high MOM scores. This pattern is reversed in 2001Q2 when low BM funds have lower than average MOM scores. During the stock market boom, growth stocks were on average also high momentum stocks; hence portfolios of "growth" funds, holding primarily low BM stocks, were also high momentum. Momentum of low BM stocks was low after the market correction in 2000, hence "growth" funds were low momentum. Since portfolios of "value" funds are invested in low, mid and high BM stocks in similar proportions, this comovement of

BM and MOM scores are much less pronounced. This pattern holds throughout the sample. The time series standard deviation of MOM scores for low BM funds is twice as high as that of higher BM funds.

3.7 Characteristics and Return

Finally, we investigate how characteristics are related to returns. Table 3.9 reports returns of stocks in Panel A and mutual funds in Panel B across characteristic quintiles. In addition to ME, BM, MS, and MOM we also report results for the two components of the MS index: Fundamental growth rates (GR) and multiples (MULT). The returns across size, book-to-market and momentum have the familiar patterns of the size, value and momentum premia. The Morningstar value index MS produces a significantly smaller return spread than the book-to-market ratio. The reason is that both MS components yield relatively small return spreads. There is no consistent return pattern across GR quintiles, and the index of multiples produces a smaller return spread than the book-to-market ratio by itself. The corresponding results for mutual funds are reported in Panel B. The mutual fund returns are after fees and overall lower than those of individual stocks but have lower volatility. Sharpe-ratios of stock and mutual fund returns are comparable.

While the well-studied characteristic premia are present in stock returns, they are much smaller on the mutual fund level. There are no consistent return patterns across ME, MOM, MS, MULT and GR quintiles. Only the book-to-market effect is present in mutual funds returns, but its magnitude is smaller than that for stocks. The BM quintile-5 to quintile-1 spread for stocks is 2.82%, and 0.78% for mutual funds, respectively. Hence, investors in mutual funds are not rewarded for return premia associated with characteristics that are present in individual stock returns.

Next, study the characteristics-return link more formally using Fama-MacBeth regressions. In each quarter t we estimate the regression

$$R_{i,t+1} - R_{f,t+1} = \boldsymbol{\beta}_t' \mathbf{X}_{i,t} + e_{i,t+1}, \tag{3.2}$$

where $R_{i,t+1} - R_{f,t+1}$ is the excess return of asset i in quarter $t+1$ and $\mathbf{X}_{i,t}$ is a vector of characteristics of asset i at time t. Then, we time-average the betas and report $\overline{\boldsymbol{\beta}} = \sum_t \boldsymbol{\beta}_t$ in Table 3.10. We estimate the model for individual stocks and mutual funds. The results for the sample of individual stocks in the top panel shows the familiar patterns. Stocks that are small size have high momentum and high "value" are associated with higher returns. The size effect is statistically insignificant while the momentum coefficient is significant. The BM coefficient is twice as large as the MS coefficient and strongly significant while the significance of the MS estimate is only marginal. This suggests that on the stock level the book-to-market ratios are more powerful return predictor than the Morningstar index, consistent with the results in Table 3.9.

The results for the sample of mutual funds, shown in Panel B, differ from those for individual stocks in important aspects. The ME coefficients are slightly more negative than those for the stock sample while the MOM point estimates are almost identical. The BM and MS coefficients, however, are slightly negative and insignificant. This in sharp contrast to the estimates for individual stocks. The "value" premium is present in individual stocks returns but not in returns of mutual funds, which is also consistent with the results in Table 3.9.

3.8 Conclusion

This paper provides a comprehensive analysis of characteristics of mutual fund portfolios. Some facts stand out. First, the BM distribution of mutual funds is strongly skewed towards low BM ratios. While there are many funds that have a BM ratio comparable to that of the "L" portfolio in HML, there very no funds with a BM ratio close to "H". Moreover, the skew towards low BM values is more pronounced for mutual funds than for individual (large) stocks. Second, "growth" funds hold almost exclusively low BM stocks in their portfolios. In contrast, portfolios of "value" funds include stocks across the entire BM distribution. In fact, on average mutual funds hold a higher share of stocks with low BM ratios that stocks with high BM ratios. The BM distributions of ETFs and hedge funds are similar to that of mutual funds. Third, mutual funds are on average almost momentum-neutral. While momentum of "growth" funds varies over time, in contrast to momentum of "value" funds, there are very few mutual funds with consistently high momentum. Fourth, size, book-to-market and momentum return spreads are smaller for mutual funds than for individual stocks and insignificant in Fama-MacBeth regressions.

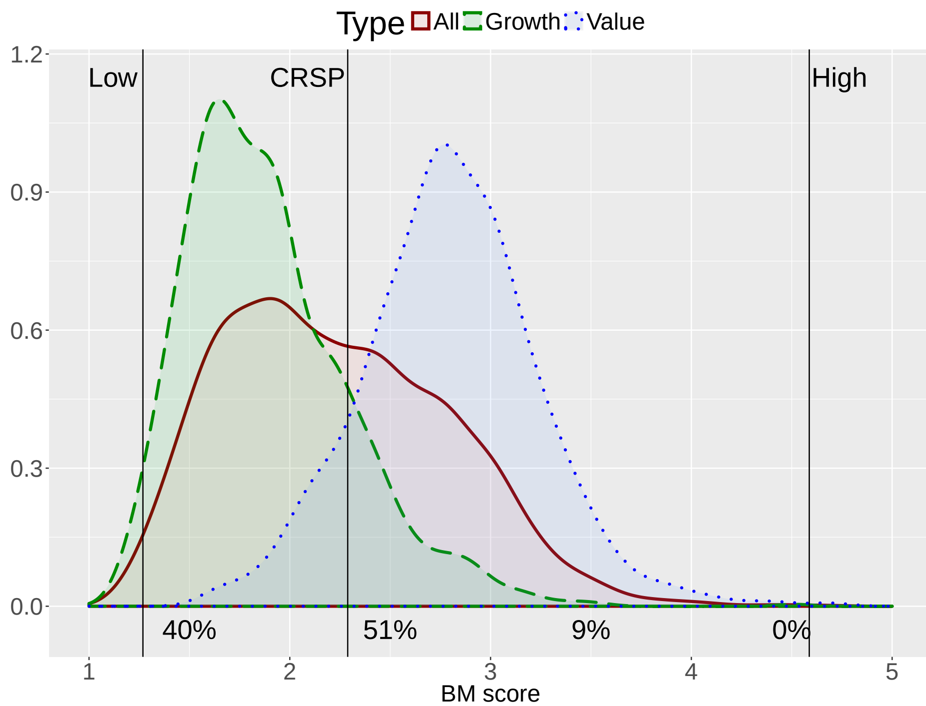
These stylized facts raise a number of questions about active mutual funds:

1. Why is the distribution of mutual fund portfolios so strongly tilted towards low book-to-market ratios and why are there virtually no high BM funds at all even though high BM stocks are associated with higher returns than low BM stocks?
2. Why do funds that label themselves as "value" funds hold more low BM stocks than high BM stocks while "growth" funds hold almost exclusively low BM stocks?
3. Why are portfolios of active mutual funds not more tilted towards characteristics that are associated with high returns, i.e. small, high BM and high momentum stocks?
4. Why don't mutual funds combine multiple strategies (e.g., high BM - high momentum) that have been shown to be more profitable than univariate strategies (Asness, Moskowitz, and Pedersen (2013)).
5. Why do mutual funds and ETFs follow strategies that emulate the Morningstar value/growth definition even though it has no return premium?

Our results have also broader implications for equity markets. Aside from the issue of delisting of funds and the implied survivorship bias, the literature takes the set of mutual funds as given and there is little research about why new funds are created. In other words, what economic forces determine the set of funds and strategies that we observe? Is the mutual fund market driven by investor's demand for certain strategies or by the supply of profitable strategies? Are there so many "growth" funds because investor's demand for "growth" stocks and the absence of high-BM funds is due to low demand? How can the stylized facts presented in this paper be reconciled with the evidence that capital flows react strongly to past performance? Since returns of high-BM stocks are on average higher than returns of low BM stocks, capital should flow from low-BM funds into high-BM mutual funds over the sample and the number of high-BM funds should increase relative to the number of low-BM funds. Yet, there is no evidence support this conjecture.

Portfolios of active mutual funds account for about 13% of total market cap (as of 2016) and their portfolio allocations are likely to have an effect on equilibrium prices. Whether factor premia are permanent or diminishing over time due to higher demand for underpriced stocks is still an open question. Our results suggest that active mutual funds do not systematically hold the stocks with characteristics associated with high returns and thus are unlikely to contribute to any shrinking of factor premia during the sample period. Our sample of mutual funds and ETFs is exhaustive but we only observe portfolio holdings of a very small subset of small hedge funds, so we cannot rule out that (larger) hedge funds tilt their portfolios towards profitable characteristics.

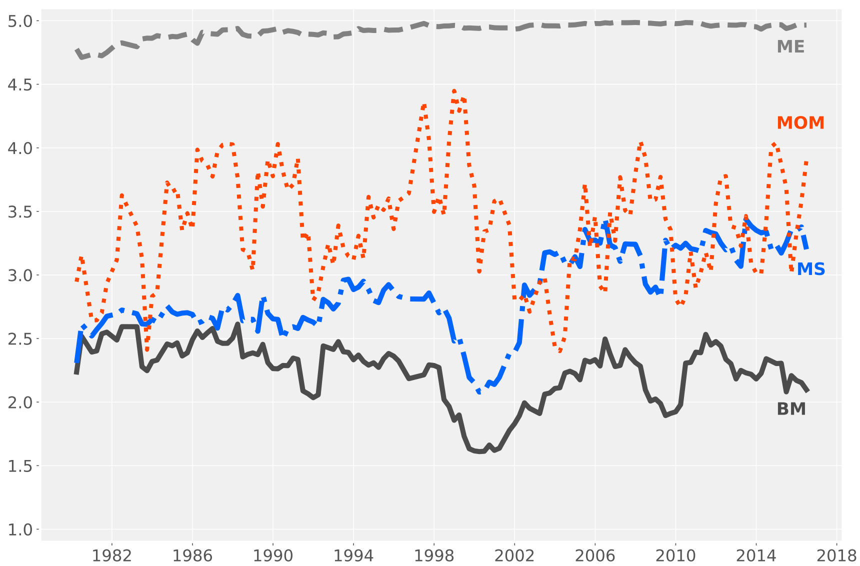
Figure 3.1: Distribution of Book-to-Market Ratios of Mutual Funds



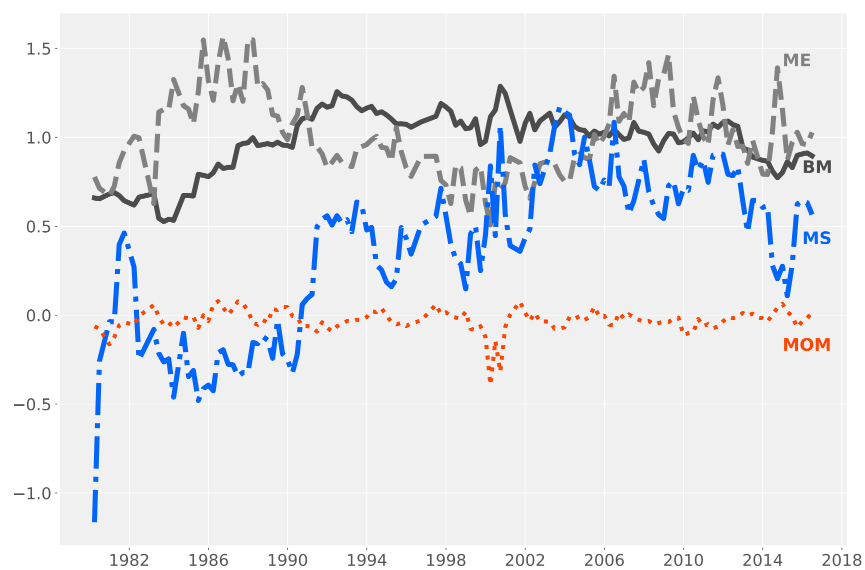
Notes: See Figure 3.3.

Figure 3.2: Characteristics of "The Investment Company of America Fund" (AIVSX)

Panel A: Characteristics scores

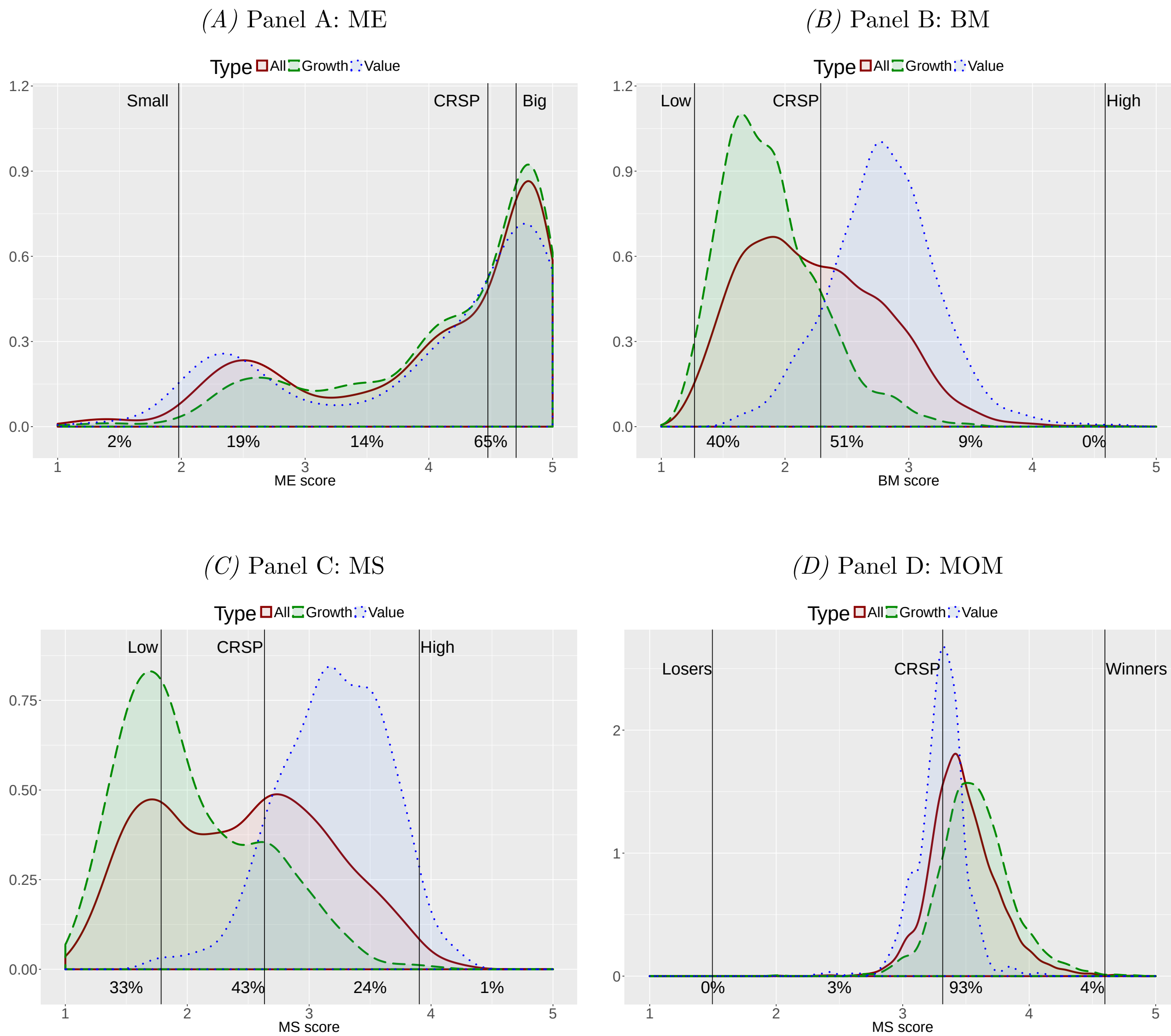


Panel B: Market-adjusted characteristics



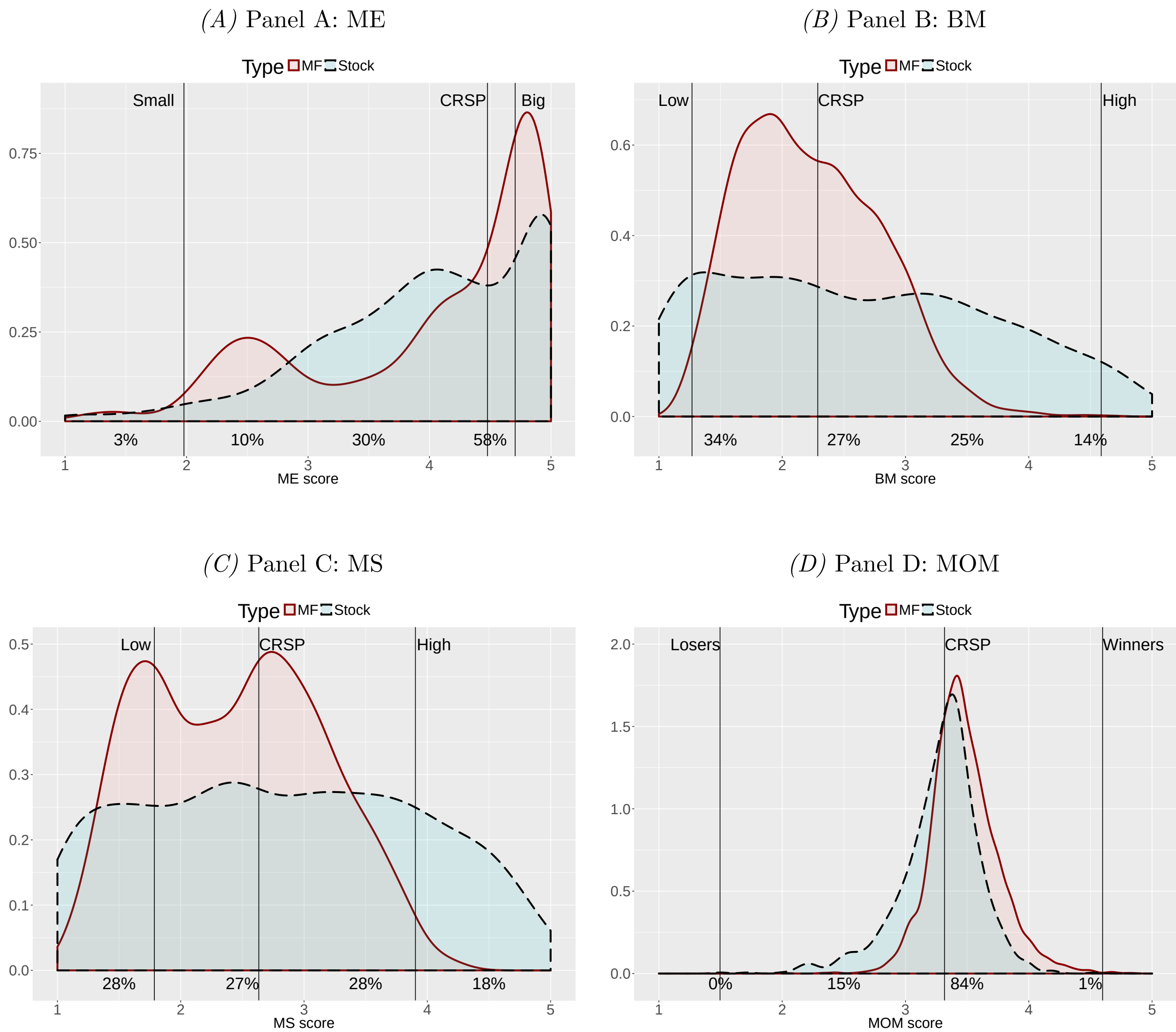
Note: This shows the time series of ME, BM, MS and MOM characteristics of the "The Investment Company of America Fund" (AIVSX) mutual fund. Panel A shows the characteristic scores. The market-adjusted characteristics are plotted in Panel B. Adjusted MS is divided by 10.

Figure 3.3: Characteristics of Mutual Funds



Note: This plot shows the histograms of the distribution of average size (ME), book-to-market (BM), Morningstar index (MS) and momentum (MOM) characteristics of mutual funds over the periods that the fund is in the sample. In each quarter, the fund characteristics are computed as the value-weighted averages of scores of holdings of the fund. The scores are computed using Fama-French quintile breakpoints. An index of '1' indicates firms in the lowest B/M quintile and firms with a score of '5' are in the highest B/M quintile. The solid black line is the histogram of all mutual funds, the dashed green line is for 'growth' funds fund and the dashed blue line is for 'value' funds. The vertical lines indicate the average score of the CRSP-VW index and the corresponding "high" and "low" portfolios of Fama-French long/short portfolios. The sample is from 1980Q1 to 2016Q2.

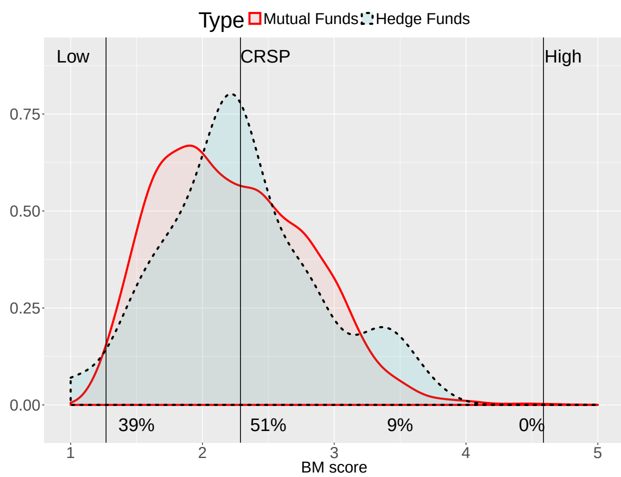
Figure 3.4: Histograms - Characteristics of Mutual Funds and Stocks



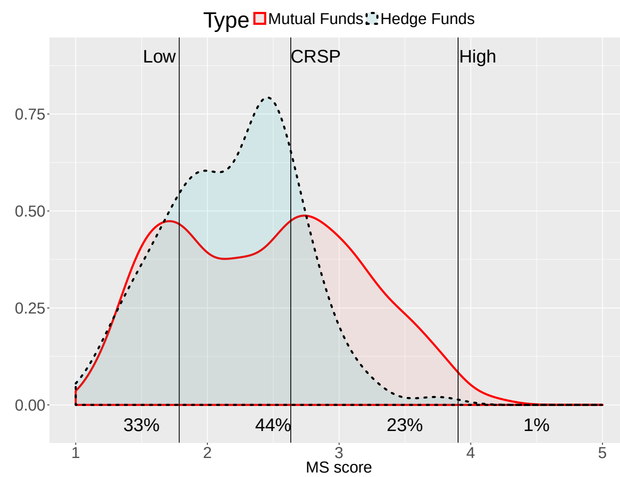
Note: This plot shows the histograms of the distribution of average size (ME), book-to-market (BM), Morningstar index (MS) and momentum (MOM) characteristics of mutual funds over the periods that the fund is in the sample as well as the histogram of average characteristics for individual S&P500 stocks. In each quarter, the fund characteristics are computed as the value-weighted averages of scores of holdings of the fund. The scores are computed using Fama-French quintile breakpoints. An index of '1' indicates firms in the lowest B/M quintile and firms with a score of '5' are in the highest B/M quintile. The solid black line is the histogram stocks and the dashed line is for mutual funds. The vertical lines indicate the average score of the CRSP-VW index and the corresponding "high" and "low" portfolios of Fama-French long/short portfolios. The sample is from 1980Q1 to 2016Q2.

Figure 3.5: Histograms - Characteristics of Hedge Funds and ETFs

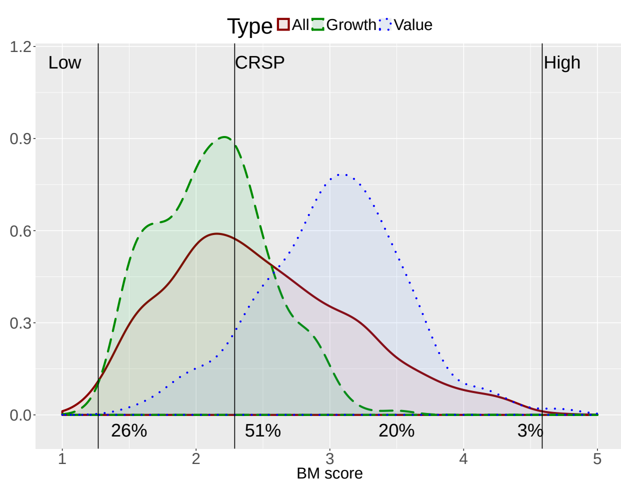
(A) Panel A: HF BM



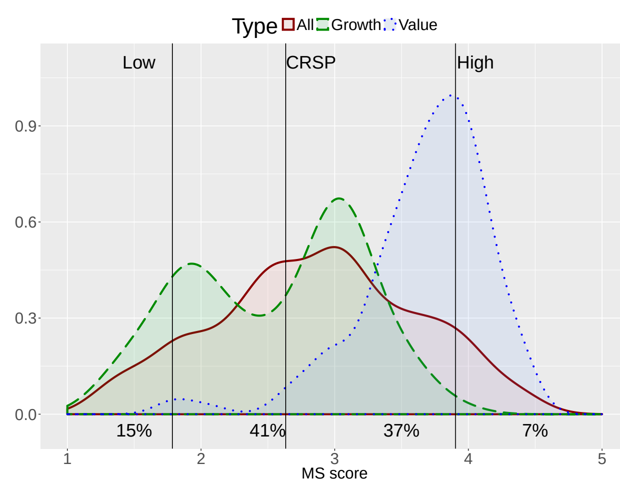
(B) Panel B: HF MS



(C) Panel C: ETF BM



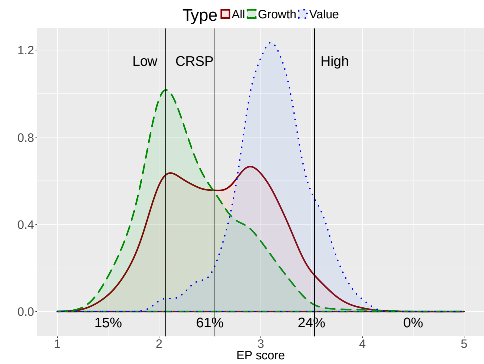
(D) Panel D: ETF MS



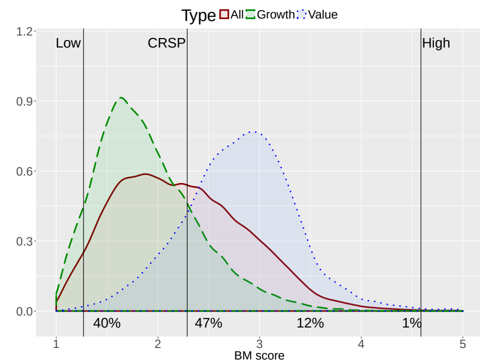
Note: This plot shows the histograms of the distribution of average book-to-market (BM) and Morningstar index (MS) characteristics of ETFs and hedge funds. In each quarter, the fund characteristics are computed as the value-weighted averages of scores of holdings of the fund. The scores are computed using Fama-French quintile breakpoints. An index of '1' indicates firms in the lowest B/M quintile and firms with a score of '5' are in the highest B/M quintile. Panels A and B are ETF histograms (for all, 'value' and 'growth' ETFs) and Panels C and D are HF histograms (dashed lines and solid lines for mutual funds). The vertical lines indicate the average score of the CRSP-VW index and the corresponding "high" and "low" portfolios of Fama-French long/short portfolios. The sample is from 1980Q1 to 2016Q2.

Figure 3.6: Characteristics of Mutual Funds – Robustness

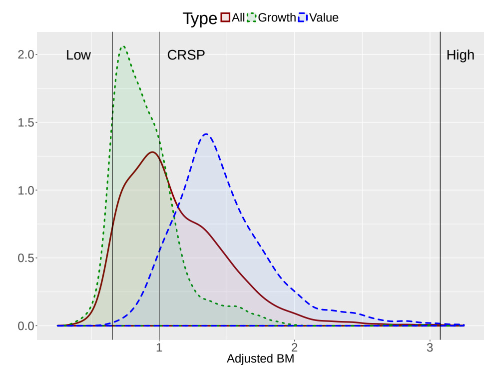
(A) EP



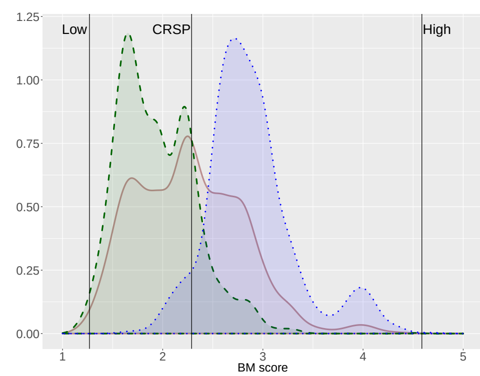
(B) Fund/Date Obs.



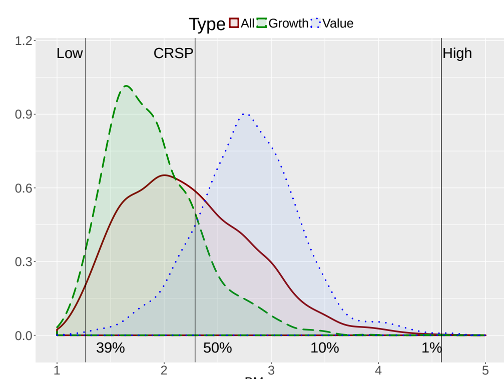
(C) Adjusted BM



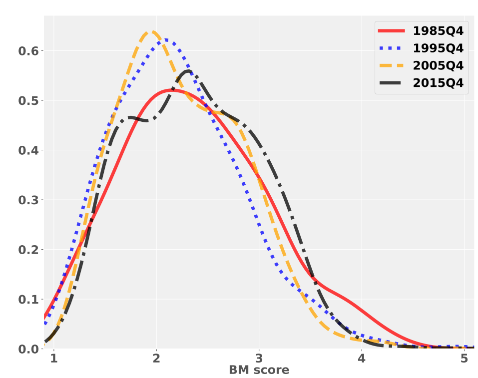
(D) AUM-weighted



(E) Incl. index/sector MFs



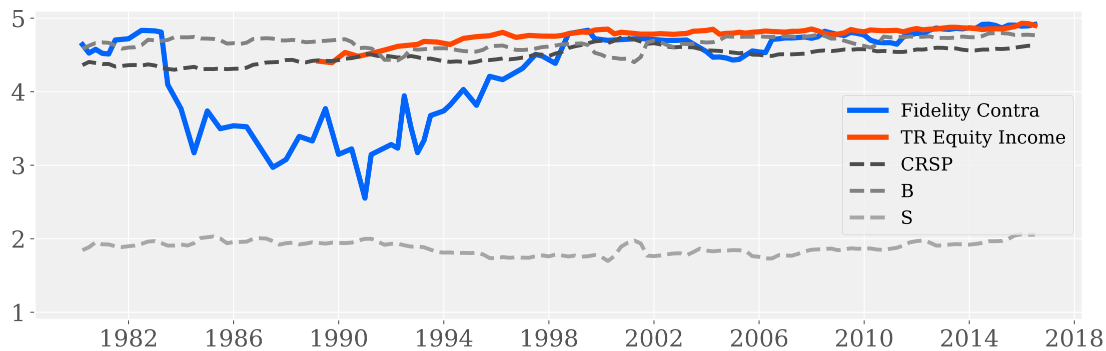
(F) Subsamples



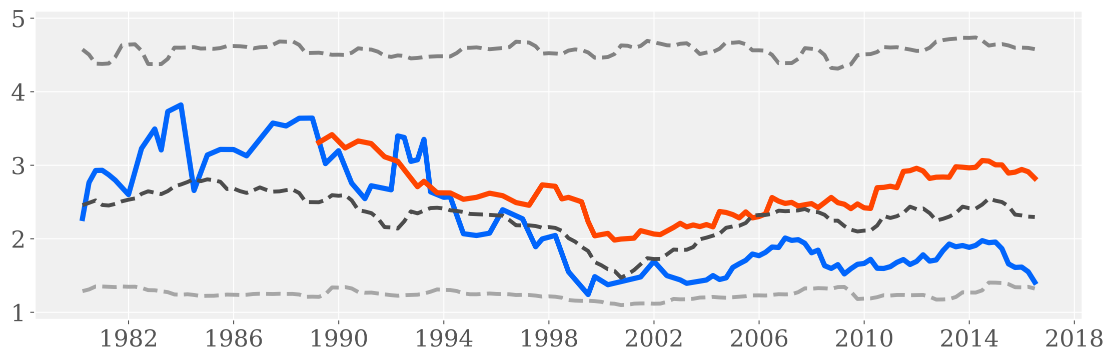
Note: This plot shows the histograms of the characteristic distributions of the earnings/price (EP) ratio, fund/date BM, the adjusted BM ratio, AUM-weighted BM, BM for all mutual funds, including index and sector funds, and the BM in four quarters. The sample is from 1980Q1 to 2016Q2.

Figure 3.7: Time-series of Characteristics of two large Mutual Funds

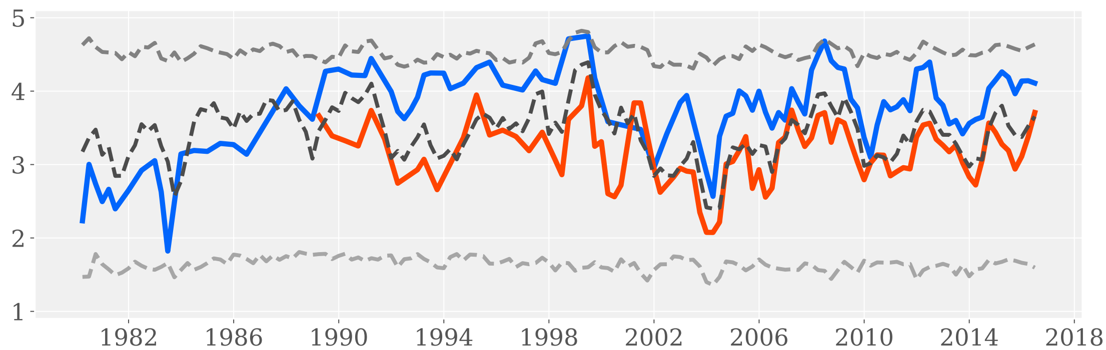
(A) Panel A: ME



(B) Panel B: BM



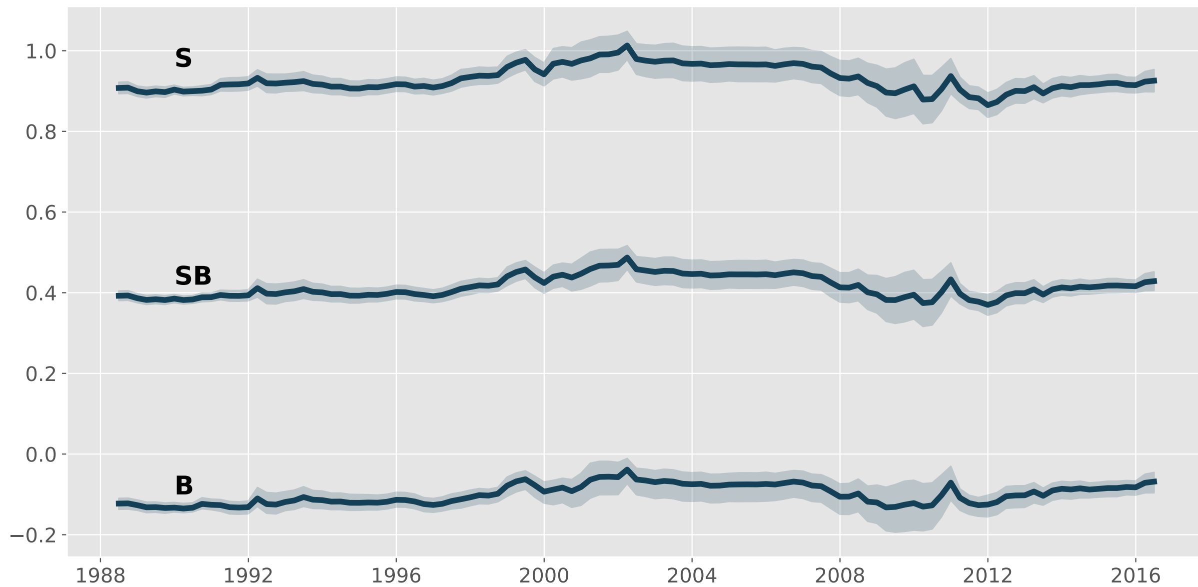
(C) Panel C: MOM



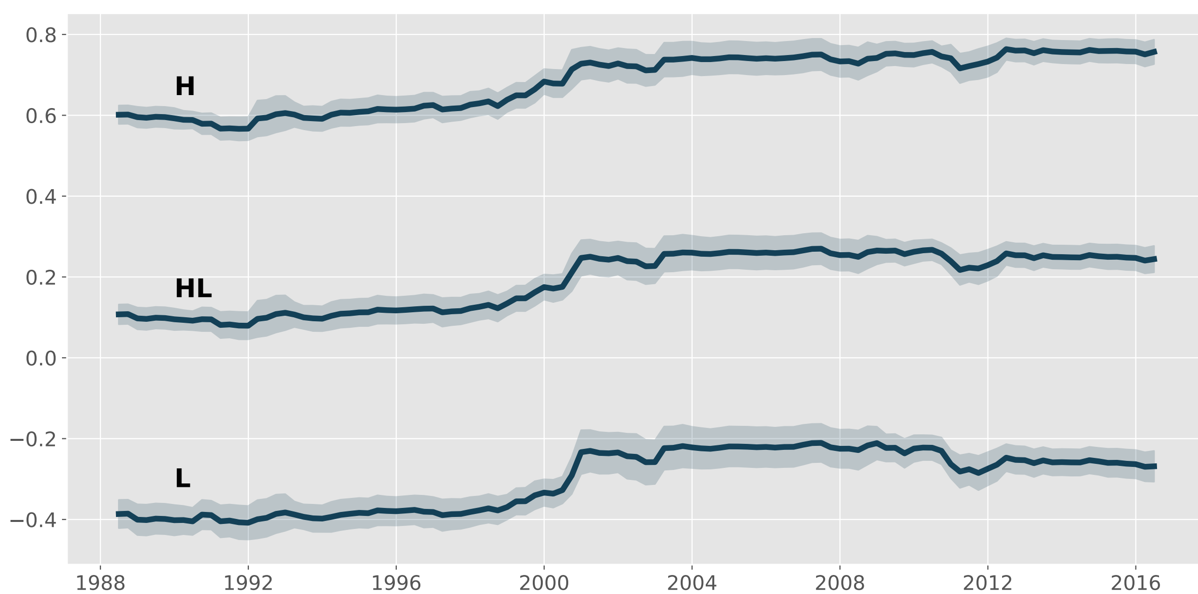
Note: This plot shows time-series of size (ME), book-to-market (BM) and momentum (MOM) characteristics of the the largest mutual fund in our sample (Fidelity Contrafund), the largest value fund (Fidelity Equity Income Fund), the CRSPVW index and the Fama-French long/short portfolios. The sample is from 1980Q1 to 2016Q1.

Figure 3.8: Loadings in Rolling 4-Factor Regressions

Panel A: SMB-β



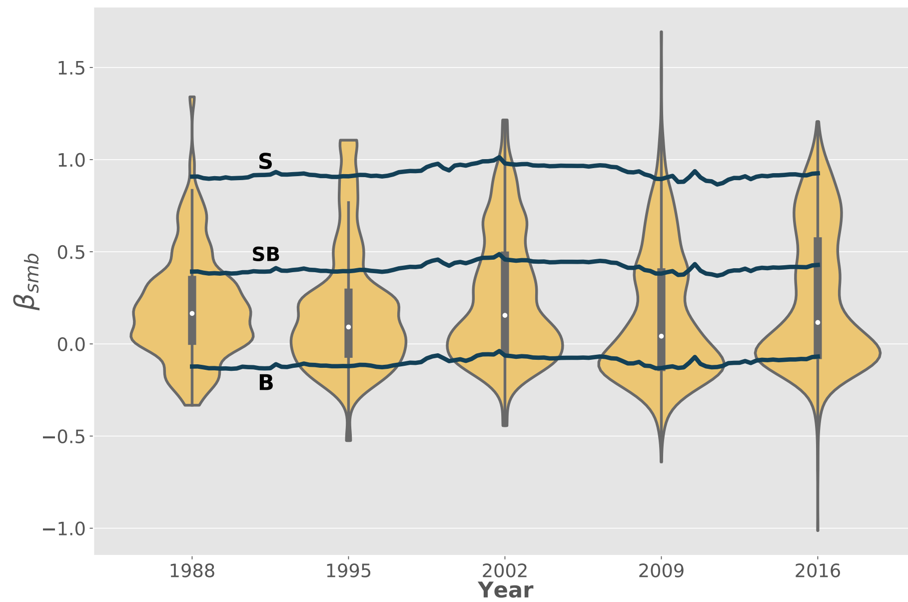
Panel B: HML-β



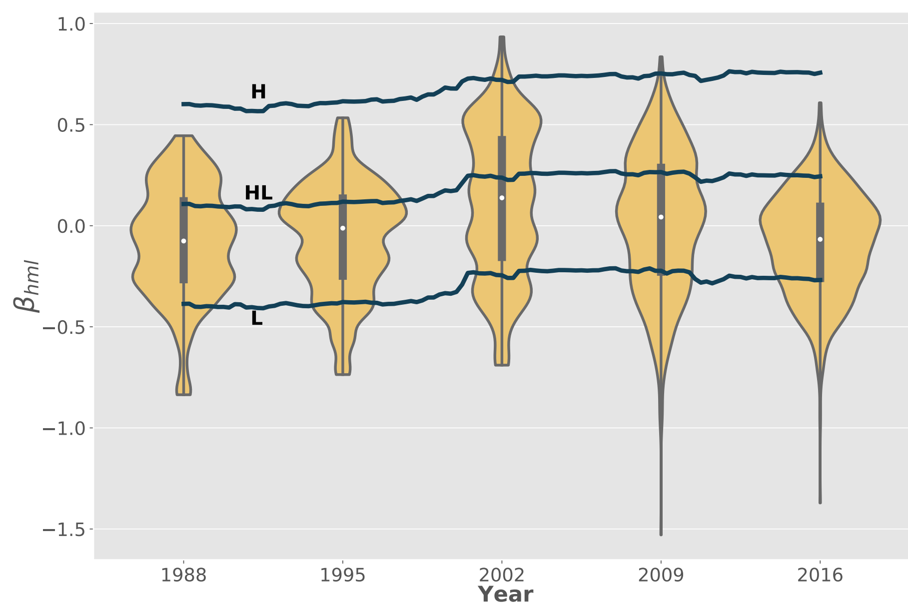
Note: This figure shows loadings in rolling window regressions of fund excess returns on the market excess return (MKT), SMB and HML. The windows size is 60 quarters.

Figure 3.9: HML and SMB Loadings in Rolling 4-Factor Regressions

Panel A: SMB-β



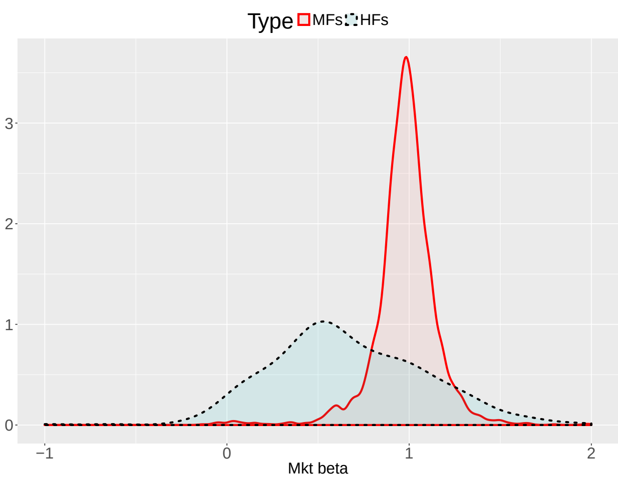
Panel B: HML-β



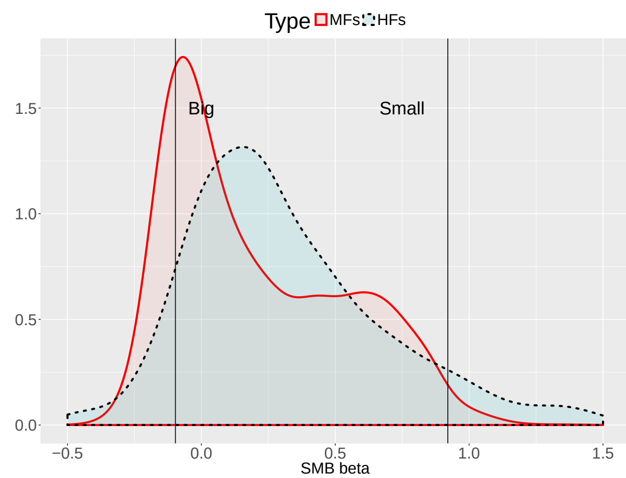
Note: This figure shows box plots of HML and SMB loadings in rolling window regressions of fund excess returns on the market excess return (MKT), SMB and HML. The windows size is 60 quarters. The years on the x-axis indicate the end-year of a window. The box plots show deciles as well as the median for all mutual funds that are included in a window. The lines indicate the rolling β estimates for “H”,“L”, “S” and “B”

Figure 3.10: Histograms - Loadings of Mutuals Funds and Hedge Funds

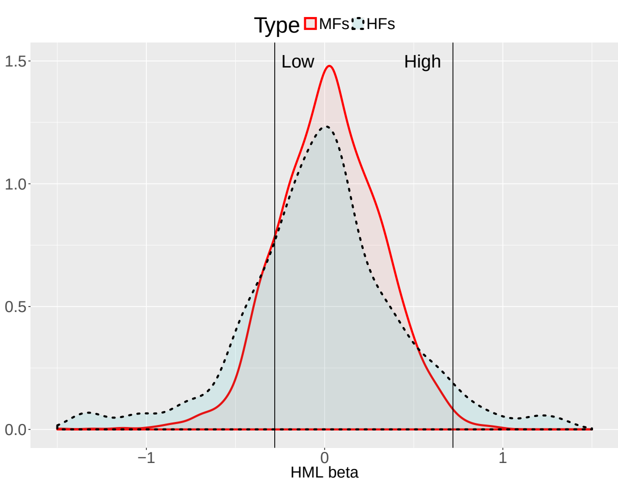
(A) Panel A: MKT



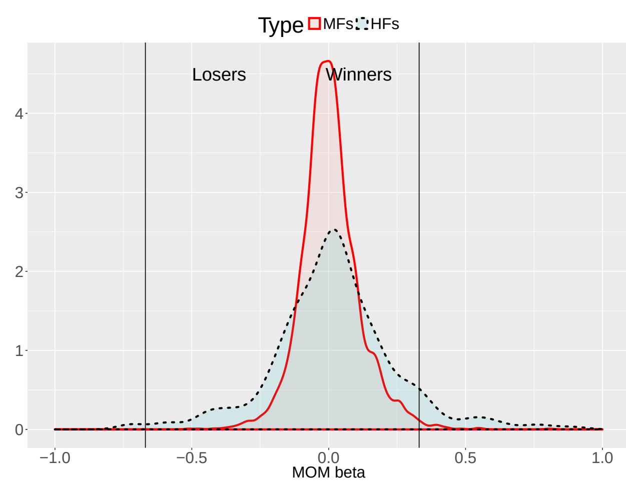
(B) Panel B: SMB



(C) Panel C: HML



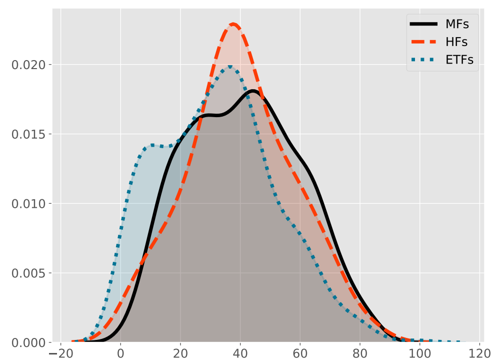
(D) Panel D: MOM



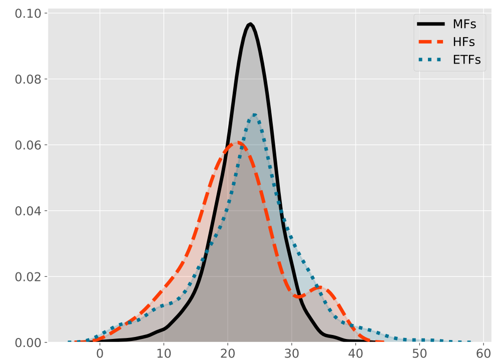
Note: This plot shows the histograms of β's of mutual funds and hedge funds in 3-factor regressions of fund excess returns on the market excess returns, SMB, HML and MOM. Hedge funds returns are from Hedge Fund Research (HFR). The vertical lines indicate β's of the components of SMB, HML and MOM The sample is from 1980Q1 to 2016Q2.

Figure 3.11: BM-Quintile Portfolio Shares of Mutual Funds, Hedge Funds and ETFs

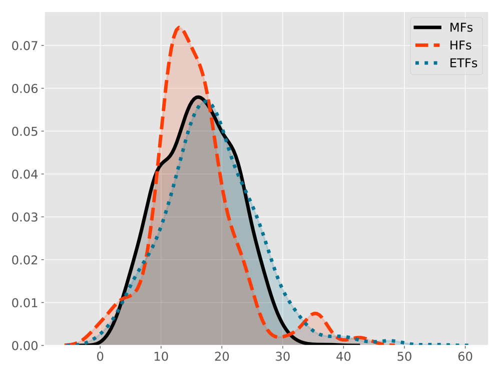
(A) Panel A:
BM Quintile 1



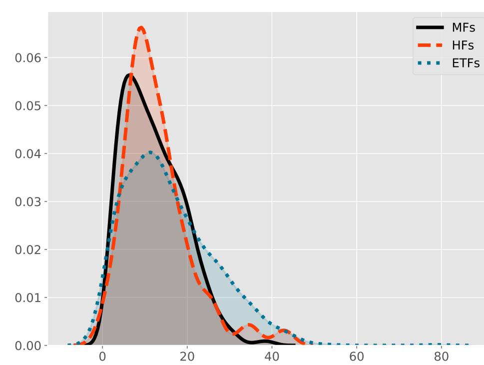
(B) Panel B:
BM Quintile 2



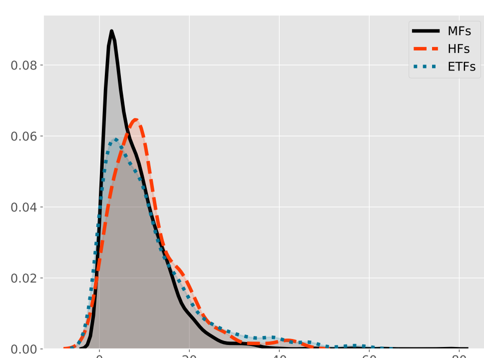
(C) Panel C:
BM Quintile 3



(D) Panel D:
BM Quintile 4



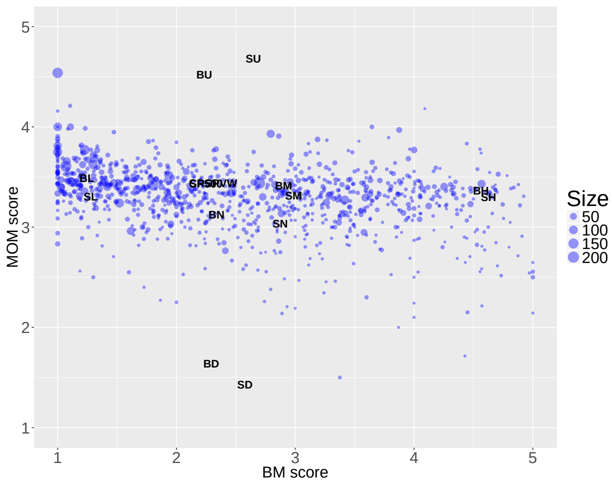
(E) Panel E:
BM Quintile 5



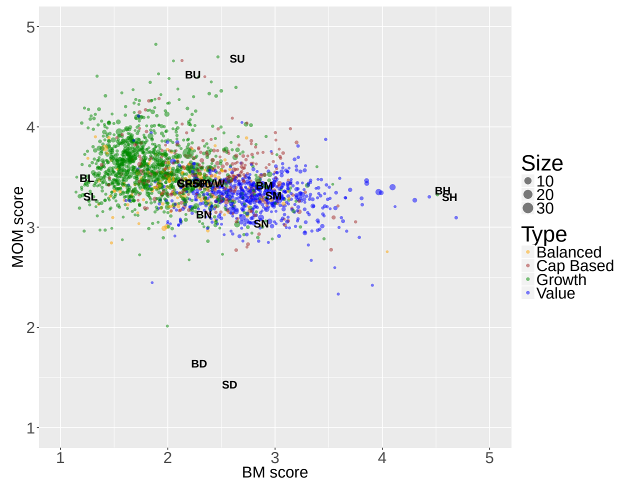
Note: The figure shows the distributions of portfolio shares in BM quintiles 1 to 5 for mutual funds, hedge funds and ETFs.

Figure 3.12: Joint Characteristics Distributions

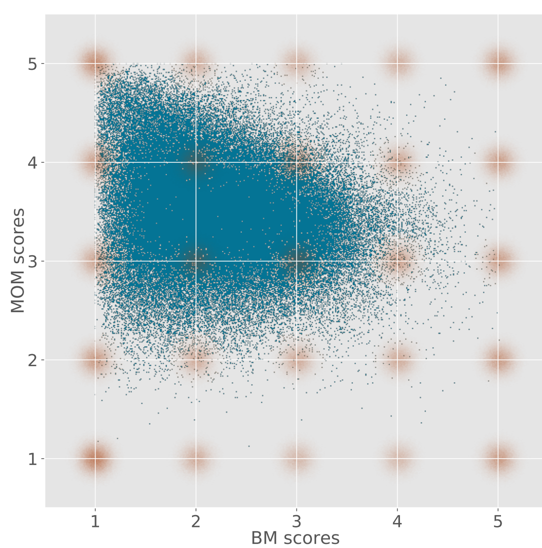
(A) Panel A: BM/MOM of S&P 500 Stocks



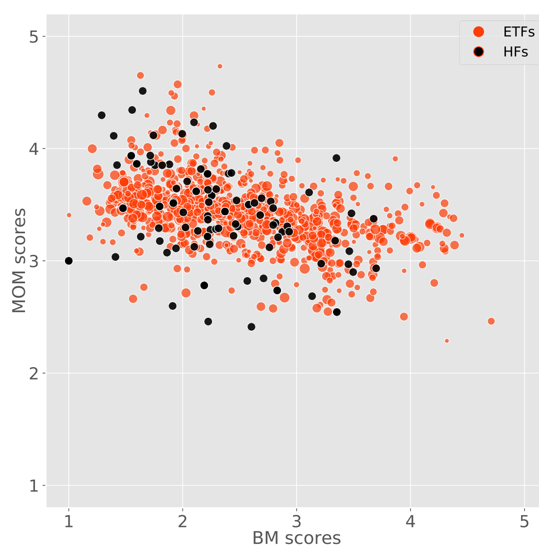
(B) Panel B: BM/MOM



(C) Panel C: BM/MOM of Stock and MF/quarter



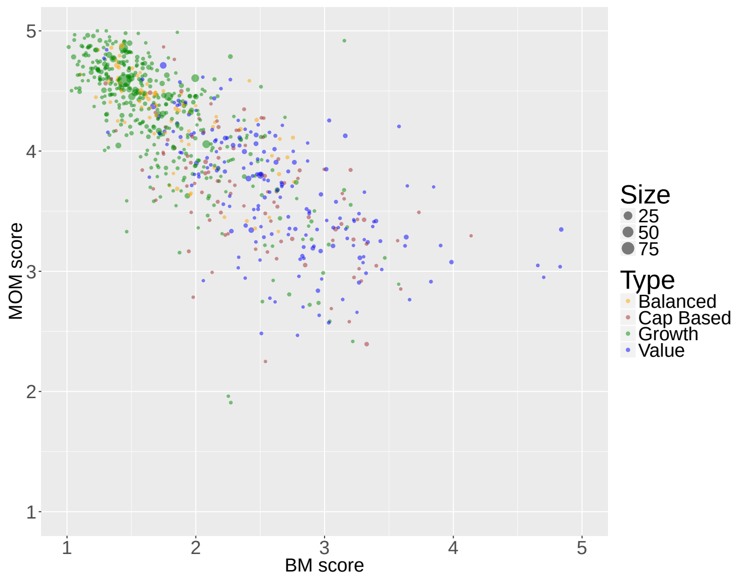
(D) Panel D: BM/MOM of HFs and ETFs



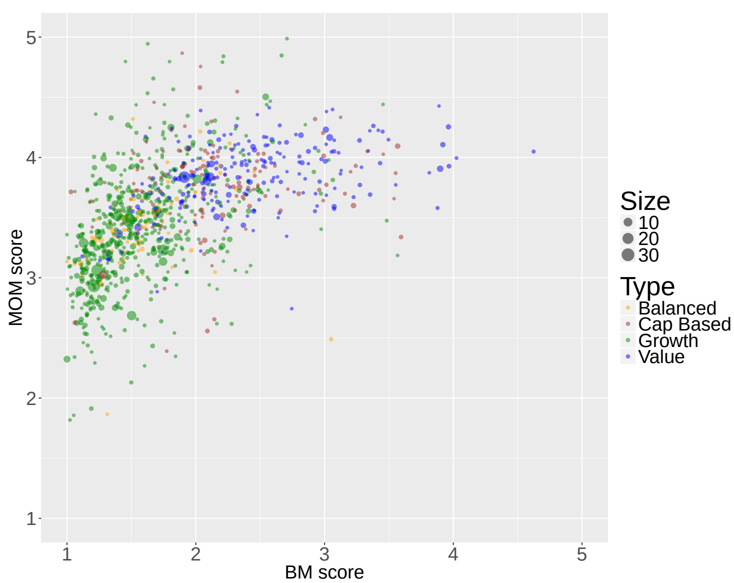
Note: This figure shows scatter plots of characteristics. Panels A and B show the BM/MOM and MS/MOM distributions of mutual funds, respectively. Panels C and D show the BM/MOM distributions of ETFs and hedge funds (Panel C) and stocks (Panel D). The sample is from 1980Q1 to 2016Q2.

Figure 3.13: Joint Characteristics Distributions: 1999Q1 and 2001Q2

(A) Panel A: 1999Q1



(B) Panel B: 2001Q2



Note: This figure shows scatterplots of the BM/MOM distribution of mutual funds in 1999Q1 and 2001Q2.

Table 3.1: Descriptive Statistics of Mutual Funds and Hedge Funds

	Mutual Funds				ETFs	HFs
	All	Value	Growth	Other		
Number of funds	2,638	574	1,130	934	955	114
AUM 12/2014 ($ bil.)	2,143	416	927	799	1,124	53
Median size ($ mil.)	149	145	150	150	167	NA
Median age (years)	11.58	9.88	11.83	12.56	5.42	NA
Median no. of stocks	54	56	51	56	99	64
Median Return over S&P 500 (% p.a.)	-0.70	-0.41	-0.74	-0.81	0.76	NA
Median 4-Factor α (% p.a)	-0.35	-0.04	-0.45	-0.41	-0.36	NA

Note: Descriptive statistics of mutual funds, ETFs and hedge funds.

Table 3.2: Characteristics of Passive Benchmark Portfolios

	CRSP-VW	BM				MOM			
		SL	BL	SH	BH	SD	BD	SU	BU
ME score	4.50	2.07	4.80	1.88	4.66	1.86	4.60	2.02	4.73
BM score	2.31	1.28	1.25	4.63	4.56	2.58	2.29	2.65	2.23
MS score	2.74	1.53	2.00	3.61	4.08	2.33	2.75	2.29	2.53
MOM score	3.44	3.30	3.49	3.30	3.36	1.43	1.64	4.68	4.52
Adjusted ME	1.00	0.02	1.29	0.01	0.65	0.01	0.69	0.02	1.04
Adjusted BM	1.00	0.66	0.62	3.31	2.98	1.46	1.12	1.54	1.04
Adjusted MS	0.00	−28.31	−14.95	18.48	26.23	−9.31	−0.69	−9.83	−5.44
Adjusted MOM	0.00	0.08	0.01	0.03	−0.04	−0.48	−0.41	0.53	0.30

Note: This table shows average characteristic scores and adjusted characteristics of the CRSP-VW index and passive Fame-French portfolios. “SL” is the small/low-BM portfolio, “BL” is the big/low-BM portfolio, etc.

Table 3.3: Characteristics of Mutual Funds and Hedge Funds

	Stocks					**All Funds**				
	mean	10th	25th	75th	90th	mean	10th	25th	75th	90th
ME score	4.01	3.47	4.80	5.00	0.31	4.02	2.42	3.38	4.81	4.91
BM score	2.62	1.69	3.46	4.17	0.69	2.23	1.55	1.78	2.63	2.99
MS score	2.82	1.89	3.71	4.36	0.82	2.44	1.51	1.82	2.98	3.42
MOM score	3.28	3.13	3.47	3.62	1.12	3.48	3.19	3.31	3.63	3.83
	Value Funds					**Growth Funds**				
	mean	10th	25th	75th	90th	mean	10th	25th	75th	90th
ME score	3.99	3.06	4.83	4.91	0.18	4.15	2.72	3.76	4.82	4.91
BM score	2.80	2.52	3.06	3.32	0.29	1.89	1.43	1.59	2.11	2.42
MS score	3.20	2.89	3.54	3.76	0.29	2.02	1.38	1.59	2.42	2.87
MOM score	3.30	3.21	3.41	3.50	0.38	3.58	3.26	3.41	3.74	3.93
	Hedge Funds					**All 13F Institutions**				
	mean	10th	25th	75th	90th	mean	10th	25th	75th	90th
ME score	3.64	3.15	4.35	4.65	0.35	4.11	2.69	3.73	4.78	4.90
BM score	2.29	1.91	2.62	3.19	0.34	2.25	1.57	1.85	2.55	3.06
MS score	2.19	1.82	2.52	2.78	0.35	2.42	1.54	2.01	2.82	3.24
MOM score	3.42	3.18	3.64	3.94	0.48	3.43	2.98	3.25	3.64	3.89

Note: The table reports the percentiles of the distributions of average characteristic scores for our sample of individual stocks, mutual funds, hedge funds and all 13F institutions.

Table 3.4: Characteristics of highest/lowest BM Mutual Funds

Fund	BM	MS	MOM	ME	AUM ($ mil.)
10 Highest BM Funds					
High BM portfolio "H"	4.59	3.90	3.30	3.25	NA
Aegis Value Fund	4.69	3.56	3.09	1.36	276
Mellon Capital S&P SMid 60	4.51	3.89	3.33	2.69	400
Franklin MicroCap Value Fund	4.44	3.45	3.30	1.11	285
Franklin Balance Sheet Investment Fund	4.30	3.77	3.27	2.89	1887
First Trust Dow Target Dividend	4.12	4.23	3.20	3.73	20
DFA US Small Cap Value Portfolio	4.10	3.23	3.40	1.88	5925
Ancora Special Opportunity Fund	4.05	3.05	2.75	1.94	7
DFA US Targeted Value Portfolio	3.99	3.74	3.39	4.74	306
SA US Value Fund	3.99	3.33	3.34	2.51	1849
DFA US Large Cap Value Portfolio	3.96	3.77	3.35	4.68	6307
10 Lowest BM Funds					
Low BM portfolio "L"	1.27	1.79	3.30	3.44	NA
AmSouth Capital Growth Fund	1.14	1.44	3.35	4.93	19
Excelsior Optimum Growth Fund	1.15	1.31	3.50	4.96	16
Armada Tax Managed Equity Fund	1.18	1.78	3.19	5.00	190
Jensen Quality Growth Fund	1.18	2.10	3.31	4.78	1374
Pioneer Papp Strategic Growth Fund	1.20	1.49	3.40	4.73	129
IAI Emerging Growth Fund	1.20	1.02	4.04	3.14	260
Bender Growth Fund	1.20	1.04	3.38	4.24	15
JPMorgan Equity Growth Fund	1.20	1.50	3.62	4.93	120
American Performance Growth Equity Fund	1.21	1.74	3.49	4.94	94
JNL/S&P Competitive Advantage Fund	1.21	2.34	3.32	4.70	1161

Note: THis table reports characteristics scores of the the 10 mutual funds with the highest BM scores as well as the scores of the 10 funds with the lowest BM scores.

Table 3.5: Probit Regressions

	All	Growth	Value
ME score	0.263***	0.382***	0.250***
	(0.014)	(0.020)	(0.016)
MOM score	0.084***	0.116***	0.032***
	(0.011)	(0.016)	(0.010)
BM score	-0.028**	-0.066***	0.022
	(0.011)	(0.014)	(0.016)
MS score	-0.024**	-0.135***	0.096***
	(0.011)	(0.015)	(0.012)
Observations	1,095,648	478,668	211,536
No. stocks	1356	1356	1356
No. funds	808	353	156
Pseudo R2	0.0616	0.130	0.0489

Note: This table shows result for the Probit model

$$P(y_{i,j,t}) = \Phi(\mathbf{X}'_{i,t}\boldsymbol{\beta}),$$

where $y_{i,t}$ is an indicator variable that is 1 if stock i is held by mutual fund j in quarter t and zero otherwise. $\mathbf{X}_{i,t}$ is a vector of ME, MOM, BM and MS characteristics of stock i in period t.

Table 3.6: 4-Factor Regressions of Passive Benchmark Portfolios

	S	B	H	L	U	D
α	0.01	0.01	0.01	0.01	0.01	0.01
MKT	1.01	1.01	1.04	1.04	1.05	1.05
SMB	**0.90**	**-0.10**	0.41	0.41	0.51	0.51
HML	0.26	0.26	**0.72**	**-0.28**	0.05	0.05
UMD	0.00	0.00	-0.01	-0.01	**0.34**	**-0.66**

Note: The tables reports coefficients of the regression

$$X_t = \alpha_{\mathrm{X}} + \beta_{\mathrm{X,MKT}}\,\mathrm{MKT}_t + \beta_{\mathrm{X,SMB}}\,\mathrm{SMB}_t + \beta_{\mathrm{X,HML}}\,\mathrm{HML}_t + \beta_{\mathrm{X,MOM}}\,\mathrm{MOM}_t + e_{\mathrm{X},t},$$

where $X \in \{\mathrm{S, B, H, L, U, D}\}$. The sample is from 1980Q1 to 2016Q2.

Table 3.7: Loadings of 25 ME-BM sorted Portfolios

	BM1	BM2	BM3	BM4	BM5
			SMB Betas		
ME1	1.37	1.30	1.13	1.13	1.15
ME2	0.90	0.91	0.74	0.72	0.86
ME3	0.65	0.62	0.51	0.45	0.57
ME4	0.47	0.33	0.31	0.21	0.27
ME5	-0.29	-0.15	-0.23	-0.23	-0.23
			HML Betas		
ME1	-0.41	0.02	0.26	0.49	0.70
ME2	-0.45	0.06	0.41	0.61	0.82
ME3	-0.45	0.16	0.42	0.60	0.79
ME4	-0.42	0.21	0.42	0.50	0.72
ME5	-0.33	0.12	0.31	0.64	0.62

Note: The tables reports SMB and HML coefficients of the regression

$$R_{i,t} = \alpha_{\mathrm{X}} + \beta_{\mathrm{X,MKT}}\,\mathrm{MKT}_t + \beta_{\mathrm{X,SMB}}\,\mathrm{SMB}_t + \beta_{\mathrm{X,HML}}\,\mathrm{HML}_t + \beta_{\mathrm{X,MOM}}\,\mathrm{MOM}_t + e_{\mathrm{X},t},$$

for 25 size/BM double-sorted portfolios. The sample is from 1980Q1 to 2016Q2..

Table 3.8: Portfolio Composition of Mutual Funds by Quintiles

	BM1	BM2	BM3	BM4	BM5
	Panel A: Funds				
CRSP-VW	39.01%	21.86%	16.93%	13.49%	8.70%
All MFs	40.96%	22.88%	16.13%	12.05%	7.97%
Growth MFs	53.10%	22.11%	12.39%	7.71%	4.70%
Value MFs	22.11%	23.02%	21.70%	19.20%	13.97%
HFs	39.98%	21.23%	15.53%	12.89%	10.37%
ETFs	32.31%	23.23%	18.18%	15.74%	10.55%
	Panel B: 5 Largest Value Funds				
T Rowe Price Equity Income Fund	29.29%	23.56%	19.28%	14.59%	13.28%
Fidelity Equity-Income Fund	19.89%	22.66%	20.49%	22.36%	14.60%
T Rowe Price Value Fund	24.97%	24.43%	20.29%	14.34%	15.96%
Fidelity Value Fund	18.10%	25.93%	23.06%	19.61%	13.29%
DFA US Large Cap Value	0.84%	4.26%	25.42%	37.98%	31.50%
	Panel C: 5 Largest Growth Funds				
Fidelity Contrafund	45.30%	19.31%	16.35%	12.16%	6.88%
Growth Fund of America	50.71%	23.49%	12.41%	8.38%	5.01%
Fidelity Magellan Fund	42.39%	22.87%	15.29%	11.08%	8.37%
Fidelity Growth Company Fund	64.34%	18.09%	9.26%	5.52%	2.79%
Fidelity Blue Chip Growth Fund	63.28%	20.82%	8.17%	4.74%	2.98%

Note: This table shows the average portfolio shares in the five BM quintiles.

Table 3.9: Returns of Stocks and Mutual Funds

Quintile	ME	BM	MS	MULT	GR	MOM
			Panel A: Stocks			
1	4.06	2.38	3.25	3.64	3.71	2.89
2	3.54	3.64	3.96	3.98	4.20	3.56
3	3.63	4.00	4.04	3.52	4.30	3.94
4	3.64	4.25	4.35	3.34	4.04	4.22
5	3.17	5.20	4.32	4.61	3.27	4.55
$5-1$	-0.88	2.82	1.07	0.97	-0.43	1.66
			Panel B: Mutual Funds			
$[1,2]$	2.37	2.17	2.23	2.20	2.37	1.88
$(2,3]$	2.75	2.38	2.39	2.41	2.21	2.09
$(3,4]$	2.84	2.48	2.32	2.30	2.42	2.63
$(4,5]$	2.11	2.95	2.17	2.24	2.24	1.12
$(4,5]-(1,2]$	-0.25	0.78	-0.05	0.04	-0.13	-0.76

Note: The table reports the mean returns by quintile (stocks) and quintile ranges (mutual funds).

Table 3.10: Fama-MacBeth Regressions

ME	MOM	BM	MS
Panel A: Stocks			
-0.26	0.39	0.54	
[-1.65]	[2.44]	[5.01]	
-0.37	0.40		0.27
[-2.36]	[2.58]		[1.98]
Panel B: Mutual Funds			
-0.45	0.39	-0.02	
[-3.11]	[1.39]	[-0.14]	
-0.43	0.39		-0.05
[-3.20]	[1.53]		[-0.28]

Note: Fama-MacBeth regressions of returns of individual stocks and mutual funds on characteristic scores. The regression coefficients are in percent per month. t-statistics are in brackets.

Bibliography

Acemoglu, Daron, David H. Autor, and David Lyle, 2004, Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury, *Journal of Political Economy* 112, 497–551.

Acemoglu, Daron, Simon Johnson, and James A. Robinson, 2001, The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review* 91, 1369–1401.

Agarwal, Vikas, Vyacheslav Fos, , and Wei Jiang, 2013a, Inferring Reporting-Related Biases in Hedge Fund Databases from Hedge Fund Equity Holdings, *Management Sciences* 59, 12711289.

Agarwal, Vikas, Wei Jiang, Yuehua Tang, and Baozhong Yang, 2013b, Uncovering Hedge Fund Skill from the Portfolio Holdings They Hide, *Journal of Finance* 68, 739–783.

Andonov, Aleksandar, Rob M.M.J. Bauer, and K.J. Martijn Cremers, 2017, Pension Fund Asset Allocation and Liability Discount Rates, *Review of Financial Studies* 30, 2555–2595.

Andonov, Aleksandar, Yael V. Hochberg, and Joshua D. Rauh, 2018, Political Representation and Governance: Evidence from the Investment Decisions of Public Pension Funds, *Journal of Finance* 73, 2041–2086.

Ang, Andrew, Bingxu Chen, and Suresh Sundaresan, 2018, Liability Investment with Downside Risk, *Journal of Portfolio Management* 40, 71–87.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and Momentum Everywhere, *Journal of Finance* 68, 929–985.

Becker, Connie, Wayne Ferson, David H. Myers, and Michael J. Schill, 1999, Conditional market timing with benchmark investors, *Journal of Financial Economics* 52, 119 – 148.

Bellows, John, and Edward Miguel, 2009, War and local collective action in Sierra Leone, *Journal of Public Economics* 93, 1144–1157.

Bergstresser, Daniel, Mihir Desai, and Joshua Rauh, 2006, Earnings Manipulation, Pension Assumptions, and Managerial Investment Decisions, *The Quarterly Journal of Economics* 121, 157–195.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan B., and Jules H. van Binsbergen, 2015a, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

Berk, Jonathan B., and Jules H. van Binsbergen, 2015b, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1 – 20.

Bernstein, Shai, Josh Lerner, and Antoinette Schoar, 2013, The investment strategies of sovereign wealth funds, *Journal of Economic Perspectives* 27, 219–38.

Besley, Timothy, 1995, Property Rights and Investment Incentives: Theory and Evidence from Ghana, *Journal of Political Economy* 103, 903–937.

Besley, Timothy, and Hannes Mueller, 2012, Estimating the Peace Dividend: The Impact of Violence on House Prices in Northern Ireland, *American Economic Review* 102, 810–833.

Bradley, Daniel, Christos Pantzalis, and Xiaojing Yuan, 2016, The influence of political bias in state pension funds, *Journal of Financial Economics* 119, 69–91.

Brown, Jeffrey R., Joshua M. Pollet, and Scott J. Weisbenner, 2015, The In-State Equity Bias of State Pension Plans, NBER Working Papers 21020, National Bureau of Economic Research, Inc.

Cornwell, Christopher, and William N Trumbull, 1994, Estimating the Economic Model of Crime with Panel Data, *The Review of Economics and Statistics* 76, 360–366.

Cremers, K. J. Martijn, and Antti Petajisto, 2009, How Active Is Your Fund Manager? A New Measure That Predicts Performance, *Review of Financial Studies* 22, 3329–3365.

Cullen, Julie Berry, and Steven D. Levitt, 1999, Crime, Urban Flight, And The Consequences For Cities, *The Review of Economics and Statistics* 81, 159–169.

Daniel, et al, Kent, 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *Journal of Finance* 52, 1035–1058.

Dell, Melissa, 2015, Trafficking Networks and the Mexican Drug War, *American Economic Review* 105, 1738–1779.

Dix-Carneiro, Rafael, Rodrigo Soares, and Gabriel Ulyssea, 2018, Economic Shocks and Crime: Evidence from the Brazilian Trade Liberalization, *American Economic Journal: Applied Economics* 10, 158–195.

Dyck, A., and L. Pomorski, 2016, Investor Scale and Performance in Private Equity Investments, *Review of Finance* 20, 1081–1106.

Dyck, Alexander, and Adair Morse, 2011, Sovereign Wealth Fund Portfolios, Working Papers 2011-003, Becker Friedman Institute for Research In Economics.

Dyck, Alexander, Adair Morse, and Luigi Zingales, 2013, How Pervasive is Corporate Fraud?, Rotman School of Management Working Paper 2222608, National Bureau of Economic Research, Inc.

Engelmann, Dirk, and Martin Strobel, 2004, Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments, *American Economic Review* 94, 857–869.

Fama, Eugene F, and Kenneth R French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus Skill in the Cross-Section of Mutual Fund Returns, *Journal of Finance* 65, 1915–1947.

Freeman, Richard B., 1999, The economics of crime, in O. Ashenfelter, and D. Card, eds., *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, chapter 52, 3529–3571 (Elsevier).

Goldstein, Markus, and Christopher Udry, 2008, The Profits of Power: Land Rights and Agricultural Investment in Ghana, *Journal of Political Economy* 116, 981–1022.

Gould, Eric D., Bruce A. Weinberg, and David B. Mustard, 2002, Crime Rates And Local Labor Market Opportunities In The United States: 1979-1997, *The Review of Economics and Statistics* 84, 45–61.

Greenbaum, Robert T., and George E. Tita, 2004, The Impact of Violence Surges on Neighbourhood Business Activity, *Urban Studies* 41, 2495–2514.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2008, Trusting the Stock Market, *Journal of Finance* 63, 2557–2600.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, ... and the Cross-Section of Expected Returns, *Review of Financial Studies* 29, 5–68.

Hochberg, Yael V., and Joshua D. Rauh, 2013, Local Overweighting and Underperformance: Evidence from Limited Partner Private Equity Investments, *Review of Financial Studies* 26, 403–451.

Hornbeck, Richard, 2010, Barbed Wire: Property Rights and Agricultural Development, *The Quarterly Journal of Economics* 125, 767–810.

Jensen, Michael C., 1968, The Performance Of Mutual Funds In The Period 19451964, *Journal of Finance* 23, 389–416.

Johnson, Simon, John McMillan, and Christopher Woodruff, 2002, Property Rights and Finance, *American Economic Review* 92, 1335–1356.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-Varying Fund Manager Skill, *Journal of Finance* 69, 1455–1484.

Karpoff, Jonathan M., D. Scott Lee, and Gerald S. Martin, 2008, The Cost to Firms of Cooking the Books, *Journal of Financial and Quantitative Analysis* 43, 581–611.

Klapper, Leora, Raphael Amit, and Mauro F. Guilln, 2010, Entrepreneurship and Firm Formation across Countries, in *International Differences in Entrepreneurship*, NBER Chapters, 129–158 (National Bureau of Economic Research, Inc).

Levitt, Steven D, 1997, Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime, *American Economic Review* 87, 270–290.

Levitt, Steven D., 2004, Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not, *Journal of Economic Perspectives* 18, 163–190.

Levitt, Steven D., and Sudhir Alladi Venkatesh, 2000, An Economic Analysis of a Drug-Selling Gang's Finances, *The Quarterly Journal of Economics* 115, 755–789.

Lorentzen, Peter, John McMillan, and Romain Wacziarg, 2008, Death and development, *Journal of Economic Growth* 13, 81–124.

Mauro, Luciano, and Gaetano Carmeci, 2007, A Poverty Trap of Crime and Unemployment, *Review of Development Economics* 11, 450–462.

Mehlum, Halvor, Karl Moene, and Ragnar Torvik, 2005, Crime induced poverty traps, *Journal of Development Economics* 77, 325–340.

Novy-Marx, Robert, and Joshua D. Rauh, 2009, The Liabilities and Risks of State-Sponsored Pension Plans, *Journal of Economic Perspectives* 23, 191–210.

Novy-Marx, Robert, and Joshua D. Rauh, 2014, Funding Soft Liabilities, Working papers.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2017, Fund Tradeoffs, NBER Working Papers 23670, National Bureau of Economic Research, Inc.

Pstor, ubo, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Raphael, Steven, and Rudolf Winter-Ember, 2001, Identifying the Effect of Unemployment on Crime, *Journal of Law and Economics* 44, 259–283.

Romano, Roberta, 1993, Public Pension Fund Activism in Corporate Governance Reconsidered, *Columbia Law Review* 93, 795–853.

Rosenthal, Stuart S., and Amanda Ross, 2010, Violent crime, entrepreneurship, and cities, *Journal of Urban Economics* 67, 135–149.

Schmalz, Martin C., David A. Sraer, and David Thesmar, 2017, Housing Collateral and Entrepreneurship, *Journal of Finance* 72, 99–132.

Shleifer, Andrei, and Robert W. Vishny, 1993, Corruption, *The Quarterly Journal of Economics* 108, 599–617.

Soares, Rodrigo R., 2006, The welfare cost of violence across countries, *Journal of Health Economics* 25, 821–846.

Special Secretariat for Strategic Affairs of Brazil, 2018, Custos Econmicos da Criminalidade no Brasil.

van Binsbergen, Jules, and Michael W. Brandt, 2015, Optimal Asset Allocation in Asset Liability Management, in Pietro Veronesi, ed., *Handbook of Fixed Income* (Elsevier).

Voors, Maarten J., Eleonora E. M. Nillesen, Philip Verwimp, Erwin H. Bulte, Robert Lensink, and Daan P. Van Soest, 2012, Violent Conflict and Behavior: A Field Experiment in Burundi, *American Economic Review* 102, 941–964.

Wang, Shing-Yi, 2012, Credit Constraints, Job Mobility, and Entrepreneurship: Evidence from a Property Reform in China, *The Review of Economics and Statistics* 94, 532–551.

Wermers, Russ, 2000, Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses, *Journal of Finance* 55, 1655–1703.

Appendix A

Chapter 1 Appendix

A.1 Proofs

Proof of the Proposition 1

Define the functions $f(\cdot)$ and $g(\cdot)$ by

$$f(\bar{w}) \equiv 1 - \bar{w},$$
$$g(\bar{w}) \equiv \min\left\{1, \max\left\{0, \frac{(1-\pi)R_c - \bar{w}}{\pi}\right\}\right\}.$$

It is easy to see that $\bar{w}$ is a long-run equilibrium wage if, and only if, it satisfies $f(\bar{w}) = g(\bar{w})$. Let

$$\bar{w}_A \equiv (1-\pi)R_c - \pi,$$
$$\bar{w}_B \equiv (1-\pi)R_c.$$

Another way to write function $g(\cdot)$ is:

$$g(\bar{w}) = 1 \text{ if } \bar{w} \leq \bar{w}_A$$
$$g(\bar{w}) = \frac{(1-\pi)R_c - \bar{w}}{\pi} \text{ if } \bar{w}_A < \bar{w} < \bar{w}_B$$
$$g(\bar{w}) = 0 \text{ if } \bar{w} \geq \bar{w}_B$$

This implies that to check the number of intersections of the graphs of $f(\cdot)$ and $g(\cdot)$, it is sufficient to look at the wages $\bar{w}_A$ and $\bar{w}_B$. Simple math shows that:

$$f(\bar{w}_A) - g(\bar{w}_A) = (1-\pi)\left[\frac{\pi}{1-\pi} - R_c\right],$$
$$f(\bar{w}_B) - g(\bar{w}_B) = (1-\pi)\left[\frac{1}{1-\pi} - R_c\right].$$

In particular, $f(\bar{w}_B) - g(\bar{w}_B) \geq f(\bar{w}_A) - g(\bar{w}_A)$. Now, let us consider three possible cases.

Case 1: $R_c \leq \pi/(1-\pi)$. In this case, $f(\bar{w}_A) - g(\bar{w}_A) \geq 0 \Rightarrow f(\bar{w}_A) \geq g(\bar{w}_A)$ and $f(\bar{w}_B) - g(\bar{w}_B) \geq 0 \Rightarrow f(\bar{w}_B) \geq g(\bar{w}_B)$. Given that $f(\cdot)$ is affine and decreasing, this implies that $\bar{w}^* = 1 > \bar{w}_B$ is the single solution of the equation $f(\bar{w}) = g(\bar{w})$. For this wage, crime C^* should be 0.

Case 2: $R_c \geq 1/(1-\pi)$. In this case, $f(\bar{w}_B) - g(\bar{w}_B) \leq 0 \Rightarrow f(\bar{w}_B) \leq g(\bar{w}_B)$ and $f(\bar{w}_A) - g(\bar{w}_A) \leq 0 \Rightarrow f(\bar{w}_A) \leq g(\bar{w}_A)$. Given that $f(\cdot)$ is affine and decreasing, this implies that $\bar{w}^* = 0 < \bar{w}_A$ is the single solution of the equation $f(\bar{w}) = g(\bar{w})$. For this wage, crime C^* should be 1.

Case 3: $\pi/(1-\pi) < R_c < 1/(1-\pi)$. In this case, $f(\bar{w}_A) - g(\bar{w}_A) < 0 \Rightarrow f(\bar{w}_A) < g(\bar{w}_A)$ and $f(\bar{w}_B) - g(\bar{w}_B) > 0 \Rightarrow f(\bar{w}_B) > g(\bar{w}_B)$. Using a similar argument of cases 1 and 2, it is easy to see that $\bar{w}^* = 0 < \bar{w}_A$ and $\bar{w}^* = 1 > \bar{w}_A$ are equilibrium wages. According to the intermediate value theorem, there is also a third equilibrium wage $\bar{w}^* \in (\bar{w}_A, \bar{w}_B)$, for which $C^* \in (0,1)$. Given that $f(\cdot)$ and $g(\cdot)$ are linear in the interval $(\bar{w}_A, \bar{w}_B)$, this third equilibrium is unique.

Proof of the Proposition 2

Let

$$\bar{\lambda} = \infty, \, T = 2.$$

I start by showing that this law enforcement intervention (using the terminology of proposition 2) can shift the economy to the prosperous equilibrium. Note that if $t \in (t_0, t_0 + 1)$, then

$$\begin{aligned}\pi_t &= 1 - e^{-\int_t^{t+1} \lambda_s ds} \\ &= 1 - e^{-\bar{\lambda}} \\ &= 1\end{aligned}$$

Which implies that if $t \in (t_0, t_0 + 1)$, then

$$(1 - \pi_t)R_c - \pi_t j \leq 0, \, \forall j \in [0, 1] \tag{A.1}$$

Therefore, any agent born in $(t_0, t_0 + 1)$ would choose formal labor over a criminal career. In mathematical terms,

$$c^s = 0, \, \forall s \in (t_0, t_0 + 1)$$

This implies that at $t_0 + 1$, the percentage of criminals is given by

$$\begin{aligned}C_{t_0+1} &= \int_{t_0}^{t_0+1} c^s ds \\ &= \int_{t_0}^{t_0+1} 0 \; ds \\ &= 0,\end{aligned}$$

Therefore, the economy converged to the prosperous equilibrium.

Now, let Λ be the set of law enforcement intensities $\bar{\lambda} \in \mathbb{R} \cup \{\infty\}$ such that, for some T, the intervention $(\bar{\lambda}, T)$ induces a shift to the prosperous equilibrium. Given that $\infty \in \Lambda$, Λ is a non-empty set. Let $\bar{\lambda}_{min} \equiv \inf \Lambda$. For each $\bar{\lambda} > \bar{\lambda}_{min}$, we have $\bar{\lambda} \in \Lambda$, so $(\bar{\lambda}, T)$ induces a shift to the prosperous equilibrium for some T. Let $\tilde{T}$ be the infimum of such possible values for T. Then, if $T > \tilde{T}$, $(\bar{\lambda}, T)$ induces an equilibrium shift, as we wanted to prove.

Appendix B

Chapter 2 Appendix

B.1 Model Solution

In this appendix we prove that the optimal manager quality chosen by the board and the optimal contract offered to the portfolio manager are given by equations (2.8) and (2.11).

Optimal Contract

First, we assume that the manager with quality s is hired, and then we calculate the optimal contact offered by the board of trustees. We can clearly assume that $b = (1-a)\kappa$, given that financial and political returns are perfectly exchangeable in our model, which implies that the board would always offer the same fraction of political and of financial returns to the portfolio manager. To find new optimal value of the risk sharing parameter a, note that the objective function of the portfolio manager simplifies to:

$$r_f + (1-a)w^\top B(s) - 1/2\lambda(1-a)^2 w^\top \Sigma w \tag{B.1.1}$$

Where w is the vector of portfolio weights, Σ is the covariance matrix of returns, and $B(s)$ is the vector $B(s) = (s\varphi_{MV}, s\varphi_P + \kappa)^\top$. The optimal response that maximizes (B.1.1) is given by:

$$w = (1-a)^{-1}\lambda^{-1}\Sigma^{-1}B(s) \tag{B.1.2}$$

Now we can write the board's objective function as follows:

$$r_f + w^\top aB(s) - c - 1/2\Lambda_B a^2 w^\top \Sigma w \tag{B.1.3}$$

Let $v = a/(1-a)$. Basic algebra shows that (B.1.3) is proportional to

$$v - \frac{1}{2}\frac{\lambda_B}{\lambda}v^2 \tag{B.1.4}$$

Which is maximized by $v = \lambda_B/\lambda$. This implies that the optimal a is given by

$$a^* = \frac{\lambda}{\lambda + \lambda_B} \tag{B.1.5}$$

Optimal Manager Quality

By plugging the optimal contract into the board objective function, we find the following indirect utility function:

$$V_B(s) = r_f + 1/(2\bar{\lambda})B(s)^\top \Sigma^{-1} B(s) - O(s) \tag{B.1.6}$$

Where $\bar{\lambda} = (\lambda^{-1} + \lambda_B^{-1})^{-1}$. The underlying first order condition for the choice of the optimal quality is:

$$\Sigma^{-1}\varphi = O'(s) \tag{B.1.7}$$

Where $\varphi = (\varphi_{MV}, \varphi_P)^\top$. It's easy to see that this implies in the following condition:

$$\frac{(\sigma_P^2\varphi_{MV}^2 - 2\rho\sigma_P\sigma_{MV}\varphi_{MV}\varphi_P + \sigma_{MV}^2\varphi_P^2)s + (\sigma_{MV}^2\varphi_P - \rho\sigma_P\sigma_{MV}\varphi_{MV})\kappa}{\bar{\lambda}\sigma_P^2\sigma_{MV}^2(1-\rho^2)} = O'(s) \tag{B.1.8}$$

B.2 Comparative Statics Computations

In this appendix we compute the signals of the partial derivatives stated on the panels A and B on the comparative statics section of the paper. First we consider the case when the outrage constraint is not binding, and after that we compare the derivatives of the biding and not-biding cases.

Partial Derivatives of Manager Quality

If the outrage constraint is not binding, then the optimal manager quality s^* maximizes the ex-ante utility function of the board $V_B(s)$, which can be written as:

$$V_B(s) = \frac{1}{2\bar{\lambda}}B(s)^\top\Sigma^{-1}B(s) - O(s) \tag{B.2.1}$$

where Σ is the covariance matrix of returns, $O(s)$ is the outside option for a manager with quality s, and $B(s)$ is a vector defined by $B(s) = (s\varphi_{MV}, s\varphi_P + \kappa)^\top$. It's easy to see that we can write the underlying first order condition as

$$\bar{\lambda}^{-1}\varphi^\top\Sigma^{-1}[s\varphi + \kappa e_2] = O'(s) \tag{B.2.2}$$

where $\varphi = (\varphi_{MV}, \varphi_P)^\top$ and $e_2 = (0,1)^\top$. Differentiating (B.2.2) with respect to the political return κ we get:

$$[O''(s^*) - \bar{\lambda}^{-1}\varphi^\top\Sigma^{-1}\varphi]\frac{\partial s}{\partial\kappa} = \bar{\lambda}^{-1}\varphi^\top\Sigma^{-1}e_2 \tag{B.2.3}$$

The term $[O''(s^*) - \bar{\lambda}^{-1}\varphi^\top\Sigma^{-1}\varphi]$ is positive by the concavity of the objective function on the maximum, while the term $[\bar{\lambda}^{-1}\varphi^\top\Sigma^{-1}e_2]$ is negative if the Sharpe ratio of the mean-variance efficient securities is sufficiently larger than the Sharpe ratio of the political assets. This implies that:

$$\frac{\partial s}{\partial\kappa} < 0 \tag{B.2.4}$$

Now differentiating (B.2.2) with respect to the political return λ we get:

$$[O''(s^*) - \bar{\lambda}^{-1}\varphi^\top \Sigma^{-1}\varphi]\frac{\partial s}{\partial \lambda} = -\bar{\lambda}^{-1}O'(s) \tag{B.2.5}$$

The term $[O''(s^*) - \bar{\lambda}^{-1}\varphi^\top \Sigma^{-1}\varphi]$ is positive, while the term $[-\bar{\lambda}^{-1}O'(s)]$ is negative, which implies that:

$$\frac{\partial s}{\partial \lambda} < 0 \tag{B.2.6}$$

Partial Derivatives of Portfolio Weights

The vector of portfolio weights will be given by:

$$w = \bar{\lambda}^{-1}\Sigma^{-1}[s\varphi + \kappa e_2] \tag{B.2.7}$$

Differentiating (B.2.7) with respect to κ we get:

$$\frac{\partial w}{\partial \kappa} = \bar{\lambda}^{-1}\det(\Sigma)^{-1}\begin{bmatrix} \sigma_{MV}\sigma_P^2(\varphi_{MV}/\sigma_{MV} - \varphi_P/\sigma_P)\partial s/\partial\kappa - \rho\sigma_{MV}\sigma_P \\ \sigma_{MV}^2 - \sigma_{MV}^2\sigma_P(\varphi_{MV}/\sigma_{MV} - \varphi_P/\sigma_P)\partial s/\partial\kappa \end{bmatrix}$$

from which follows that:

$$\frac{\partial w_{MV}}{\partial \kappa} < 0,\ \frac{\partial w_P}{\partial \kappa} > 0 \tag{B.2.8}$$

Similar algebra shows that (i) the investment in fixed income is increasing on the risk aversion, and (ii) the investment on the mean-variance efficient security is decreasing on the risk aversion.

Comparison between Constrained and Unconstrained Cases

Now we compare the values of the partial derivatives with respect to the exogenous variables when boards are constrained and unconstrained. Its easy to see that:

$$\left[\frac{\partial w}{\partial \kappa}\right]_{unconstrained} = \bar{\lambda}^{-1}\Sigma^{-1}[\frac{\partial s}{\partial \kappa}\varphi + \kappa e_2] \tag{B.2.9}$$

$$\left[\frac{\partial w}{\partial \kappa}\right]_{constrained} = \bar{\lambda}^{-1}\Sigma^{-1}e_2 \tag{B.2.10}$$

And therefore:

$$\Delta\frac{\partial w}{\partial \kappa} = \left[\frac{\partial w}{\partial \kappa}\right]_{constrained} - \left[\frac{\partial w}{\partial \kappa}\right]_{unconstrained} = \bar{\lambda}^{-1}\Sigma^{-1}\frac{\partial s}{\partial \kappa}\varphi \tag{B.2.11}$$

And therefore:

$$\Delta\frac{\partial w_{MV}}{\partial \kappa} < 0,\ \Delta\frac{\partial w_P}{\partial \kappa} > 0$$

A similar argument shows that:

$$\Delta\frac{\partial w_{MV}}{\partial \lambda} > 0,\ \Delta\frac{\partial w_P}{\partial \lambda} > 0$$