ESSAYS IN ECONOMICS AND FINANCE

BY

MAHDI RASTAD

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Graduate College of the University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Daniel Mark Bernhardt, Chair
Professor Heitor Almeida
Professor Roger Koenker
Professor George Pennacchi
Abstract

Chapter 1: A large body of the corporate finance literature is devoted to capital structure. This literature examines whether firms have a target capital structure, and whether they actively rebalance their capital structure towards a target. Conversion of a convertible bond causes a drop in leverage, which target capital structure theory suggests should be rebalanced in the future. Consistently, evidence is provided that following a realized conversion, firms rebalance their positions in less than a year. When the stock price passes the conversion price threshold for a convertible bond, the firm expects this drop in leverage to occur in the near future. Using a regression discontinuity design around the conversion price threshold, for those conversions that are decided by investors, not by the firm it is documented that a 21% increase in leverage before an actual drop in leverage. That is to say, firms do not wait for the realization of leverage shocks, but rather respond to anticipated shocks. A quantile treatment effect analysis reveals the effect to be a hump-shaped function of leverage, with a peak for firms with a conditional leverage ratio around the 70th percentile.

Chapter 2: This chapter provides a theory of collusion under demand uncertainty by cartels of countries such as OPEC that do not care directly about profits, per se, but rather the utility derived by their risk averse citizens who receive those profits; and who face positive fixed operating costs. The chapter provides conditions under which it is most difficult for cartel members to collusively restrict output when demand is especially low, but it also becomes difficult to support collusion when demand is very high, showing both that cartel members must be risk averse and operating costs must be positive. Further it is established that when cartel members are more risk averse or fixed operating costs are higher, then it becomes more difficult to support collusion in bad demand states, but easier in booms.

Chapter 3: The last chapter presents a model of a public pension fund’s choice of portfolio risk. Optimal portfolio allocations are derived when pension fund management maximize the utility of wealth of a representative taxpayer or when pension fund management maximize their own utility of compensation. The
model’s implications are examined using annual data on the portfolio allocations and plan characteristics of 125 state pension funds over the 2000 to 2009 period. Consistent with agency behavior by public pension fund management, evidence is provided that funds chose greater overall asset–liability portfolio risk following periods of relatively poor investment performance. In addition, pension plans that select a relatively high rate with which to discount their liabilities tend to choose riskier portfolios. Moreover, consistent with a desire to gamble for higher benefits, pension plans take more risk when they have greater representation by plan participants on their Boards of Trustees.
To My Family
I am especially grateful to my main adviser, Dan Bernhardt, and my committee members Heitor Almeida, Roger Koenker and George Pennacchi for their invaluable guidance. I would also like to thank Murillo Campello for his guidance and delicate feedback. I thank Sadra Amiri Moghadam, Jeffrey Brown, Jaewon Choi, Prachi Deuskar, Vyacheslav Fos, Charles Kahn, Mathias Kronlund, Neil Pearson, seminar participants at University of Illinois as well as participants at National Bureau of Economic Research State and Local Pensions Conference for their helpful comments.

I would like to give my highest gratitude to my wife, Fatemeh, for being incredibly supportive and patient during the course of my Ph.D. program. I am also thankful to all my teachers, including my parents, without whom I could not have accomplished this journey.

I gratefully acknowledge financial support from Boltz fellowship. All remaining errors are mine.
Chapter 1  Capital Structure Pre-Balancing  ................................................. 1
  1.1 Introduction ................................................................................. 1
  1.2 Data and Sample Selection ......................................................... 6
    1.2.1 Convertible Bond Data ......................................................... 6
    1.2.2 Sample Construction ......................................................... 6
    1.2.3 Debt Conversion .............................................................. 10
  1.3 Debt Conversion and Financing Decisions : Theory and Practice ................................................................................. 13
    1.3.1 Capital Structure Rebalancing ........................................... 13
    1.3.2 Debt Conversion and Capital Structure ................................. 16
  1.4 The Response of Leverage to a Realized Conversion ...................... 19
    1.4.1 Graphical Analysis ............................................................. 19
    1.4.2 Diff-in-Diff Analysis at Conversion ....................................... 19
    1.4.3 Leverage Dynamics After Conversion ..................................... 22
  1.5 The Response of Leverage to an Anticipated Conversion ............... 24
    1.5.1 Empirical Strategy: A Regression Discontinuity Design .......... 25
    1.5.2 Unconditional Response ...................................................... 26
    1.5.3 Parametric and Non-parametric Analysis ............................... 28
    1.5.4 Heterogeneity in Leverage Response : Quantile Analysis .......... 32
  1.6 Concluding Remarks .................................................................. 34
  1.7 Appendix ................................................................................. 35

Chapter 2  Risk Averse Cartels ............................................................. 38
  2.1 Introduction ................................................................................. 38
  2.2 The Model ................................................................................. 40
  2.3 Conclusion ................................................................................. 50
Chapter 3 Portfolio Allocation for Public Pension Funds

3.1 Introduction

3.2 Related Literature on Public Pension Fund Portfolio Allocation

3.3 A Public Pension Fund Model

- 3.3.1 The Public Pension Fund’s Investment Objective
- 3.3.2 Public Pension Fund Portfolio Choice

3.4 Data and Variable Construction

3.5 Empirical Evidence

- 3.5.1 Tracking Error Minimizing Allocations and Alternative Portfolio Allocations
- 3.5.2 Tracking Error Volatility and Plan Characteristics

3.6 Conclusion

References
Chapter 1

Capital Structure Pre-Balancing

1.1 Introduction

In this paper, I study the capital structure decisions of firms that have issued convertible bonds, whose stock price is close to the conversion price thresholds at which bondholders can exchange the bond for equity. From the perspective of the firm, these decisions by bondholders act as exogenous shocks to their capital structure, altering the debt-equity composition, instantaneously reducing leverage by significant amounts. Target capital structure theory would suggest that firms should respond to such shocks by rebalancing. My paper investigates how firms respond in practice. In doing so, I answer a set of questions: How fully do firms respond to what are effectively exogenous shocks to their capital structure by re-establishing their original capital structure? What is the timing of firm responses? How quickly do firms rebalance their capital structures in response to exogenous changes to their capital structure? Do firms anticipate likely future changes to their capital structure, by preemptively beginning to rebalance prior to the actual shocks to their capital structure? In other words, does a firm’s management passively wait until a change actually occurs, or does it respond preemptively when the current stock price suggests that future changes are likely?

Previous research has focused on capital structure decisions in response to realized deviations from target levels. The literature has largely ignored the real-time capital structure decisions that firms must make as a function of anticipated likely changes in capital structure. This is partially due to the fact that it is typically hard to quantify anticipated changes in capital structure. More importantly, given that capital structure compositions is a choice variable and therefore an internal decision to the firm, finding a clear answer to these questions requires an exogenous source of variation in capital structure, a feature that has been completely ignored by prior studies. That is, one needs to distinguish the effect of capital structure from the effects of other firm characteristics that influenced the actual capital structure choice. The goal of this paper is to
exploit the discrete changes in capital structure caused by conversion features of convertible bonds to answer these questions.

Static trade-off theory hypothesizes that firms identify an optimal target for their composition of debt and equity to balance financial distress and agency costs (of the conflict between lenders and shareholders) of debt to its tax benefits. Empirical evidence of the relevance of target capital structure theory is mixed. Titman and Wessels (1988), Rajan and Zingales (1995), Graham (1996) and Hovakimian et al. (2001) provide evidence of an association between leverage and firm characteristics consistent with a target capital structure. Moreover, Jalilvand and Harris (1984), Auerbach (1985), Flannery and Rangan (2006) and Faulkender et al. (2011) provide evidence consistent with firms actively rebalancing their capital structure toward a target. Conversely, the negative association between past profitability and the leverage ratio is widely presented as evidence against trade-off theory (Hovakimian et al. (2001)). In addition, inconsistent with the target capital structure hypothesis, studies have documented that changes in capital structure are mainly the result of internal financing deficits (Shyam-Sunder and C. Myers (1999)), historical stock returns (Welch (2004)), or management attempts to time the market (Baker and Wurgler (2002)). A possible reason for why these studies have had mixed results is lack of identification strategy which makes it really hard to measure the causal impact of leverage deviations from the target on future financing decisions of the firm.

The conversion option of convertible bonds provides a unique opportunity to study the empirical relevance of target capital structure hypothesis by examining how the financing decisions of firms respond to both realized and anticipated changes in capital structure. Bondholders find it optimal to convert their bonds into equity only after the stock price passes a predetermined threshold (the conversion price). Conversion causes an instant drop in the leverage ratio, a drop that is exogenous from the perspective of the firm. Likewise, the “expected leverage” ratio falls as the share price rises, approaching the conversion threshold. I search for evidence that the conversion threshold serves as a trigger for firms, inducing them to change capital structure when the stock price hits the threshold. I design my tests so that conversions and anticipated conversions, detected by the conversion option becoming in the money, can be seen as treatment. Following Lee (2008) I argue that around the conversion threshold, the assignment of treatment is close to a random assignment. This provides a quasi-experimental design to study the effect of changes in anticipated leverage ratio on financing decisions.

Two facts make this opportunity a credible framework for investigating target capital structure theories. First, as Stein (1992) points out convertible bonds are an important source of financing for many firms.
Essig (1991) provides evidence that more than 10% of all COMPUSTAT companies had at least one-third of their total debt in the form of convertible debt in the period 1963-1984. Second, voluntary conversions (i.e., conversion that are decided by the bondholder and not by the firm) cause significant changes in capital structure. As I document below the median size of a voluntary conversion shock is approximately 7.3% relative to total debt. Therefore, capital structure consequences of a conversion are both frequent and sizeable.

While theory provides a variety of explanations for why firms issue convertible bonds, it seems less clear that theory predicts how capital structure responds to a conversion. The presence of the conversion option is rationalized by its ability to mitigate the underinvestment problem caused by “risk-shifting” (Green (1984)), avoiding adverse-selection costs associated with direct equity issuance when shares are undervalued (Stein (1992)), and controlling over-investment or free cash flow problems (Jensen (1986)) by matching firm’s financing and investment options (Mayers (1998)). Contrary to others, Mayers (1998) provides insights into firm’s investment and financing behavior at conversion. He predicts that a drop in the leverage ratio due to a conversion would free up debt capacity. If firms use this new capacity to raise new debt to finance investment options, consistent with rebalancing theory, this would cause leverage ratio to bounce back quickly. However, his empirical evidence—that both debt and equity issuances increase after conversion—makes it ambiguous as to what the direction is. This calls for a clean empirical analysis to identify the direction of financing decision in response to capital structure changes.

To study the link between changes in capital structure and financing policy, one needs to address the endogeneity of capital structure. Without that, one may wonder whether the effect comes from changes in capital structure per se or from common underlying factors that determine the choice of capital structure. To alleviate this concern, I exploit the discrete changes in capital structure around the conversion threshold. In particular, the discontinuity in leverage due to conversion feature of convertible bonds enables me to employ a regression discontinuity design as an empirical strategy to identify the impact of anticipated leverage changes on capital structure decisions. The quasi-experimental set-up in regression discontinuity design provides a clean experiment in which changes in capital structure around the conversion threshold can be considered as exogenous. This approach allows me to compare otherwise similar firms, for which the stock [roce is just below (above) the conversion price threshold and therefore conversion option is out of (in) the money. This comparison sheds light on how an anticipated change in capital structure (due to the conversion option being in the money) affects financing policy of firms.
I also examine firms’ financing responses to a realized change in capital structure, by comparing their behavior before and after a realized conversion to the behavior of a group of similar firms that did not experience a conversion in that period. More specifically, I use a difference-in-difference approach for the changes in the leverage ratios before and after conversions. To do this, I use a group of firm-quarter observations with a voluntary conversion as the treatment group. Perhaps the closest control group to the treatment group is the group of firm-quarter observations with outstanding convertible bonds that eventually have a conversion. I restrict the treatment to only voluntary conversions so that the resulting drop in leverage is as exogenous to the firm as possible. To explore leverage dynamics following a conversion, I continue comparing the leverage ratio between the two groups during the periods following a conversion up to a point when the effect disappears.

My first finding shows that a realized conversion on average reduces leverage of about 4.5%, which is more than 12% with respect to average capital structure. This finding is robust to the inclusion of conventional leverage control variables as well as year and quarter fixed effects in the model specification. Further, nearly identical results are obtained by either expanding the control group to the universe of all convertible bond issuers, or by restricting the experiment to a subset of more exogenous conversions (i.e., that occur during the hard call protection period). Although conversion creates a mechanical drop in the leverage ratio, documenting the economic significance of this relationship is a pre-requisite for the subsequent analysis.

I next show that following a realized conversion, firm rebalance their position in less than a year. This evidence is consistent with the trade-off theory’s prediction that firms actively rebalance their capital structure in response to deviations from the target level. What distinguishes my finding from prior studies is that in line with Graham and Leary (2011)’s suggestion. I locate a source of variation in capital structure outside the partial adjustment model (Jalilvand and Harris (1984), Auerbach (1985), Flannery and Rangan (2006), Faulkender et al. (2011)). In addition, compared to other papers using variations outside the partial adjustment model (e.g., Harford et al. (2009), Baker and Wurgler (2002), Kayhan and Titman (2007), Alti (2006) and Hovakimian et al. (2001)) I use an arguably exogenous source of change in capital structure, i.e., voluntary conversion of convertible bonds by investors and not by the firm. This enables me to pin down the causal effect of capital structure changes on financing policy.

I also find strong evidence that firms actively “pre-balance”. Using a regression discontinuity design to compare firms with convertible bonds for which the conversion option is in the money vs. out of the money, I find that firms increase their leverage even before a realization of the drop in leverage. More specifically,
I find that the book value of long term debt rises as soon as the stock price crosses conversion threshold by approximately 7.4% of assets –a 21% increase with respect to the average leverage level before passing the threshold. To use the entire sample, I control for the potential information contained in stock price distance to conversion threshold by a polynomial in this distance, labelled as conversion premium. This result is robust to the inclusion of control variables including conventional leverage controls, as well as firm, year and quarter fixed effects. Both parametric and non-parametric analysis indicate that pre-balancing begins a quarter before the stock price passes the conversion threshold and continues until a quarter afterwards. These findings suggest highly active capital structure management in response to deviations from target capital structures.

Finally, I show that the degree of pre-balancing varies as a function of the leverage level. A quantile treatment effect (QTE) analysis indicates that pre-balancing is a hump-shaped function of leverage ratio. More specifically, firms with extremely high or low conditional leverage levels pre-balance by increasing leverage levels by 3% to 5%; but this number surges up to 13% for firms that are moderately levered.

My primary contribution to the capital structure literature is to provide evidence in support of capital structure rebalancing in response to an exogenous variation in capital structure. My paper is the first to show that re-balancing may begin even before the realization of a change in capital structure. Using conversion feature of convertible bonds, I propose a novel method to track anticipations of future changes in capital structure. My paper also relates to literature explaining the rationale behind issuing convertible bonds. In particular, my finding is consistent with Mayers (1998)’s theory of staged financing, predicting an increase in debt financing and investment activities around conversion time.

My findings for variation of pre-balancing in low and middle quantiles of the leverage ratio may also be interpreted as a counter example to debt overhang (Myers (1977), Lang et al. (1996), Hennessy (2004)). Debt overhang hypothesis would predict the pre-balancing to be a decreasing function in leverage. In contrast, QTE analysis shows that pre-balancing is increasing in leverage for a wide range of conditional leverage levels, i.e. from the 5th to about the 70th. Only in high leverage quantiles (above the 70th quantile), do I find a decreasing pre-balancing function in leverage. Thus, debt overhang hypothesis seems to be relevant only for highly-levered firms.

The remainder of this paper is organized as follows. Section 1.2 discusses data and sample construction. Section 1.3 presents the theoretical motivation for the study by outlining how a conversion affect capital
structure and what theory predicts about the response of financing decision to such changes. Section 1.4 presents the results of the analysis investigating capital structure response to a realized conversion. In section 1.5 I provide evidence for pre-balancing using regression discontinuity design analysis. Section 1.6 concludes the paper and proposes a number of questions for future research.

1.2 Data and Sample Selection

1.2.1 Convertible Bond Data

The convertible bond information used in this study comes from the Fixed Income Securities Database (FISD), in which the basic unit of observation is a bond issue. Given that firms typically have multiple issues outstanding at the same time, I exploit the dataset to aggregate information on all bonds that are issued by the same issuers. FISD provides general information on bond characteristics such as issue date, issue amount, yield and maturity as well as information that is more specific to convertible bonds. This includes the beginning and ending dates of the conversion option, conversion price, conversion premium (a distance measure between conversion price and current market price of the underlying commodity that the bond is converted to), the level of protection against calls by the firm management, and the relative size of each issue with respect to the underlying commodity. Moreover, FISD records the history of events (called “actions” in FISD terminology) per each issue. Examples of such events are conversions, calls, exchanges, new offerings, etc. Table 1.1 provides a complete list of those events along with the frequency of such events in convertible bonds sample. For any typical event, FISD reports the following information: the effective date of the event, the event amount (for example the amount that is converted) and the remaining outstanding balance of the issue. Although bond characteristics have been available in FISD since the 19th century, the fact that event history data is fairly limited before 1995, forces us to focus our attention on the sample of convertible bonds with event dates between 1995 and 2010.

1.2.2 Sample Construction

I begin with all 4,269 convertible bond issues in FISD issued between 1967 and 2010, where 1967 is the earliest year for which convertible bond data is available in FISD. First, I drop issues lacking conversion price or conversion date information and exclude financial issuers (SIC codes 6000-6999). This results in
a sample size of 3,347 issues. Second, I match frequency of observations across datasets. Events for each issue in FISD are available at daily frequencies. However, given that the highest frequency accounting data available from COMPUSTAT is quarterly, I aggregate the issue-event data, into issue-quarter levels so that data frequencies are harmonized between the two sources. This procedure aggregates 13,096 total daily events into 12,940 quarterly events (or 11,204 events after exclusion of events for financial firms). This indicates there are 156 multiple events in quarters in my sample. For example for the case of conversion events, there are 56 multiple conversions for an issue with more than one conversion in the same quarter. Aggregating multiple conversions for a single issue to quarterly basis makes a sample of 436 quarter-conversions. These issue conversions correspond to 318 issuer firms between 1986 and 2010.

Among the events affecting capital structure listed in Table 1.1, I focus on conversions. A (voluntary) conversion is a type of event that is not initiated or planned by the firm and therefore can be considered as a treatment chosen outside a management’s jurisdiction. Table 1.1 shows 428 cases recorded in FISD of conversions not initiated by the firm. Conversions that are the result of decisions by the firms, e.g., calls, are more numerous (total of 1088 calls), but they are less appealing for identification purposes. More importantly the table shows that conversion shocks are sizeable with a median of about $40 million. The right panel of this table shows that relative to the size of equity the value of converted securities in a conversion shock is on average about 7.4%, the second largest shock in this table. Lastly, notice that almost all convertible bond issues are either converted or called before they mature. The last row of Table 1.1 shows that among the total of 2,097 events recorded in this table, only 9 of them are related to the final maturity of an issue.

<table>
<thead>
<tr>
<th>Event</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>(SD)</th>
<th>Amount (M$)</th>
<th>Amount as Fraction of Equity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issue Converted</td>
<td>428</td>
<td>91.35</td>
<td>39.64</td>
<td>(140.88)</td>
<td>296</td>
<td>7.39</td>
</tr>
<tr>
<td>Balance of Issue Called</td>
<td>331</td>
<td>85.62</td>
<td>34.84</td>
<td>(125.98)</td>
<td>153</td>
<td>4.58</td>
</tr>
<tr>
<td>Entire Issue Called</td>
<td>664</td>
<td>200.44</td>
<td>105.96</td>
<td>(242.46)</td>
<td>235</td>
<td>8.29</td>
</tr>
<tr>
<td>Part of an Issue Called</td>
<td>95</td>
<td>87.55</td>
<td>13.96</td>
<td>(182.42)</td>
<td>50</td>
<td>1.38</td>
</tr>
<tr>
<td>Issue Repurchased</td>
<td>572</td>
<td>110.72</td>
<td>36.67</td>
<td>(196.22)</td>
<td>471</td>
<td>2.76</td>
</tr>
<tr>
<td>Issue Matured</td>
<td>9</td>
<td>102.40</td>
<td>99.57</td>
<td>(113.74)</td>
<td>8</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Table 1.1: Summary of shocks to capital structure: The table presents a summarized list of events affecting capital structure for a bond issue in FISD between 1995 and 2010.

To ensure that conversion is not driven by a merger or acquisition, I match the resulting FISD sample to the Securities Data Corporation (SDC) dataset and exclude 17 issuers that were the target of a merger or acquisition deal between 1980 and 2010. Similarly, to make sure that conversions are not driven by a call due
to an IPO, I drop 7 observations with an IPO-clawback \(^1\) provision. To measure the pure effect of conversion—and no other treatment, I also require that no other event listed in table 1.11 happens in the same quarter in which conversion occurs\(^2\). This includes events which affect capital structure such as conversion-forcing calls and repurchase. The resulting sample is 307 conversion events for 217 issuer firms between 1985 and 2008 (a total of 3,657 issues for all events). Next, I move from issue-quarter units to firm-quarter units by aggregating observations across multiple issues outstanding for a single firm. Similar to Mayers (1998), I calculate the sum of event-amounts for each event-quarter across different issues outstanding for each firm. For example, if a firm experiences \(X\) dollar conversions on a 5-year bond issue and \(Y\) dollar conversion on a 10-year bond issue during the same quarter, I consider \(X + Y\) as the total conversion in that quarter. This results in 275 firm-quarter observation corresponding to 183 firms, a total of 1,362 firms for all events.

Finally, I match the resulting dataset to quarterly accounting data from COMPUSTAT. The reason for using quarterly data, the highest frequency in COMPUSTAT, is to get as close as possible to the event dates in FISD, so that I can generate the most accurate estimate for the impact of a conversion. This reduces the size of the conversion sample to 204 firm-quarter observations for 129 firms. I then draw treatment and control samples for each of the first and second sections of the paper, the realized and anticipated conversion analysis, as follows. For the former the treatment group contains firm-quarter observations in which only a conversion event and no other event is recorded. To show robustness of results for the first experiment, I use three different control groups: (a) all convertible bond issuers in quarters that did not experience a conversion, (b) firm-quarter observations for those convertible issuers who eventually experience a conversion but not in the current quarter, and (c) a subset of group (b), conversions happening during the hard call protection period. Conditioning my second and third control groups on having a conversion eventually in the future, controls for potential unobservable factors affecting taking the treatment, i.e., voluntary conversion by the bond holders. To make sure I measure the pure effect of treatment, I put a four-quarter and one-quarter cushions of no event around both treatment and control quarters. The final treatment sample contains 145 firm-quarter observations for 103 firms. For the second experiment, the treatment group contains firm-quarter observation in which (1) there is a convertible bond outstanding, (2) which eventually will be converted in the future (but not yet), and (3) for which the stock price passed the conversion price.

\(^1\)The IPO-clawback option exists when the issuer has the right to call the issue with the proceeds of an initial or subsequent stock offering

\(^2\)Except for the “Review” event which is used when FISD staff review an issue that did not experience any event in that time period
The control group here is identical to the treatment group, except for condition (3) which is reversed: stock price did not reach the conversion price threshold. This results in a treatment (control) group containing 203 (283) firm-quarter observations for 48 (40) firms.

**Selection**

Given that sample construction is not a random procedure, potential selection concerns may limit my ability to generalize or infer beyond my sample. However, Tables 1.2 and 1.3 address this concern by comparing the universe of non-financial firms in COMPUSTAT and FISD during the same years to my sample. To limit the effect of outlier, I trim COMPUSTAT ratios at upper and lower 5% and all and other samples at 1%. Table 1.2 shows that while convertible and conversion samples are fairly similar to COMPUSTAT in terms of profitability, tangibility and investment to capital ratio, they differ along other dimensions. In particular, consistent with Essig (1991), the conversion sample contains firms with higher leverage levels, higher market to book ratio and larger cash flow volatility. Stein (1992) interprets these findings as evidence of a higher cost of financial distress for convertible bond issuers. Moreover, Table 1.2 also shows that firms in the conversion sample are relatively larger, have higher cash flow ratio and are more likely to pay dividends than the COMPUSTAT sample. Comparing conversion to convertible sample in the same table, shows that those convertible issuers who experience a conversion are more likely to be larger, have better growth opportunities as measured by market to book ratio (similar to Lewis et al. (1999)), have higher cash flow volatility and have more tangible assets.

Similarly, Table 1.3 reflects the differences of characteristics of bond issues in the convertible sample, the conversion sample and all bonds in the FISD universe. We Immediately see sharp differences in the convertible bond sample relative to other bonds: convertible bonds on average have longer maturity, (over 13 years compared to about 8 years for the FISD sample), as a premium for the conversion option, have a lower yield rate and are issued in larger amounts. However, the convertible sample and its sub-sample of bonds that experienced a voluntary conversion (conversion sample) are quite similar in many dimensions. Perhaps the main difference between the two samples are: (i) in the convertible sample the majority of bonds become immediately convertible after issuance, whereas in conversion sample there is at least a 9-month delay for the majority of the sample; and (ii) the convertible period is shorter in the conversion sample. Thus with the exception of these two dimensions, convertible and conversion samples which constitute the treatment
Table 1.2: Firm sample selection comparison: This table presents summary statistics—averages, medians, and (standard deviations)—for three samples of firm-quarter observations. COMPUSTAT Sample consists of all firm-quarter observations from non-financial firms in COMPUSTAT between 1995 and 2010. Convertibles Sample consists of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database that have at least one convertible bond outstanding from 1995 to 2010. Conversions Sample consists of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database for which at least for one of their convertible bond issues a conversion event is recorded in FISD between 1995 and 2010. In both Convertible and Conversion samples, when there is more than one bond issue outstanding per firm data is aggregated across issues.

and control groups in the first part of the paper are identical in other dimensions.

1.2.3 Debt Conversion

Conversion is a unique feature of convertible debt instruments, which gives the right to either the debt holder or borrower to exchange debt for equity under certain conditions. These conditions are typically related to the price of the underlying commodity, e.g., a stock must be traded above a “conversion price” threshold for 30 trading days. If a borrower enforces a conversion this event is recorded as a “call” or a “conversion-forcing call”. On the other hand, if a debt holder exercises the conversion option, it is recorded as a (voluntary) “conversion”.

I focus attention on conversions that are decided by debt holders (voluntary conversions) and not conversions by the borrower or firm. Conversion-forcing calls, which are clearly planned by the firm, do not
Table 1.3: Bond sample selection comparison: This table presents summary statistics - averages, [medians], and (standard deviations) - for the three samples of firm and issue observations. FISD sample consists of all bonds in FISD dataset issued by non-financial firms outstanding at least for year between 1995 and 2010. Convertibles sample consists of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database that have at least one convertible bond outstanding from 1995 to 2010. Conversions sample consists of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database for which at least for one of their convertible bond issues a conversion event is recorded in FISD between 1995 and 2010. In both Convertible and Conversion samples, when there is more than one bond issue is outstanding per firm data is aggregated across issues. Offering Yield is the yield to maturity at the time of issuance for fixed rate issues, Offering Amount is the par value of debt initially issued and Fraction of Equity is the percentage of the total conversion commodity available through conversion of the issue. All time periods are in years.

represent anticipated shocks to capital structure; rather, they are almost deterministic from the firm’s management point of view.

In Table 1.4, I examine several features of voluntary conversions. Beginning with their frequency of occurrence, 11% of convertible bond issuers in the sample experienced at least one conversion between 1995 and 2010, which constitutes about 16% of firm-quarter observations. Notice that some firms experienced multiple conversions in this period. As a fraction of issues, 7% of convertible bond issues in FISD are classified with at least one voluntary conversion event. This table also shows that the average amount of conversion with respect to assets is 7% and therefore is an important event. First conversions in most cases occur early: after 2.25 years for a typical convertible bond with 10 years of life.

---

3As pointed out by Asquith and Mullins (1991) when a firm calls an issue, investors are typically given a 30 day notice period to convert their bonds into equity. Therefore, most calls are conversion-forcing calls unless the stock price falls drastically enough during the 30-day notice that the conversion value becomes less than the call price. In that case the firm has to redeem the issue in cash instead of forcing its conversion into equity.
Table 1.4: Summary of conversions: This table presents summary statistics—averages, [medians], and (standard errors)—for a sample of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database for which at least for one of their convertible bond issues a conversion event is recorded in FISD between 1995 and 2010. When there is more than one bond issue outstanding per firm data is aggregated across issues.

Conversion Premium

The conversion premium, π, is the distance between the conversion price, \( \bar{p} \), and the current market price, \( p \), for the underlying commodity. To make it more tractable, I flip the normal definition for conversion premium by multiplying it by \(-1\) as in equation (1.1). In this definition, the larger the commodity price is, the larger the conversion premium. When \( \pi > 0 \) (\( \pi < 0 \)) the conversion option is in the money (out of the money).

\[
\pi = \frac{p - \bar{p}}{p} \tag{1.1}
\]

While mathematically straightforward, empirically measuring the conversion premium is challenging. First, notice that FISD only reports the conversion premium at an event point such as a conversion or a call. For all other points of time such as before or after the stock price passes the conversion threshold I have to compute this number as follows. I keep track of slight changes in conversion price over time using FISD’s Convertible History table. To measure how close the market price gets to the conversion price threshold, I take the average of the highest monthly stock price for the three months in each quarter. I choose to do this instead of using the monthly average stock price, since what triggers a conversion in reality is typically a 30-day trade of the stock at a price above the conversion price. Therefore what matters is the upper tail of the stock price, and not the mean or median. Obviously, this is an imperfect measure of the conversion premium as it only uses three points of data in every quarter. For example, if the stock trades only for a few days above the conversion price, the measure will underestimate the conversion premium, i.e., the distance
to conversion price.

The last two rows of Table 1.4 summarize the conversion premium at issuance and at the time of conversion. Notice that the thick left tail of conversion premium distribution results in a negative mean, even at the time of conversion. However, the sample median at conversion shows that for the majority of issues the stock price is about 17% above the conversion threshold. This is similar to the delay in calls reported by Ingersoll (1977), Asquith (1995) and Asquith and Mullins (1991) among others. The median conversion premium at the time of offering shows that for most of the bonds, the conversion price is set 24% above the market price.

1.3 Debt Conversion and Financing Decisions: Theory and Practice

Why would a debt conversion affect the financing decisions of a firm? This section answers this question by breaking it into two parts? (i) How does financing policy respond to changes in capital structure. And, (ii) how does a debt conversion change the capital structure?

1.3.1 Capital Structure Rebalancing

Does a change in capital structure affect future financing choices between debt and equity? In contrast to Modigliani and Miller (1958a) view that advocates for the irrelevance of capital structure in absence of market frictions, static the trade-off model provides evidence supporting the existence of this link. Trade-off theory predicts that firms form an optimal leverage target and actively rebalance their positions in response to any shock that causes a deviation from the target. Similar to the empirical literature on its competing theory, pecking order (Myers and Majluf (1984)), which suggests a hierarchy for a firm's preferences of financing sources, empirical literature on target capital structure provides mixed conclusions.

The literature on capital structure rebalancing is divided into two broad categories: (i) partial adjustment models (PAM) and its critiques, and (ii) papers outside partial adjustment model.
Partial Adjustment Model (PAM)

The partial adjustment model links changes in leverage to its lagged deviation from a target. Several empirical papers examine the sensitivity of this relationship, known as “speed of adjustment” (SOA). For example, consistent with rebalancing toward a target capital structure, Jalilvand and Harris (1984) and Auerbach (1985) find statistically significant parameters for mean reversion in leverage. Further, Jalilvand and Harris (1984) find that SOA is related to the firm size, expectation about interest rate and stock price level. Nevertheless, PAM has been criticized from two different perspectives.

First, more recent studies such as Fama and French (2002), Baker and Wurgler (2002), Welch (2004) and Iliev and Welch (2010) suggest that the rate of reversion to target is too slow to have a first-order effect on capital structure. In particular, Iliev and Welch (2010) argue that once accounting for dynamic misspecification in previous studies, partly due to boundedness of the leverage ratio between 0 and 1, the SOA becomes zero or negative, i.e., no rebalancing. Second, Shyam-Sunder and C. Myers (1999) as well as Chang et al. (2006) question the power of PAM in distinguishing between trade-off and pecking order theories. Using a Monte Carlo simulation they show that boundedness of leverage ratio can generate an artificial mean reversion when there is no mean-reversion in generated data. In response to the first critique, Flannery and Rangan (2006) provide evidence that adjustment speed increases significantly after controlling for expected equity price changes in the target. Along the same line, Faulkender et al. (2011) argue that cash flow deficit or surplus increases a firm’s willingness to interact with capital markets and therefore provides an opportunity for leverage rebalancing. They show that SOA becomes significantly faster, if cash flow realization is taken into account. Going back to the big picture, as Graham and Leary (2011) point out, perhaps:

“more work refining estimates of partial adjustment parameters may not be the most fruitful path to answering this question. ... Rather, new methods outside the partial adjustment framework may be necessary to identify the circumstances under which firms make deliberate, value-relevant financing decisions and when they fail to do so.”

My paper is precisely an example for such an exercise recommended by Graham and Leary (2011): looking at leverage response to a shock outside the PAM framework. I use changes in capital structure caused by conversion of a convertible debt into equity and examine how capital structure reacts to a realized or an anticipated conversion. The next section is an overview of other examples addressing target capital structure

---

4Iliev and Welch (2010) and Leary and Roberts (2005) also make the same point using Monte Carlo simulation.
response to a factor outside PAM.

**Outside Partial Adjustment Model: Opportunities for Rebalancing**

Other than studies that confirm association between leverage ratio and firm characteristics (e.g., firm size, investment opportunities and marginal tax rate) that are consistent with trade-off theory (for example Titman and Wessels (1988), Rajan and Zingales (1995), Graham (1996) and Hovakimian et al. (2001)) a few studies examine capital structure rebalancing using leverage deviations caused by a factor outside PAM. My paper adds to this literature by introducing another deviation in leverage, namely debt conversion.

Although leverage deviations in papers outside PAM originate from different factors inside or outside the firm, they all seem to be consistent with the trade-off theory. Using deviations in capital structure caused by an accumulation of earnings or losses over time Hovakimian et al. (2001) show that when firms raise or retire capital (including convertible debt) they use those opportunities to move toward their target leverage levels rather than to offset the deviations caused by earning and loss accumulations. They emphasize the role of capital repurchase/retirement as opposed to issuance as an opportunity for moving towards the target leverage. Using a discrete choice model, they find that leverage deviation predicts the sign and the size for debt and equity repurchase, but not for their issuance. Contrary to active rebalancing theory, Baker and Wurgler (2002) provide evidence for passive “market timing” by managements. By factoring market to book ratio with the amount of capital raised (financial deficit) they conclude that current leverage ratios are the result of historical stock price movements. In other words, high leverage ratios tend to mean revert, by issuing equity but only in favorable market conditions (market timing).

Related to market timing, Alti (2006) uses deviation in capital structure due to “hotness” of the IPO market at the time of initial public offering. He finds that firms that time the market by going public when the IPO market is hot, i.e., high volume of IPO activity, tend to raise more equity and therefore have lower leverage ratios than those who enter a cold market. However, subsequent reverse pattern in new capital issuance causes leverage differences to fade within two years. Kayhan and Titman (2007) investigate long-term capital structure adjustments in response to deviations from the target caused by historical stock price movements and financial deficit. Consistent with Baker and Wurgler (2002) they confirm the role of historical changes in cash flow, capital expenditure and stock price on leverage ratios which tend to persist over years. However, their findings also show that long-term reverse adjustments will eventually offset those
deviations in such a way that they move capital structure back toward its target level.

Finally, Harford et al. (2009) use deviations in leverage due to the type of financing instrument chosen in large merger and acquisition deals. Consistent with target capital structure theory, they find that deviation from a firm’s target leverage before an acquisition influences both the choice of financing method in an acquisition and how the bidder rebalances its position afterwards. Specifically, they show that immediately following a debt-financed acquisition, the bidders rebalance their position in such a way that compensates for 75% of the merger induced increase in leverage in 5 years.

A distinguishing feature of my paper as compared to the above examples is the exogeneity of deviations in capital structure. A second look at the examples above indicates that the leverage deviations from the target is usually caused by a variable of choice, and therefore is an endogenous firm decisions rather than a shock. Deviations caused by the choice of financing method in Harford et al. (2009), (delays in) equity issuance in Baker and Wurgler (2002), historical cash flow or capital expenditure choices in Kayhan and Titman (2007), the choice of when to enter the IPO market in Alti (2006) and the accumulation of earnings or losses in Hovakimian et al. (2001) are all endogenously determined by management decision. In contrast, in my setting voluntary conversion of a convertible debt (or anticipation of such conversion) which is triggered by the stock price passing the conversion threshold is decided by the debt holder, and not the firm. Therefore, under the assumption of “no price manipulation” by firm management around the conversion price threshold, resulting deviations in capital structure can be considered as exogenous to management decisions. This is certainly a step forward toward the identification of capital structure response to its deviation and ultimately to testing the static trade-off theory.

1.3.2 Debt Conversion and Capital Structure

Conversion of a convertible bond into equity creates a drop in leverage that might be endogenous or exogenous to the firm. When converted, the long-term debt item on the balance sheet would decrease by the face value of the converted amount and the equity term would increase roughly by the same amount (as a function of the amount converted, times the conversion ratio). Clearly, as a result, the leverage ratio would drop by the ratio of the converted amount to total assets. Whether this drop is endogenous or exogenous to the firm depends on who decides on conversion. Evidently, if the firm redeems an issue using a conversion-forcing call, the conversion decision is endogenous to the firm. Depending on the terms of the contract regarding the call
provision the firm may be given the privilege to call a convertible bond when it is in the money or even out of the money. However, for in the money conversions, if the conversion is initiated by bond holders this may be seen as an exogenous event to the firm, under the assumption that there is no stock price manipulation around the conversion threshold.

Whether a conversion falls into the endogenous or exogenous category depends on call provisions in the debt contract that determine the level of protection for an investor against a call by the firm. Generally speaking, convertible bonds may be (i) callable anytime (no protection), (ii) callable conditional on stock price passing a threshold and staying there for certain number of days (soft protection), (iii) not callable for a certain period (hard protection) and (iv) not callable at all (absolute protection). One suspects that voluntary conversions are more likely to be detected in categories (iii) and/or (iv). A subset of those conversions that occur during the call protection period are less likely to be influenced by the firm and therefore are considered the most exogenous conversions in section 1.4.

Why Issue Convertible Bonds?

A convertible bond is a hybrid security that mixes a straight bond features with those of a standard equity. More importantly, convertible bonds are state contingent securities, which may be transformed from a bond type into a equity type structure in different states of the world, usually as a function of the stock price. The contract specifies the conditions under which the conversion option may be exercised either at borrower’s discretion, i.e., a call or at lender’s discretion, i.e., a voluntary conversion. Given the state dependent structure of convertible bonds, we ask what is the use of the conversion option? And what does theory predict about capital structure policy at the conversion point?

There are two broad classes of theories explaining why firms issue convertible bonds. According to these theories convertible bonds are issued as a solution to problems concerning either (i) asymmetric information or (ii) agency problem. The first reason why firms issue convertible bonds addresses the asymmetric information between insiders (existing shareholders or management) and outsiders (lenders/bondholders) about the firm’s type. In a pooling equilibrium this results in higher financing costs for good types. Kim (1990) provides a model in which good type firms differentiate from lemons in a separating equilibrium by signally their type via the “conversion ratio.” Good types, with favorable information about future earnings, prefer to

---

5Consistent with sequential financing theory (Mayers (1998)) Korkeamaki and Moore (2004) show that the level of protection is inversely related to future investment: the lower the call protection, the sooner and the larger the investment.
absorb more equity risk by lowering dilution in equity after conversion and therefore offer a lower conversion ratio in equilibrium. The market reads this as a signal for good future earnings, and therefore reacts more favorably. In Stein (1992)' model firms can avoid the adverse-selection costs associated with direct equity sales by issuing equity “through the back door”. In this model, high quality firms issue convertible bonds instead of costly straight equity in order to delay their stock issuance to a time when uncertainty about the firm type is resolved. What separates a good type from a lemon is a special feature of the call option that can be exercised only when the stock price is higher than the conversion threshold. This model also rationalizes the mild negative announcement effect for issuance of convertible bonds. (Mikkelson (1985) and Lewis et al. (1999))

The second motivation for issuing convertible bonds addresses the agency problem between existing shareholders and bondholders. Agency costs can potentially cause two inefficient outcomes: (i) underinvestment due to “risk-shifting” behavior (Myers and Majluf (1984)), and (ii) over-investment or free cash flow (Jensen (1986)) due to “sequential-financing” problem. Risk-shifting arises when an investment decision is not observable by the bondholders and therefore is not contractable. Due to limited liabilities, shareholders (or management) may have incentive to invest in very risky projects at the expense of debt holders resulting in underinvestment in equilibrium. Green (1984) presents a model in which convertible bonds (or warrants) mitigate the risk-shifting problem by reducing the upside benefits of excessive risk taking for existing shareholders. In good states once a bond is converted, former bondholders would become new shareholders and therefore would share the upside gains with existing shareholders. Thus, by alleviating the incentive problem, convertible bonds offer a solution to underinvestment. Sequential-financing which involves an investment option with a future maturity date, is another cause of the incentive conflict between bondholders and management. Upfront funding (leaving the funds in the hands of a manager before uncertainty about the profitability of an investment project is resolved) may result in overinvestment in current non-profitable investments. In Mayers (1998)'s model convertibility solves the future financing problem by matching financing options with investment options. When the investment option is in the money, and therefore the stock price is on the rise, the conversion option will also be in the money. Conversion of debt into equity will free up debt capacity which then can be used to raise new capital for investment projects. This provides a solution to overinvestment, since up front financing is no longer required.

Although all explanations suggest that convertible bonds can be designed to induce managers to make efficient capital expenditure decisions, the only theory predicting a capital structure reaction to a conversion
is Mayers (1998). He predicts that a drop in leverage because of conversion frees up a firm’s debt capacity to finance new investments. He also provides empirical evidence showing an increase in financing activities (issuance of long-term debt and common stocks) following a conversion. However, the direction of reaction is not completely clear. On one hand Mayers (1998) provides support for capital structure rebalancing by predicting new debt financing after conversion. On the other hand the rise in equity financing following a conversion would predict just the opposite. Therefore, capital structure response to a conversion remains an empirical question.

1.4 The Response of Leverage to a Realized Conversion

1.4.1 Graphical Analysis

Figure 1.1 presents the time trend of the average leverage ratio, defined as the ratio of total book value of debt to total assets, relative to the quarter in which voluntary conversion occurs. This figure also presents the 90% confidence intervals, indicated by the bands around the averages. It is clear that there is a drop in average leverage immediately at conversion quarter. During the years before conversion the average leverage ratio is 34.0%, and 31.8% in $t-2$ and $t-1$ respectively while over the first year after conversion the leverage ratio is 27.7%. Therefore, the figure suggests a temporal break or discontinuity in leverage coinciding with voluntary conversion. Moreover, consistent with re-balancing theory, Figure 1.1 illustrates that the average leverage ratio bounces back between $t+1$ and $t+3$ to its original level at $t-8$. Another interesting observation in Figure 1.1 is the temporal rise in the average leverage ratio between periods $t-6$ to $t-3$ which is consistent with the pre-balancing theory in section 1.5. Whether or not these patterns are significant in the presence of other determinants of capital structure is investigated in the next section.

1.4.2 Diff-in-Diff Analysis at Conversion

To examine the impact of a realized conversion on leverage, we use Baker and Wurgler (2002) specification for difference in leverage, as follows:

$$\Delta L_{it} = \beta_0 + \beta_1 \text{Conversion Dummy}_{it} + \beta_2 X_{it} + \beta_3 L_{it-1} + \lambda_t + \epsilon_{it},$$

(1.2)
Figure 1.1: Leverage in event time leading to conversion: The sample includes firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database for which at least for one of their convertible bond issues a conversion event during the hard call protection period is recorded in FISD between 1995 and 2010. The figure presents average book leverage in event time relative to when one of the convertible bonds outstanding is voluntarily converted. Squares present the average leverage ratio for the sample at each quarter and bands around squares denote 90% confidence interval.

where $\Delta L_{it}$ is the first difference in leverage ratio for firm $i$ in year-quarter $t$, Conversion Dummy$_{it}$ is a treatment indicator variable that switches on in a conversion quarter, $X_{it}$ is a vector of conventional control variables or leverage including Log of Total Assets, Market to Book ratio, Profitability, Tangibility, Industry Median Leverage and Cash Flow Volatility. $\lambda_t$ captures the time fixed effect, which is equal to the sum of the year and quarter fixed effects. The coefficient of interest in this regression is $\beta_1$ which is effectively a difference-in-difference coefficient. It compares leverage changes before and after a conversion, between firm-quarter observations with and without a conversion. The benefit of diff-in-diff specification as compared to a simple difference equation for before and after conversion is that, by contrasting with a control group, it takes non-treatment changes in leverage into account. The results are unchanged if we use a difference model instead of difference-in-difference, where the dependent variable is the level of leverage instead of its first difference. The results are also robust to the inclusion of a lagged leverage ratio as a control.  

Table 1.5 presents the estimation results for equation (1.2) for three samples. In “Convertibles Sample” the leverage ratio for voluntary conversions is compared to that of all convertible bond issuers. Comparing

---

6The estimation results for both alternative specifications are not reported but are available upon request.
firms under conversion to the ones that never experience a conversion raises a valid concern about the effect of unobservable factors in the first group that may affect the changes in leverage, and therefore may weaken the comparison result. To address this concern, in “Conversions Sample” the former control group is restricted only to the firms that eventually experience a conversion. This ensures that possible genetic differences that might relate conversion to leverage changes are present in both the treatment and control groups. Finally, to compare the impact of all voluntary conversions to the more exogenous conversions that occur during the call protection period (when a firm is most restricted from influencing the conversion decision) in “Restricted Conversions” I change the treatment definition to the subset of voluntary conversions that occur during the call protection period. Similarly, the control group is chosen from the firms that eventually have such a conversion.

<table>
<thead>
<tr>
<th>Dependent Variable: Changes in Book Leverage</th>
<th>Converting Firms Sample Δ</th>
<th>Converting Firms Sample</th>
<th>Restricted Conversions Sample Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversion Dummy</td>
<td>-0.051</td>
<td>-0.054</td>
<td>-0.042</td>
</tr>
<tr>
<td></td>
<td>(-11.80)</td>
<td>(-14.12)</td>
<td>(-6.44)</td>
</tr>
<tr>
<td>Log(Assets)</td>
<td>0.001</td>
<td>0.005</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(2.25)</td>
<td>(1.18)</td>
<td>(2.57)</td>
</tr>
<tr>
<td>Market to Book</td>
<td>-0.003</td>
<td>-0.005</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(-6.39)</td>
<td>(-1.76)</td>
<td>(.84)</td>
</tr>
<tr>
<td>Profitability</td>
<td>-0.198</td>
<td>-0.403</td>
<td>-0.689</td>
</tr>
<tr>
<td></td>
<td>(-14.11)</td>
<td>(-4.46)</td>
<td>(-3.80)</td>
</tr>
<tr>
<td>Tangibility</td>
<td>0.012</td>
<td>0.041</td>
<td>0.171</td>
</tr>
<tr>
<td></td>
<td>(5.06)</td>
<td>(1.74)</td>
<td>(2.89)</td>
</tr>
<tr>
<td>Industry Med Lev</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>(5.61)</td>
<td>(.60)</td>
<td>(-.83)</td>
</tr>
<tr>
<td>Cash Flow Vol</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(.38)</td>
<td>(.09)</td>
<td>(.29)</td>
</tr>
<tr>
<td>Dividend Payer</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(-.61)</td>
<td>(-.17)</td>
<td>(-.22)</td>
</tr>
<tr>
<td>(Book Leverage) t-1</td>
<td>-0.167</td>
<td>-0.059</td>
<td>-0.248</td>
</tr>
<tr>
<td></td>
<td>(-45.93)</td>
<td>(-23.05)</td>
<td>(-12.26)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.050</td>
<td>0.029</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(12.58)</td>
<td>(2.80)</td>
<td>(7.85)</td>
</tr>
<tr>
<td>Year and Quarter FE</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>(17278)</td>
<td>(15189)</td>
<td>(873)</td>
</tr>
<tr>
<td>Obs</td>
<td>0.053</td>
<td>0.056</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>0.45</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Table 1.5: Change in leverage at conversion: This table presents estimation results for three samples. Convertibles sample consists of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database that have at least one convertible bond outstanding from 1995 to 2010. Conversions sample consists of all firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database for which at least for one of their convertible bond issues a conversion event is recorded in FISD between 1995 and 2010. Restricted conversions sample is a subset of Conversions sample, including only firm-quarter observations that a conversion happened during the call protection period. The table presents regression results, where the depended variable in each regression is the first difference in book leverage. t-statistics are reported in parenthesis.

Columns (1), (3) and (5) present simple correlations between Conversion Dummy and the leverage ratio. Similar to Figure 1.1 in all three samples a conversion results in a significant drop in leverage ratio of the
order of 4.4% to 5.1%. Relative to the average leverage ratio of 34% in quarters with no conversion this estimate translates into a relative decrease in leverage of 12% to 15%.

Specifications (2), (4) and (6) include additional controls variables, used in previous studies (Titman and Wessels (1988), Rajan and Zingales (1995), Graham et al. (1998) and MacKay and Phillips (2005)). The sign and significance of the coefficient estimates on these control variables are consistent with previous studies. Nevertheless, their inclusion has little effect on the estimated impact of the conversion on leverage. The only exception is the drop in significance of the impact in the last sample, Restricted Conversions although the sign stays the same. This might be the result of the shrinkage of the sample size (i.e., 59 conversions). In summary, Table 1.5 shows a statistically and economically significant impact of a realized conversion on leverage ratio that is robust to the inclusion of control variables and time fixed effects.

1.4.3 Leverage Dynamics After Conversion

In this section I investigate whether or not firms actively rebalance their capital structure in response to the significant changes illustrated in section 1.4.2. I use a specification similar to that for equation (1.2), where the dependent variable is the leverage level instead of its first difference. To compare the leverage ratio between firms who had a conversion in the past vs. those who did not, I run a separate regression for each quarter following a conversion for upto six quarters subsequent to the conversion. More precisely, in each regression the Conversion Dummy\(_{it-s}\) variable switches on if there is a conversion \(s\) quarters prior. Therefore, \(\beta_1\) captures the difference of the leverage ratio between observations with and without a conversion \(s\) quarters after the conversion. A significant \(\beta_1\) coefficient indicates no or partial rebalancing, whereas an insignificant estimate for \(\beta_1\) is shows full rebalancing.

Table 1.6 presents the estimation results for three samples, in panels A, B and C. On each row, columns (1) to (7) present the regression results for each of the \(s\) quarters following a conversion, \(s \in \{0, 1, 2, ..., 6\}\), i.e., a total of 21 regressions. For each regression, only the coefficient on Conversion Dummy\(_{it-s}\) is reported in this table, but all regressions include the control variables used in Table1.5 as well as year and quarter fixed effects. Panel A includes observations from the “Conversion Sample” in Table 1.5, where I use a one-quarter “no-event” cushion around each control and treatment quarter. In other words, I restrict the sample to only those observations for which no other event (conversion, call, etc.) is recorded in the quarter immediately before or after. Similarly, panel C uses a four-quarter cushion around both control and treatment
observations. Panel B which uses exactly the same sample as the Conversion Sample in Table 1.5 uses a one-year cushion for the treatment and four-year cushion for the control group. Adding “no-event” cushions helps to purify the treatment effect, here conversion impact. For example, when there is a series of conversions the outcome behavior might be different than when there is a single conversion. Similarly, a good control or treatment observation should not have any recent call or a conversion, otherwise one has to deal with a mixture of different treatments. These examples illustrate the benefits of the “no-event” cushion and how lack of such cushions may result in an unreliable estimates.

### Table 1.6: Leverage dynamics post-conversion

This table presents estimation results for three samples in seven periods, total of 21 regressions. All three panels contain firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database that have at least one convertible bond outstanding which eventually will be converted at some point between 1995 and 2010. In panel A observations that have any event in period before or after \( t \) are dropped from both treatment and control groups. In panel B observations that have any event in previous or next period are dropped from treatment and the ones that have any event in the past or next four periods with respect to \( t \) are dropped from control group. In panel C observations that have any event in past or next four periods are dropped from both treatment and control groups. The table presents regression results, where the dependent variable in each regression is the book leverage level in periods marked on the second row of the table. Quarter \( t \) is the conversion quarter, quarter \( t+1 \) is one year after conversion, quarter \( t+2 \) is two years after and so on. Conversion dummy is an indicator variable that switches on when there is a conversion in period \( t \). All regressions include control variable used in Table 1.5 as well as year and quarter fixed effects. t-statistics are reported in parenthesis.

<table>
<thead>
<tr>
<th>Dependent Variable: Book Leverage</th>
<th>t+1</th>
<th>t+2</th>
<th>t+3</th>
<th>t+4</th>
<th>t+5</th>
<th>t+6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversion Dummy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-0.0671</td>
<td>-0.0577</td>
<td>-0.0336</td>
<td>-0.0199</td>
<td>-0.0095</td>
<td>-0.0173</td>
</tr>
<tr>
<td>R²</td>
<td>0.21</td>
<td>0.22</td>
<td>0.22</td>
<td>0.24</td>
<td>0.25</td>
<td>0.24</td>
</tr>
<tr>
<td>Obs</td>
<td>1374</td>
<td>1379</td>
<td>1371</td>
<td>1367</td>
<td>1364</td>
<td>1357</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variable: Book Leverage</th>
<th>t+1</th>
<th>t+2</th>
<th>t+3</th>
<th>t+4</th>
<th>t+5</th>
<th>t+6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversion Dummy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-0.0663</td>
<td>-0.0577</td>
<td>-0.0336</td>
<td>-0.0199</td>
<td>-0.0095</td>
<td>-0.0173</td>
</tr>
<tr>
<td>R²</td>
<td>0.21</td>
<td>0.22</td>
<td>0.22</td>
<td>0.24</td>
<td>0.25</td>
<td>0.24</td>
</tr>
<tr>
<td>Obs</td>
<td>801</td>
<td>802</td>
<td>800</td>
<td>795</td>
<td>796</td>
<td>792</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variable: Book Leverage</th>
<th>t+1</th>
<th>t+2</th>
<th>t+3</th>
<th>t+4</th>
<th>t+5</th>
<th>t+6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversion Dummy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-stat</td>
<td>-0.0664</td>
<td>-0.0577</td>
<td>-0.0336</td>
<td>-0.0199</td>
<td>-0.0095</td>
<td>-0.0173</td>
</tr>
<tr>
<td>R²</td>
<td>0.21</td>
<td>0.22</td>
<td>0.22</td>
<td>0.24</td>
<td>0.25</td>
<td>0.24</td>
</tr>
<tr>
<td>Obs</td>
<td>734</td>
<td>733</td>
<td>732</td>
<td>729</td>
<td>729</td>
<td>726</td>
</tr>
</tbody>
</table>

The findings in Table 1.6 support a fairly quick rebalancing in capital structure in response to negative leverage shock caused by debt conversion. Columns (1) to (3) in Panel B of Table 1.6 show that the 5.1% drop in leverage level in the conversion quarter, shrinks to 3.5% and then to 2.6% in the second and third quarters and then to zero (insignificant) thereafter. Similarly, in panel C the leverage gap between the control and treatment group closes after three quarters. Finally, in panel A, where I use only a one-year no-event cushion around the treatment and control quarters, I observe fairly similar results. Although the estimated
drop in leverage is slightly larger than panel B and C, it fades away after four quarters. The 6.7% drop in leverage caused by conversion, contracts to 5.8% in the second quarter, 3.4% in the third quarter, 2.0% in the fourth quarter and finally disappears after that. This corresponds to a 1.7% increase in leverage in each quarter following conversion, which translates into an approximately 5% speed of adjustment (SOA) relative to the average leverage ratio before conversion. In sum, consistent with capital structure rebalancing, I observe that a drop in the leverage ratio in response to a voluntary conversion is reversed in less than a year.

One must acknowledge the possibility of estimation bias due to selection (Heckman (1979)) in my results. After all, conversion is not a randomly assigned treatment. While previous studies (Constantinides and Grundy (1987) and Asquith and Mullins (1991)) found conversion motivations which boils down to observable cash flow differences between coupon vs. dividend streams, the presence of unobservable factors affecting a conversion makes a causal interpretation virtually impossible. In the interest of saving space, I intentionally skip correcting for such a bias in the estimates for two reasons. First, the choice of control group in the Conversion Sample limits the comparison to only firms that eventually have a conversion. This guarantees that if there is any unobservable factor in their genes affecting the decision to take the treatment, (i.e. to convert,) it is common between the treatment and control group. The robustness of my results makes it less likely for leverage changes to be affected by such unobservables. Second, the main purpose of this section is to offer a descriptive rather than a causal explanation for what occurs after a conversion. Having said that, I keep the caveat in mind and use it as a motivation for the quasi-experiment analysis in the next section.

1.5 The Response of Leverage to an Anticipated Conversion

In this section I examine the possibility of rebalancing even before the realization of a leverage shock. I ask if there is any evidence confirming a preemptive increase in leverage when a drop due to conversion is anticipated. The key challenge here is to measure when such a leverage shock is truly anticipated. The unique feature of convertible bonds that links debt conversion to the level of the underlying commodity price enables us to have a tractable measure for “shock anticipation.” When the stock price passes the conversion price threshold (i.e., conversion premium=0) the conversion option becomes in the money, meaning that bond holders have the right to exercise their option by exchanging their bonds for equity. This generates a discontinuity (negative jump) in the expected amount of debt, and consequently in the anticipated leverage ratio. In the next section, I exploit this discontinuity for identification using regression discontinuity design.
1.5.1 Empirical Strategy: A Regression Discontinuity Design

An ideal framework to study the existence of pre-balancing in the data, i.e., whether or not the leverage ratio responds to changes in anticipated leverage, is a randomized experiment, where treatment (change in anticipated leverage) is assigned randomly to a subset of observations (treatment group) and their resulting behavior is contrasted to an untreated (control) group. Given that running a controlled experiment in corporate finance is virtually impossible to implement, I turn to the closest proxy: a quasi-experiment. Regression discontinuity design provides such a quasi-experimental framework where the probability of receiving the treatment changes discontinuously as a function of one of the variables, here stock price. As Lee (2008) points out in such a design once one gets close to the discontinuity threshold (i.e., conversion price), one can assume that observations are randomly assigned to each side of the threshold. Firm-quarter observations for which the stock price is greater (less) than the conversion price and therefore are located to the right (left) of the threshold constitute the treatment (control) group. Specifically, the treatment variable $ITM_{it}$ (In The Money) is defined as:

$$ITM_{it} = \begin{cases} 
1 & \pi_{it} > 0 \\
0 & \pi_{it} \leq 0 
\end{cases}$$

(1.3)

where, $i$ and $t$ index firm and year-quarter observations, and $\pi_{it}$ is the conversion premium for the underlying convertible bond (the distance between stock price and conversion price as defined in equation (1.1)). My sample consists of non-financial firm-quarter observations in merged FISD-COMPUSTAT between 1995 and 2010 for which a convertible bond is outstanding, eventually, but not currently, experience a voluntary conversion and have four quarters of “no-event” cushions before and after the quarter they show up in the sample.

Table 1.7 presents the summary statistics for the treatment ($ITM = 1$) and control ($ITM = 0$) subsamples. As the first indication of pre-balancing, the second and third rows reveal an economically and statistically significant increase in book and market leverage ratios when comparing out of the money (control) to in the money (treatment) samples. From left to right of the conversion threshold, the average book leverage increases by 0.03, from 0.27 to 0.30, an 11% increase. Similarly, the average market leverage leaps by 0.06, from 0.20 to 0.26, a 32% raise. In contrast to changes in average leverage ratio, there is no significant heterogeneity in any of the leverage related firm characteristics on other rows (all p-values are
greater than 0.10). This makes the two sub-samples a close match, which is consistent with the notion of random assignment of treatment, making perfect treatment/control groups for the experiment. Although we control in my regression model for those firm characteristics, these similarities make my inferences from unconditional comparison in non-parametric sections a valid analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Out of the Money Sample</th>
<th>In the Money Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs</td>
<td>Mean</td>
</tr>
<tr>
<td>Conversion Premium</td>
<td>104</td>
<td>-21.71</td>
</tr>
<tr>
<td>Book Leverage</td>
<td>98</td>
<td>0.27</td>
</tr>
<tr>
<td>Market Leverage</td>
<td>98</td>
<td>0.20</td>
</tr>
<tr>
<td>Log(Assets)</td>
<td>104</td>
<td>7.67</td>
</tr>
<tr>
<td>Market to Book Ratio</td>
<td>104</td>
<td>2.01</td>
</tr>
<tr>
<td>Profitability</td>
<td>104</td>
<td>0.03</td>
</tr>
<tr>
<td>Tangibility</td>
<td>104</td>
<td>0.32</td>
</tr>
<tr>
<td>Cash Flow Volatility</td>
<td>104</td>
<td>0.35</td>
</tr>
<tr>
<td>Dividend Payer</td>
<td>104</td>
<td>0.38</td>
</tr>
<tr>
<td>Cash Flow</td>
<td>104</td>
<td>0.15</td>
</tr>
<tr>
<td>Investment/Capital</td>
<td>104</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Table 1.7: Effect of anticipated conversions: This table presents summary statistics - averages, medians, and standard deviations - for two samples. Out of the money sample consists of non-financial firm-quarter observations in merged FISD-COMPUSTAT between 1995 and 2010 that satisfy the following conditions: (1) have at least one convertible bond outstanding, (2) experience at least one conversion after the quarter they show up in this table, (3) do not experience any of the events listed in Table 1.11 such as conversion, call, repurchase, etc. in neither the past four quarters, nor in the next four quarters and (4) their stock price is below the conversion price, i.e., conversion premium < 0. In the money sample has the same conditionals (1) to (3) as out of the money sample, but with an opposite last condition, i.e., (4'): their stock price is above the conversion price, i.e., conversion premium > 0. One sided p-values are calculated for on two sample mean comparisons test for each row.

1.5.2 Unconditional Response

Figure 1.2 presents the average leverage ratio for firm-quarter observations in treatment and control groups in four sub-samples constructed based on the conversion premium for the underlying convertible bond for the firm i at time t. The four bins of conversion premium are: deep out of the money, with π_{it} ∈ (-50%, -25%); out of the money with π_{it} ∈ (-25%, 0%); in the money with π_{it} ∈ (0%, +25%) and deep in the money, π_{it} ∈ (+25%, +50%). Similar to Figure 1.1, squares represent book leverage averages and surrounding bands denote the 90% confidence intervals. Immediately, we observe that there is a discontinuous increase in leverage ratio, around the conversion threshold. When comparing firms located to the left of conversion threshold (π_{it} ∈ (-25%, 0%)) with the ones to the right (π_{it} ∈ (0%, +25%)) in period t, the average leverage ratio jumps from 0.30 to 0.34 in the same period. This effect seems to be persistent in (t - 1) and (t + 1), indicating that pre-balancing begins one quarter prior and continues to the quarter immediately after t, t being the quarter in which stock price crosses the conversion threshold. Note, Figure 1.2 provides two simple falsification tests at π_{it} = -25 and π_{it} = +25: unlike at the conversion threshold (π_{it} = 0) we observe no
break in average leverage ratio in either of these two thresholds. This confirms that my findings are not driven by a general pattern in data and are specific to the conversion threshold.

![Image of Figure 1.2: Leverage response to proximity to conversion]

**Figure 1.2: Leverage response to proximity to conversion:** The sample includes firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database with a convertible bond issue outstanding. Observations are sorted based on conversion premium at quarter t into each of the four bins between -50%, -25%, 0, 25% and 50% on horizontal axis. The figure presents average book leverage in period t (as well as (t-1) and (t+1) ) for the portfolio of firms in each bin. Squares present the average leverage ratio and bands around squares denote 90% confidence interval for the average leverage in quarter t. Marked bands denote the confidence interval for period t.

Table 1.8 presents the same data as in Figure 1.2, but for an extended time horizon, beginning from four quarters before and continuing to four quarters after crossing the conversion threshold. Similar to Figure 1.2, columns (1) to (4) present the average leverage ratio for four bins of conversion premiums. The last column compares the average leverage ratio for the middle two bins, using a two sample mean comparison test. The p-values show that data rejects the null hypothesis in favor of a jump at the conversion threshold, only in periods (t-1), t and (t+1), at 0.10, 0.06 and 0.05 levels. This shows that pre-balancing does not begin until the period before crossing the conversion threshold and stops one period after that.

In sum, consistent with the pre-balancing hypothesis, the unconditional comparison of the average leverage ratio for the firm-quarter observations to the left and right of the conversion threshold reveals an economically and statistically significant increase in leverage. This preemptive increase in leverage tends to start one quarter before crossing the threshold and continues during and after that quarter.
Table 1.8: Leverage response to distance from anticipated conversion: For each period (row) columns (1) to (4) presents average book leverage value for non-financial firm-quarter observations in merged FISD-COMPUSTAT between 1995 and 2010 with a convertible bond issue outstanding, for which conversion premium falls into one of the four bins between -50%, -25%, 0, 25% and 50%. For example, column (1) shows the average leverage value in quarter (t-4) to (t+4) for the group of firms with an outstanding convertible bond for which $-50% < \text{conversion premium} < -25%$ in quarter t. Last column reports the one-sided p-values for testing if leverage ratio for firms in column (2) is less than leverage ratio for the ones in (3). All leverage means are reported in percentage.

1.5.3 Parametric and Non-parametric Analysis

In this section, I expand the size of the window around the conversion threshold to the entire sample. I do so taking the standard approach illustrated in Lee and Lemieux (2010) by controlling for possible information contained in the distance between stock price and the conversion price threshold using a polynomials in the conversion premium $\pi_{it}$. Allowing for different polynomials for observations on the left-hand side of the threshold $P_l(\pi_{it}, \gamma^l)$ and on the right-hand side of the threshold $P_r(\pi_{it}, \gamma^r)$ gives:

$$L_{it} = \beta_0 + \beta_1 ITM_{it} + \beta_2 X_{it} + P_l(\pi_{it}, \gamma^l) + P_r(\pi_{it}, \gamma^r) + \lambda_t + \mu_i + \epsilon_{it}$$

(1.4)

where $L_{it}$ is the leverage ratio for firm $i$ in year-quarter $t$, $ITM_{it}$ is the treatment dummy (defined in equation (1.3)) which isolates the effect of discontinuity in anticipated leverage, $X_{it}$ is a vector of conventional control variables or leverage including Log of Total Assets, Market to Book ratio, Profitability, Tangibility, Industry Median Leverage and Cash Flow Volatility. $\mu_i$ is a firm fixed effect, $\lambda_t$ is a time fixed effect, which is equal to the sum of the year and quarter fixed effects, and $\epsilon_{it}$ is a random error term. The coefficient of interest in this regression is $\beta_1$ which represents the impact of a drop in anticipated leverage due to near future debt conversion on the current leverage ratio. A positive and significant $\beta_1$ rejects the null hypothesis of no pre-balancing (i.e., $H_0 : \beta_1 = 0$) in favor of the pre-balancing alternative.
Table 1.9 presents the estimation results for the entire sample\(^7\). All specifications include year and quarter fixed effects and results for both models with and without firm fixed effects are presented as indicated at the bottom of the table. The first column, presenting results for the linear polynomial case, reveal that as soon as the conversion option becomes in the money, meaning a drop in leverage is anticipated, average leverage leaps by 7.4%. Relative to an average book leverage ratio of approximately 34.1%, this estimate translates into a relative preemptive increase in leverage of almost 21%. The second column confirms a positive jump in leverage, although at a lower significance level due to inclusion of the firm fixed effect terms.

Columns (1) and (2) also show that there is indeed information about leverage changes contained in the distance between stock price and the conversion threshold. The negative slope for the linear term for the first-order polynomial in these columns (as well as for higher order polynomials in other columns) on both sides of the conversion threshold shows a general decreasing trend in leverage as the stock price increases. This pattern is also depicted in Figure 1.4 in appendix which presents the graphical version of the results in column (1). The downward slope of leverage in conversion premium or stock price might be due to debt retirement and equity issuance as pointed out by Hovakimian et al. (2001). They find that when stock price rises, firms generally issue more equity and retire debt.

My result is robust for a higher order polynomial and inclusion of leverage control variables. Specifications (3) to (4) expand the model by using fourth-order\(^8\) polynomial instead of linear. Similarly, in columns (5) and (6) the model is further expanded by incorporating additional control variables for leverage introduced in equation (1.2). None of these modifications have any impact on the estimated preemptive increase in leverage. Even in the most saturated model in column (6) which includes a fourth-order polynomial, leverage control variables plus time and firm fixed effects, still there is a significant jump in leverage of 5.4% around the conversion threshold, which translates to about 16% relative to the average leverage.

Finally, Table 1.10 presents results for the non-parametric analysis. In this table I consider three possible window sizes around the conversion threshold, i.e., \((-12\%, +12\%)\), \((-25\%, +25\%)\), and \((-50\%, +50\%)\). Each column presents the difference in average leverage ratio for observations to the right and to the left of the conversion threshold in each of the three windows. Panel A to E presents changes in the average leverage ratio in period \((t-2)\) to \((t+2)\), where \(t\) is the quarter in which the stock price passes the conversion threshold. Columns (1) and (2) show that for a narrow bandwidth of 12% the result disappears. However, if one

---

\(^7\)I excluded outliers with extremely negative conversion premium values, i.e., \(\pi_{it} < -1200\%\). This results in losing only 9 observations out of 639—about 1% reduction in sample size.

\(^8\)In an unreported regression I also use a polynomial of order five, but the main result is unchanged.
<table>
<thead>
<tr>
<th>Coefficients</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the Money (ITM)</td>
<td>0.074</td>
<td>0.019</td>
<td>0.072</td>
<td>0.070</td>
<td>0.061</td>
<td>0.054</td>
</tr>
<tr>
<td>(2.90)</td>
<td>(1.33)</td>
<td>(2.46)</td>
<td>(2.42)</td>
<td>(2.25)</td>
<td>(2.01)</td>
<td></td>
</tr>
<tr>
<td>Conversion Premium</td>
<td>-0.032</td>
<td>-0.002</td>
<td>-0.136</td>
<td>-0.139</td>
<td>-0.103</td>
<td>-0.111</td>
</tr>
<tr>
<td>(-4.96)</td>
<td>(-.46)</td>
<td>(-4.40)</td>
<td>(-4.51)</td>
<td>(-3.52)</td>
<td>(-1.84)</td>
<td></td>
</tr>
<tr>
<td>Conversion Premium*ITM</td>
<td>-0.142</td>
<td>-0.192</td>
<td>-0.582</td>
<td>-0.562</td>
<td>-0.857</td>
<td>-0.745</td>
</tr>
<tr>
<td>(-2.37)</td>
<td>(-.47)</td>
<td>(-1.31)</td>
<td>(-1.28)</td>
<td>(-2.09)</td>
<td>(-1.84)</td>
<td></td>
</tr>
<tr>
<td>Conversion Premium$^2$</td>
<td>-0.065</td>
<td>-0.069</td>
<td>-0.040</td>
<td>-0.044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-3.93)</td>
<td>(-4.20)</td>
<td>(-2.60)</td>
<td>(-2.81)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion Premium$^2$*ITM</td>
<td>3.274</td>
<td>3.223</td>
<td>4.666</td>
<td>3.965</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.58)</td>
<td>(1.57)</td>
<td>(2.42)</td>
<td>(2.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion Premium$^3$</td>
<td>-0.010</td>
<td>-0.011</td>
<td>-0.006</td>
<td>-0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-3.37)</td>
<td>(-3.67)</td>
<td>(-2.16)</td>
<td>(-2.36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion Premium$^4$*ITM</td>
<td>-6.374</td>
<td>-6.368</td>
<td>-8.876</td>
<td>-7.432</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-1.79)</td>
<td>(-1.81)</td>
<td>(-2.67)</td>
<td>(-2.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion Premium$^4$</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-2.85)</td>
<td>(-3.13)</td>
<td>(-1.82)</td>
<td>(-2.00)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion Premium$^4$*ITM</td>
<td>3.861</td>
<td>3.922</td>
<td>5.259</td>
<td>4.327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.95)</td>
<td>(2.01)</td>
<td>(2.85)</td>
<td>(2.37)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leverage Med Industry</td>
<td>0.368</td>
<td>0.078</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.67)</td>
<td>(4.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log(Assets)</td>
<td>0.016</td>
<td>0.095</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1.43)</td>
<td>(3.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market to Book</td>
<td>0.032</td>
<td>0.038</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5.82)</td>
<td>(6.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profitability</td>
<td>-0.536</td>
<td>-0.473</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-3.22)</td>
<td>(-2.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tangibility</td>
<td>0.197</td>
<td>0.197</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.63)</td>
<td>(2.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash Flow Volatility</td>
<td>0.063</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dividend Payer</td>
<td>-0.024</td>
<td>-0.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-1.39)</td>
<td>(-1.40)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.374</td>
<td>0.380</td>
<td>0.052</td>
<td>-0.600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12.07)</td>
<td>(13.49)</td>
<td>(.55)</td>
<td>(-2.94)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year and Quarter Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Fixed Effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>486</td>
<td>486</td>
<td>486</td>
<td>486</td>
<td>458</td>
<td>458</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.21</td>
<td>0.87</td>
<td>0.097</td>
<td>0.88</td>
<td>0.29</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table 1.9: Discontinuity regressions for the entire sample: The sample includes all firm-quarter observations between 1995 and 2010 from non-financial firms in the merged FISD-COMPUSTAT database that satisfy the following conditions: (1) have at least one convertible bond outstanding, (2) experience at least one conversion after the quarter they show up in this table, and (3) do not experience any of the events listed in Table 1.11 such as conversion, call, repurchase, etc. in neither the past four quarters, nor the next four quarters. The table presents regression results, where the depended variable in each regression is the book leverage. ITM (In The Money) is a dummy variable which switches on when stock price is above conversion price, i.e., conversion option is in the money. t-statistics are reported in parenthesis.
expands the bandwidth to 25% in columns (3) and (4), one observes a significant increase in leverage in the range of 3.3% to 4.9% in quarters t and t+1. Likewise, further increases the bandwidth to 50%, suggest pre-balancing in leverage in magnitudes of 4.2% to 5.4% in quarters t-1, t and t+1. The reason we do not observe the result in columns (1) and (2) can reflect two possibilities. First, the minimum sample size is only 66, as opposed to 139 and 266 for the wider windows. Second, the measurement error in estimating the conversion premium described in section 1.2.3 makes it a noisier measure in narrower windows.

| Variable | Conversion Premium Range | | | | | |
|----------|-----------------|-----------------|-----------------|-----------------|-----------------|
| ITM      | (-12,+12)       | (-12,+12)       | (-25,+25)       | (-25,+25)       | (-50,+50)       | (-50,+50)       |
|          | ITM (-1.11)     | ITM (.14)       | ITM (.78)       | ITM (1.11)      | ITM (1.48)      | ITM (1.38)      |
|          | 0.0006           | 0.028           | 0.038           | 0.03            | 0.03            | 0.03            |
|          | 0.27 [66]        | 0.00029 [66]    | 0.2 [141]       | 0.0088 [141]    | 0.14 [266]      | 0.0071 [266]    |
|          | R² [Obs]         |                  |                  |                  |                  |                  |
| Panel (A) |                  |                  |                  |                  |                  |                  |
| ITM      | 0.023            | 0.014            | 0.016            | 0.031           | 0.045           | 0.042           |
|          | (-.81)           | (.34)            | (.85)            | (1.20)          | (2.04)          | (2.05)          |
|          | 0.33 [67]        | 0.0018 [67]      | 0.22 [139]       | 0.01 [139]      | 0.13 [268]      | 0.016 [268]     |
|          | R² [Obs]         |                  |                  |                  |                  |                  |
| Panel (B) |                  |                  |                  |                  |                  |                  |
| ITM      | -0.012           | 0.03             | 0.033            | 0.049           | 0.064           | 0.051           |
|          | (-.51)           | (.71)            | (1.87)           | (1.84)          | (2.54)          | (2.43)          |
|          | 0.37 [66]        | 0.0077 [66]      | 0.22 [141]       | 0.024 [141]     | 0.15 [268]      | 0.022 [268]     |
|          | R² [Obs]         |                  |                  |                  |                  |                  |
| Panel (C) |                  |                  |                  |                  |                  |                  |
| ITM      | -0.007           | 0.034            | 0.038            | 0.049           | 0.05            | 0.046           |
|          | (-.33)           | (.77)            | (1.90)           | (1.77)          | (2.30)          | (2.13)          |
|          | 0.41 [66]        | 0.0091 [66]      | 0.19 [140]       | 0.022 [140]     | 0.16 [268]      | 0.017 [268]     |
|          | R² [Obs]         |                  |                  |                  |                  |                  |
| Panel (D) |                  |                  |                  |                  |                  |                  |
| ITM      | -0.018           | -0.003           | 0.008            | 0.014           | 0.03            | 0.025           |
|          | (-.69)           | (.08)            | (.29)            | (.57)           | (1.02)          | (1.21)          |
|          | 0.18 [67]        | 0.0001 [67]      | 0.13 [143]       | 0.0023 [143]    | 0.12 [270]      | 0.0054 [270]    |
| Year and Quarter FE | No | No | No | No | Yes | Yes |
| R² [Obs] |                  |                  |                  |                  |                  |                  |

Table 1.10: Non-parametric regressions for windows of conversion premium: The sample includes firm-quarter observations between 1995 and 2010 from non-financial firms in the merged FISD-COMPUSTAT database that satisfy the following conditions: (1) have at least one convertible bond outstanding, (2) experience at least one conversion after the quarter they show up in this table, (3) do not experience any of the events listed in Table 1.11 such as conversion, call, repurchase, etc. in neither the past four quarters, nor the next four quarters, and (4) fall in each of the three conversion premium windows of (-12,+12) in columns (1) and (2), (-25,+25) in columns (3) and (4), or in (-50,+50) in columns (5) and (6). The table presents non-parametric regression results, where the depended variable in each regression is the book leverage. ITM (In The Money) is a dummy variable which switches on when stock price is above conversion price, i.e, conversion option is in the money. Leverage levels are de-trended and deseasonalized using year and quarter fixed effects. t-statistics are reported in parenthesis.

In sum, I find robust evidence of pre-balancing using different parametric and non-parametric model specifications. Firms tend to increase their leverage level preemptively by 10% to 21% relative to their current leverage level as soon as they anticipate a drop in leverage. This finding is consistent with the trade-off theory stating that managers actively rebalance their firms’ capital structure towards a target. I find that they not only rebalance, but also pre-balance when they anticipate a change in leverage in the near future.
1.5.4 Heterogeneity in Leverage Response: Quantile Analysis

In this section, I examine whether the pre-balancing effect which we found in the previous section is a universal phenomena for all firms, or if it varies as a function of their leverage ratio. I repeat the exercise in the previous section by estimating a quantile regression version of the model in equation (1.4), adopting a polynomial of order one as well as year, quarter and industry fixed effects. The main difference is that in the previous section, $\beta_1$ was an estimate for the average treatment effect (ATE), whereas here it estimates the quantile treatment effect (QTE) (Koenker (2005)) for quantiles of the book leverage. In the interest of saving space, I focus only on the $\beta_1$ and skip reporting estimation results for the full model.

![Leverage Pre-balancing Coefficient vs Quantiles of Leverage](image)

**Figure 1.3: Quantile treatment effect (QTE) for leverage response**. The sample includes all firm-quarter observations between 1995 and 2010 from non-financial firms in the merged FISD-COMPUSTAT database that satisfy the following conditions: (1) have at least one convertible bond outstanding, (2) experience at least one conversion after the quarter they show up in this table, and (3) do not experience any of the events listed in Table 1.11 such as conversion, call, repurchase, etc. in neither the past four quarters, nor the next four quarters. The solid line in this figure presents pre-balancing coefficient in different quantiles of the leverage level, before crossing the conversion price (ITM coefficient for quantile version of the results in Table 1.9), dotted lines present 90% confidence intervals for this coefficient and dashed line is the OLS estimation result for the same coefficient.

Figure 1.3 presents the pre-balancing coefficient, $\beta_1$, for quantiles of conditional leverage ratio in equation (1.4). The horizontal axis denotes quantiles of the conditional book leverage, and the vertical axis shows the coefficient value. Overall, we observe a hump-shaped curve with a maximum of 0.13 around the 70th percentile. This figure shows that the pre-balancing coefficient increases in leverage for a wide range of low
and medium conditional leverage ratios. Specifically, the pre-balancing coefficient rises from $\beta_{1,\tau=0.95} = 0.03$ to $\beta_{1,\tau=0.70} = 0.13$. For high leverage ratios the pattern reverses. The increase in the conditional leverage ratio falls from $\beta_{1,\tau=0.70} = 0.13$ to $\beta_{1,\tau=0.95} = 0.05$.

The increasing section of Figure 1.3 may be seen as a counter example for the debt overhang hypothesis (Myers (1977), Lang et al. (1996), Hennessy (2004)). In our context, the debt overhang hypothesis would predict a decreasing pattern for pre-balancing as a function of leverage. This is true because as we move from lower to higher conditional leverage levels on the horizontal axis, the increase in conditional leverage ratio implies that the firm’s existing size of debt is larger relative to the total assets. The larger size of debt, which increases expected financial distress costs, makes it harder for the firms to raise new debt. This then predicts that the company has to forgo more new positive investment projects. Therefore, during the pre-balancing the amount of new debt or equivalently the increase in the leverage ratio must be a decreasing function in current size of debt. This is in contradiction to the increasing section of the Figure 1.3, which represents a wide range of conditional leverage ratios between the 5th to the 70th percentiles. As we see in this figure, it is only in high conditional leverage ratios, (i.e. above the 70th percentile,) that consistent with debt overhang prediction, the pre-balancing is decreasing in conditional leverage.

Note that my identification strategy rules out the possibility that the hump-shaped pattern in Figure 1.3 could be the result of changes in observable firm characteristics such as firm credit rating or investment opportunities (which I control for). This is due to the fact that under a mild assumption of no discontinuity in error terms around the conversion threshold one can assume that observations are randomly assigned to the control and treatment groups, and therefore selection is not a concern in this design. Another such example is the effect of the size of convertible debt issue. Essig (1991) provides evidence that the relative size of convertible debt is larger for firms with larger leverage ratios. Therefore, one may argue that the increasing pattern in the first part of Figure 1.3 is due to the relatively larger size of the convertible issues for firms in higher quantiles of leverage ratio. In other words, firms pre-balance more because they have larger issues of convertible bonds and therefore larger anticipated increases in leverage. To investigate such possibilities, I include an additional variable $\text{Convertible Size} = \frac{\text{TotalConvertible Debt Outstanding}}{\text{Total Debt}}$ in equation (1.4). However, inclusion of the $\text{Convertible Size}$ has virtually no impact on the hump-shaped pattern in Figure 1.3.
1.6 Concluding Remarks

This study shows that firms actively re-balance their capital structure positions in response to both realized and anticipated change in capital structure. Using a regression discontinuity design I find that firms pre-balance their leverage ratio by 0.054 to 0.074 (a 16% to 21% increase relative to its previous level) in response to an anticipated drop in leverage caused by near future voluntary debt conversion. I also show that once an issue is voluntarily converted, the gap in the leverage ratio between converted and non-converted groups disappears within a year. Thus, my finding is consistent with a trade-off theory of capital structure, where firms actively re-balance and even pre-balance their capital structures toward a target.

Additionally, I find that pre-balancing in capital structure varies as a function of leverage ratio. More specifically, a quantile treatment effect analysis indicates that pre-balancing is a hump-shaped function of leverage ratio, with a peak around the 70th percentile of conditional leverage ratio. I argue that the upward-sloping section of this curve which covers a wide range of low and medium leverage ratios, may be interpreted as a counter example for the debt overhang theory.

While shedding light on how capital structure responds to an anticipated change, this study also provides opportunities for addressing new questions. One such question concerns the effect of an anticipated change in capital structure on other financial policy choices such as cash holding policy, dividend policy, and so on. Another issue that is still an open question according to Graham and Leary (2011) concerns the value relevance of capital structure. Given that most of the current studies find a flat value function one wonders whether capital structure is truly irrelevant or whether better identification strategies may result in different conclusions? A regression discontinuity design to study stock return changes around discontinuity threshold (which has been recently used by other researchers (Cuat et al. (forthcoming))), would provide an effective identification strategy to answer this question. Additionally, another question concerns real and financial implications of a change in capital structure. With realized conversions in data exogenous to firm’s decisions, one could potentially answer this question in a local average treatment effects analysis (LATE) design (Angrist and Pischke (2008)) using ITM as the treatment assignment indicator and actual conversion as the compliance (treatment-taking) indicator. These questions are all on my future research agenda.
1.7 Appendix

Figure 1.4: Leverage discontinuity at conversion price: The sample includes firm-quarter observations from non-financial firms in the merged FISD-COMPUSTAT database between 1995 and 2010 that satisfy the following conditions: (1) have at least one convertible bond outstanding, (2) experience at least one conversion after the quarter they show up in this table, and (3) do not experience any of the events listed in Table 1.11 such as conversion, call, repurchase, etc. in neither the past four quarters, nor the next four quarters. Dots denote local averages for each bin of book leverage, where bin bandwidth = 6.13%, and two separate linear fits on each side of zero conversion premium (or equivalently where stock price = conversion price).
<table>
<thead>
<tr>
<th>Code</th>
<th>Action</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Balance of Issue Called</td>
<td>278</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>291</td>
</tr>
<tr>
<td>C</td>
<td>Issue Converted</td>
<td>351</td>
<td>13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>365</td>
</tr>
<tr>
<td>E</td>
<td>Entire Issue Called</td>
<td>545</td>
<td>17</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>564</td>
</tr>
<tr>
<td>I</td>
<td>Initial Offering of An Issue</td>
<td>1633</td>
<td>461</td>
<td>14</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td>2160</td>
</tr>
<tr>
<td>II</td>
<td>Initial Offering Increase</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>IL</td>
<td>Initial Load</td>
<td>310</td>
<td>14</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>325</td>
</tr>
<tr>
<td>IM</td>
<td>Issue Matured</td>
<td>354</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>365</td>
</tr>
<tr>
<td>IRM</td>
<td>Issue Remarked</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>IRP</td>
<td>Issue Repurchased</td>
<td>464</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>494</td>
</tr>
<tr>
<td>L</td>
<td>Release Of Property</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>N</td>
<td>Not Available</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>O</td>
<td>Optional Sinking Fund Increase</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>OA</td>
<td>Over-Allotment</td>
<td>871</td>
<td>43</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>915</td>
</tr>
<tr>
<td>P</td>
<td>Part of an Issue Called</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>76</td>
</tr>
<tr>
<td>PIK</td>
<td>Pay-in-kind</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>R</td>
<td>Reorganization</td>
<td>122</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>148</td>
</tr>
<tr>
<td>REV</td>
<td>Review</td>
<td>1755</td>
<td>184</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td>1955</td>
</tr>
<tr>
<td>RO</td>
<td>Reopening</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>S</td>
<td>Sinking Fund Payment</td>
<td>226</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>227</td>
</tr>
<tr>
<td>SA</td>
<td>Sale of Asset</td>
<td>14</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>T</td>
<td>Issue Tendered</td>
<td>141</td>
<td>18</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>160</td>
</tr>
<tr>
<td>X</td>
<td>Issue Exchanged</td>
<td>901</td>
<td>63</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td>975</td>
</tr>
</tbody>
</table>

Table 1.11: Frequency of Events in FISD: This table presents frequency of issues for each firm-year observation, for the sample of non-financial firms which have a convertible bond outstanding.
<table>
<thead>
<tr>
<th>Industry</th>
<th>Converted</th>
<th>Total Issued</th>
<th>Conversion Prob. (%)</th>
<th>Median Leverage</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemicals &amp; Allied Products</td>
<td>34</td>
<td>781</td>
<td>4.35</td>
<td>0.21</td>
<td>1.28</td>
</tr>
<tr>
<td>Electric, Gas &amp; Sanitary Services</td>
<td>10</td>
<td>548</td>
<td>1.82</td>
<td>0.39</td>
<td>0.52</td>
</tr>
<tr>
<td>Business Schools</td>
<td>19</td>
<td>523</td>
<td>3.63</td>
<td>0.15</td>
<td>1.36</td>
</tr>
<tr>
<td>Oil and Gas Extraction</td>
<td>26</td>
<td>418</td>
<td>6.22</td>
<td>0.28</td>
<td>0.66</td>
</tr>
<tr>
<td>Industrial &amp; Commercial Machinery</td>
<td>7</td>
<td>235</td>
<td>2.98</td>
<td>0.19</td>
<td>1.06</td>
</tr>
<tr>
<td>Electronic and Electrical Equipment</td>
<td>9</td>
<td>206</td>
<td>4.37</td>
<td>0.19</td>
<td>1.18</td>
</tr>
<tr>
<td>Communications</td>
<td>22</td>
<td>206</td>
<td>10.68</td>
<td>0.41</td>
<td>1.13</td>
</tr>
<tr>
<td>Health Services</td>
<td>5</td>
<td>206</td>
<td>2.43</td>
<td>0.33</td>
<td>1.29</td>
</tr>
<tr>
<td>Primary Metal Industries</td>
<td>8</td>
<td>169</td>
<td>4.73</td>
<td>0.24</td>
<td>0.96</td>
</tr>
<tr>
<td>Fishing, Hunting and Trapping</td>
<td>7</td>
<td>146</td>
<td>4.79</td>
<td>0.15</td>
<td>0.8</td>
</tr>
<tr>
<td>Eating &amp; Drinking Establishments</td>
<td>5</td>
<td>143</td>
<td>3.50</td>
<td>0.31</td>
<td>1.06</td>
</tr>
<tr>
<td>Measure, Analyze &amp; Control Instruments</td>
<td>4</td>
<td>122</td>
<td>3.28</td>
<td>0.19</td>
<td>1.13</td>
</tr>
<tr>
<td>Transportation Equipment</td>
<td>3</td>
<td>104</td>
<td>2.88</td>
<td>0.30</td>
<td>0.9</td>
</tr>
<tr>
<td>Air Transportation</td>
<td>2</td>
<td>104</td>
<td>1.92</td>
<td>0.39</td>
<td>1.44</td>
</tr>
<tr>
<td>Apparel &amp; Accessory Stores</td>
<td>3</td>
<td>97</td>
<td>3.09</td>
<td>0.17</td>
<td>1.11</td>
</tr>
<tr>
<td>Printing &amp; Publishing</td>
<td>2</td>
<td>92</td>
<td>2.17</td>
<td>0.23</td>
<td>0.92</td>
</tr>
<tr>
<td>Miscellaneous Retail</td>
<td>1</td>
<td>92</td>
<td>1.09</td>
<td>0.28</td>
<td>0.99</td>
</tr>
<tr>
<td>Coal Mining</td>
<td>5</td>
<td>77</td>
<td>6.49</td>
<td>0.33</td>
<td>0.93</td>
</tr>
<tr>
<td>Agricultural - Crops</td>
<td>0</td>
<td>73</td>
<td>0.00</td>
<td>0.33</td>
<td>0.61</td>
</tr>
<tr>
<td>Home Furniture, Furnishings &amp; Equip Stores</td>
<td>1</td>
<td>63</td>
<td>1.59</td>
<td>0.22</td>
<td>1.06</td>
</tr>
<tr>
<td>Rail road Transportation</td>
<td>2</td>
<td>58</td>
<td>3.45</td>
<td>0.30</td>
<td>1.1</td>
</tr>
<tr>
<td>Engineering, Accounting, Research Services</td>
<td>2</td>
<td>54</td>
<td>3.70</td>
<td>0.13</td>
<td>1.28</td>
</tr>
<tr>
<td>Fabricated Metal Products</td>
<td>1</td>
<td>52</td>
<td>1.92</td>
<td>0.31</td>
<td>0.76</td>
</tr>
<tr>
<td>General Merchandise</td>
<td>1</td>
<td>51</td>
<td>1.96</td>
<td>0.29</td>
<td>0.95</td>
</tr>
<tr>
<td>Stone, Clay, Glass &amp; Concrete</td>
<td>1</td>
<td>49</td>
<td>2.04</td>
<td>0.38</td>
<td>0.67</td>
</tr>
<tr>
<td>Building Contractors</td>
<td>1</td>
<td>44</td>
<td>2.27</td>
<td>0.45</td>
<td>1.14</td>
</tr>
<tr>
<td>Wholesale trade-Non-durable goods</td>
<td>3</td>
<td>40</td>
<td>7.50</td>
<td>0.28</td>
<td>0.98</td>
</tr>
<tr>
<td>Construction- Special Trade</td>
<td>1</td>
<td>33</td>
<td>3.03</td>
<td>0.23</td>
<td>0.38</td>
</tr>
<tr>
<td>Water Transportation</td>
<td>0</td>
<td>29</td>
<td>0.80</td>
<td>0.46</td>
<td>1.07</td>
</tr>
<tr>
<td>Heavy Construction</td>
<td>2</td>
<td>26</td>
<td>7.60</td>
<td>0.22</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Total 187 | 4,853 | 3.85 | 0.28 | 0.98

Table 1.12: Frequency of Issuance and Conversion for Industries: This table presents frequency of issuance and conversion of convertible bonds in FISD as well as average leverage and beta for all industries between 1995 and 2010. Industry classification is based on two-digit SIC codes. Industries are sorted based on Total Issued column, which indicates total number of convertible bonds for each industry between 1995 and 2010.
2.1 Introduction

In their pioneering theoretical model of collusion under demand uncertainty, Rotemberg and Saloner (1986) show that when demand is independently and identically distributed over time, and firms observe demand before taking actions, collusion is harder to support when demand is higher. The intuition is compelling: the period incentive to cheat on the cartel rises with demand, but expected continuation payoffs are unchanged. While this setting is stark, Kandori (1991) establishes that the essence of the Rotemberg and Saloner result extends to serially-correlated demand—effectively, the period incentive to cheat on the cartel is more sensitive to current market conditions than are continuation payoffs. Haltiwanger and Harrington (1991) establish related results for deterministic cyclical demand—collusion is hardest to support at the peak of a cycle. Bagwell and Staiger (1997) generalize these results melding a Markov demand growth process on top of the i.i.d. transitory shocks.

In contrast to this robust theoretical prediction, we have Scherer (1980)’s summary of his empirical work: “Yet it is precisely when business conditions turn sour that price cutting runs most rampant among oligopolistic with high fixed costs,” and Aiginger et al. (1998)’s survey of 113 experts found that most believed that price wars are more likely when demand is low. Staiger and Wolak (1992) assert “the conventional empirical wisdom [is] that tacit collusion tends to break down when business conditions turn sour.” Empirical studies providing support for the premise that collusion is harder to support in downturns include Porter (1985), Scherer and Ross (1990), Suslow (2005), and Ellison (1994). However, Domowitz et al. (1987), Chevalier and Scharfstein (1996) and Borenstein and Shepard (1996) provide empirical evidence consistent with cartels being more likely to breakdown in booms. This empirical research primarily focuses on prices and price-cost margins, and with output competition, procyclical price-cost margins are consistent with
collusion being more difficult to support when demand is high. Still, even among commodity cartels, there is evidence that collusion is difficult to support in downturns. For example, the two largest production wars in OPEC occurred in 1986 and 1997 when demand was low; in 1986, Saudi Arabia increased production from 2 to 5 mbpd when the price fell under $10.

One approach to trying to reconcile observations such as these is to argue that the central premise of Rotemberg and Saloner is wrong, and that firms can only imperfectly monitor collusion, for example, seeing equilibrium prices, but not demand realizations. Green and Porter (1984) and the vast imperfect monitoring literature that followed takes this approach. In this literature, low prices trigger punishment and price wars because cartel members cannot distinguish whether they are due to low demand or to cheating by a defecting cartel member.

We take a different approach. We return to the insights implicit in Scherer (1980), and explore collusion by risk-averse cartel members who face positive fixed operating costs. Our premise that cartel members are risk averse captures the observation that many commodity cartels consist of “country cartels” that do not care about profits per se, but rather about the utility their citizens derive from those profits. For example, OPEC is a cartel of oil-producing countries; and similar country cartels have existed in many manufacturing and commodity cartels (natural resources such as minerals, chemicals, raw materials, metals, etc., including sugar, rayon, zinc, aluminum, copper, coffee, tea, and so on; see Suslow (2005) for an extensive list.).

Because most commodity cartels choose output levels, we model output competition rather than price competition when defections from cartel quotas result in reversion to static Nash equilibrium output levels, but we otherwise focus on the classical i.i.d. demand, constant marginal cost environment studied by Rotemberg and Saloner. We measure the extent of collusion as captured by the ratio of output relative to monopoly levels that can be supported in different demand states.

When cartel members are risk averse (with power utility), they value a marginal dollar of profit by more when profits are lower. One might therefore conjecture that risk aversion alone could be enough to make collusion more difficult to support in low demand states. This conjecture is false: to overturn the Rotemberg-Saloner result that higher demand always makes it more difficult for firms to collusively restrain output, cartel members must not only be sufficiently risk averse, but fixed operating costs of production must be sufficiently

---

1Among non-country cartels, one might argue either that cartel members inherit the risk aversion of managers, or that financially-constrained firms with large debt levels, in effect, are risk averse, valuing a marginal dollar in bad times by more because it may stave off liquidation. Consistent with this, Busse (2002) provides evidence that price wars in airline industry are unilaterally initiated by financially troubled firms.
high. Fixed operating costs magnify the marginal utility derived from an additional dollar of revenues in bad times, sharply raising the period incentive to cheat on the cartel when demand is especially low. More provocatively, we establish that collusion is easiest to support when demand is intermediate, neither too low, nor too high. The fixed costs of production mean that preferences effectively exhibit decreasing relative risk-aversion. Thus, when demand is especially low, the very high marginal valuation of an additional dollar of profit induced by the fixed operating costs make the incentive to cheat on the cartel very high; and when demand is much higher, the decreasing risk aversion implies that the classical effect dominates—as demand increases, there are more dollars to be gained from cheating on the cartel. We further establish that although the incentive to cheat on the cartel is a U-shaped function of the level of demand, the incentive rises more sharply in low demand states than in higher ones. This U-shaped predicted relationship between demand and the ability to collusively support restrain output levels suggests that efforts to tease out the true relationship in the data must be more nuanced; and that this relationship may account for the somewhat mixed empirical findings.

We then establish that greater fixed operating costs or risk aversion make it harder to support collusion when demand is low, but easier to support collusion when demand is high. Greater fixed costs or risk aversion raising the net continuation payoff from collusion by enhancing the threatened Nash reversion punishment for cheating on the cartel. However, greater fixed costs or risk aversion also raise the potential period utility gains from cheating on the cartel. The impact of higher operating costs on period incentives dominates when demand is especially low, making collusion more difficult to support; but the higher net continuation payoffs dominate when demand is especially high, making collusion easier to support.

We next present the model and analysis. A conclusion follows. Proofs are in an appendix.

2.2 The Model

Our framework features an infinitely repeated output game played by two agents 1 and 2 (e.g., country members of OPEC) that sell a homogeneous good in a market where demand evolves stochastically according to an i.i.d. process.\(^2\) Date \(t\) demand is given by

\[
P_t = \theta_t - Q_t, \tag{2.1}
\]

\(^2\)Extensions to \(N > 2\) agents are routine.
where $\theta_t$ is identically and independently distributed, with associated distribution function $F(\theta)$ on its positive support $[a, b]$, with $a > 0$, and $Q_t = q^t_1 + q^t_2$ is aggregate output. Without loss of generality, we normalize the constant marginal costs of production to zero. The agents also incur fixed operating costs each period of $c \geq 0$, where $c \leq \frac{a^2}{9}$.

In the classical Rotemberg and Saloner (1986) framework, the agents are risk-neutral firms whose period payoffs equal their period profits, $U^i(\pi^i_t) = \pi^i_t$. Thus, the marginal value that a firm derives from a dollar of profit does not vary with the level of profit, and fixed operating costs are irrelevant for a firm’s decision making (assuming that exit is not a strategic consideration). We depart from this classical setting to investigate collusion by risk-averse agents that face positive fixed operating costs. Agent $i$ derives period utility from profit $\pi^i_t$ of $U^i(\pi^i_t) = (\pi^i_t)^\alpha$, where $0 < \alpha \leq 1$. For simplicity, we assume that each period, a cartel member consumes its period profits, i.e., there is no saving and borrowing. As a result, a cartel member values an extra dollar of profit by more when profits are lower. Moreover, fixed operating costs enter decision making non-trivially, as they especially magnify the marginal value of an additional dollar in bad times. Firms use a common discount factor $\beta \in (0, 1)$ to discount future payoffs.

We focus on the maximal period collusion profits that can be supported by threats to revert to the non-cooperative static Nash equilibrium outputs forever if a cartel member ever deviates from their collusive agreement. We do this because we want agents to be able to provide their citizens positive consumption in all states of the world. Our focus on output competition rather than price competition together with the assumption that $\frac{a^2}{9} > c$ ensures that the profits from Nash outputs always cover the period fixed operating costs, providing citizens subsistence consumption. Harsher threats are unlikely to be credible (for instance, failing to provide a minimal subsistence level might result in the state's overthrow). As important, output competition captures our real world motivating example of a country commodity cartel, in which cartel members clearly choosing levels of outputs rather than prices.

After observing period demand, the two cartel members simultaneously choose outputs. Define $q^{C}(\theta)$ to be the collusive firm output supported along the equilibrium path when the demand shock is $\theta$. Given that deviations from collusive outputs result in static Nash outputs in the future, an agent that cheats on the cartel agreement will produce the $q^{F}(\theta)$ that maximizes period profit, and hence period utility, solving

$$
\max_{q'(\theta)} (\theta - q^{C}(\theta) - q'(\theta))q'(\theta) - c \Rightarrow q^{F}(\theta) = (\theta - q^{C}(\theta))/2.
$$
Let $\pi_C(\theta) = (\theta - 2q_C(\theta))q_C(\theta) - c$ and $\pi_F(\theta) = (\theta - q(\theta))^2/4 - c$ denote the respective period profits from cooperating and cheating on the cartel, and let $q^C(\theta) = \theta/3$ be the Nash output and $\pi^F(\theta) = \theta^2/9 - c$ be the associated Nash period profit. Finally, let $U_C \equiv E[U(\pi_C(\theta))]$ be the expected period utility from cooperation along the equilibrium path, and let $U^P \equiv E[U(\pi^F(\theta))]$ be the expected period utility along the punishment path. Then, for each given demand shock $\theta$, incentive compatibility requires

$$U(\pi_C(\theta)) + \left(\frac{\beta}{1 - \beta}\right) U^C \geq U(\pi^F(\theta)) + \left(\frac{\beta}{1 - \beta}\right) U^P. \tag{2.2}$$

Equation (2.2) can be re-arranged in terms of the "incentive to cheat":

$$U(\pi^F(\theta)) - U(\pi_C(\theta)) \leq \left(\frac{\beta}{1 - \beta}\right) (U^C - U^P) \equiv v. \tag{2.3}$$

That is, for a cartel production schedule to be incentive compatible, the net period utility payoff from cheating when demand is $\theta$, $U(\pi^F(\theta)) - U(\pi_C(\theta))$, cannot exceed the net expected payoff from future cooperation rather than punishment, $v$.

**Cartel’s Problem.** The cartel’s objective is to find the incentive compatible production schedule that maximizes their joint utility along the equilibrium path, $\sum_{t=1}^{\infty} \beta^{t-1} E[U_1(\theta) + U_j(\theta)]$. With power utility, we can write the cartel’s problem as

$$\max_{\eta(\theta)} \int_a^b (q(\theta)(\theta - 2q(\theta)) - c) dF(\theta) \quad \tag{2.4}$$

s.t. \quad \left(\frac{\theta - q(\theta))^2}{4} - c\right)^\alpha - (2q(\theta)q(\theta) - c)^\alpha \leq \left(\frac{\beta}{1 - \beta}\right) (U^C - U^P) \equiv v, \quad \forall \theta \in [a,b].$$

We measure the cartel’s ability to support collusion when demand is $\theta$ by the ratio $q^C(\theta)/q^m(\theta) \geq 1$, i.e., by the ratio of output relative to monopoly levels. A higher ratio indicates that collusion is more difficult to support in that demand state. It is important to note that most empirical researchers measure collusion in price-cost margins (which with constant marginal costs reduces to measuring collusion in prices). Obviously, if price-cost margins fall with $\theta$, then $q^C(\theta)/q^m(\theta)$ also rises with $\theta$. However, with output competition, $q^C(\theta)/q^m(\theta)$ can rise with $\theta$, indicating a reduced ability of the cartel to support collusion in higher demand states, even though price-cost margins rise uniformly with $\theta$. Phrased differently, with output competition,
the procyclical price-cost margins found empirically do not imply that collusion is easier to support in high demand states.

For the special case of linear utility, $U(\pi_i(\theta_t)) = \pi_i(\theta_t)$, the cartel’s objective reduces to the output-competition variant of Rotemberg and Saloner (1986). In that setting, it immediately follows that the incentive to cheat increases in $\theta$, as with i.i.d. demand, expected continuation payoffs do not vary with $\theta$, but the current payoffs from cheating on the cartel rise when the stakes are higher. As a result, $q^C(\theta)/q^m(\theta)$ is constant when demand is low enough that monopoly profits can be supported, and is strictly increasing when demand is high enough that threats to deviate to Nash outputs cannot support monopoly profits.

One might conjecture that risk-aversion alone, i.e., $\alpha < 1$, would be enough to reverse the result that it is more difficult for the cartel to collusively restrict output toward monopoly levels when demand is high, i.e., to reverse the result that $q^C(\theta)/q^m(\theta)$ rises with $\theta$. That is, one might conjecture that since the marginal utility derived from another dollar of profit is higher when profits are lower, collusion might be more difficult to support when demand is low and cartel members are sufficiently risk averse. This conjecture is false. The following proposition establishes necessary conditions for it to be harder to support collusion when demand is low than when it is high: not only must agents be risk averse, $\alpha < 1$, but they must also have positive fixed costs of operation, $c > 0$.

**Proposition 1** Suppose that either $c = 0$ or $\alpha = 1$. Then over-production relative to monopoly levels rises with the level of demand, i.e., $q^C(\theta)/q^m(\theta)$ is non-decreasing in $\theta$.

Thus, both non-zero fixed operating costs and risk-aversion are necessary for overproduction not to rise with $\theta$. Intuition for why more than risk aversion is required can be gleaned from looking at those demand states $\theta$ where the net value of future cooperation $v$ is high enough that the IC constraint is slack. For such $\theta$, the cartel’s optimization problem simplifies to a point-wise maximization of the objective. The associated first-order condition is

$$(q(\theta)(\theta - 2q(\theta)) - c)^{\alpha-1} (\theta - 4q) = 0,$$

with solution $q^C(\theta) = \theta/4$. The two agents jointly produce the monopoly output, $\theta/2$, and each earns half of the monopoly profits net of operating costs, $\theta^2/8 - c$; and the associated fink output is $3\theta/8$, which delivers profits of $\pi^F(\theta) = 9\theta^2/64 - c$. To see how the incentives to cheat on the cartel hinge on the level of demand,
the extent of risk aversion and the fixed operating costs, define the (period) incentive to cheat on monopoly output as

\[ f(\theta; \alpha, c) = U(\pi^F(\theta)) - U(\pi^C(\theta)) = \left( \frac{9\theta^2}{64} - c \right)^\alpha - \left( \frac{\theta^2}{8} - c \right)^\alpha. \tag{2.5} \]

When there are no fixed costs, \( f(\theta, \alpha, c = 0) \) simplifies to

\[ f(\theta; \alpha, c = 0) = \left( \frac{9\theta^2}{64} \right)^\alpha - \left( \frac{\theta^2}{8} \right)^\alpha = \left( \frac{\theta^2}{64} \right)^\alpha (9^\alpha - 8^\alpha), \]

which rises with \( \theta \). In essence, the extent of risk aversion scales the period incentive to cheat on the cartel, preserving monotonicity in \( \theta \). A similar result holds when monopoly output cannot be supported. Relatedly, when \( \alpha = 1 \) the incentive to cheat is a quadratic increasing function of \( \theta \).

We now show that for the incentive to cheat on the cartel not to rise with the level of demand, the impact of risk aversion must be higher in low demand states than high, i.e., preferences induced by the fixed operating costs must exhibit decreasing relative risk aversion. The fixed costs of production mean that preferences effectively take a subsistence utility form, and the associated Arrow-Pratt measure of relative risk-aversion is \( RRA(W) = -WU''(W)/U'(W) = (1 - \alpha) \frac{W}{W-c} \), which decreases in \( W \) if and only if both \( c > 0 \) and \( \alpha < 1 \). Then, when demand is low, the very high marginal valuation of an additional dollar of profit induced by the fixed operating costs cause the incentives to cheat on the cartel to rise further when demand drops lower, and agents become even more desperate for another marginal dollar of profits. In contrast, when demand is much higher, the decreasing risk aversion implies that risk aversion matters less, with the result that the standard effect dominates—as demand increases, there are more dollars to be gained from cheating on the cartel. Putting these two observations together suggests that the incentive to cheat on the cartel will be a U-shaped function of the level of demand, \( \theta \). We now formalize this intuition.

We also begin to answer the question of exactly where the separation between good and bad times occurs. The theorem shows that to deliver the U-shaped relationship, agents must have intermediate levels of risk aversion: for the incentive to cheat on the cartel not to rise monotonically with \( \theta \), agents must be sufficiently risk averse; and for the incentive not to fall monotonically with \( \theta \), they must not be too risk averse.

**Theorem 1** There exist critical levels of risk aversion, \( \alpha(c) \) and \( \bar{\alpha}(c) \), indexed by the fixed costs \( c \), such that if and only if cartel members have intermediate levels of risk aversion, \( \alpha(c) < \alpha < \bar{\alpha}(c) < 1 \), then \( f \) is a
U-shaped function of $\theta$, achieving a minimum at $\hat{\theta}(\alpha, c) \in (a, b)$. That is, $f'(\theta)<0$ for $\theta < \hat{\theta}(\alpha, c)$, and $f'(\theta)>0$ for $\theta > \hat{\theta}(\alpha, c)$. Further, $\alpha(c), \bar{\alpha}(c)$ and $\hat{\theta}(\alpha, c)$ rise with the fixed cost $c$, and $\hat{\theta}(\alpha, c)$ increases in risk aversion $\alpha$.

The proof shows that there is a unique intermediate demand level $\hat{\theta}$ that minimizes the incentive to cheat. As demand falls below $\hat{\theta}$, the incentive to cheat rises due to the high marginal valuation of another dollar of profit; and as demand rises above $\hat{\theta}$, so too does the incentive to cheat due to the greater profit that can be gained. Further, the comparative statics reveal that when agents are more risk averse or fixed costs are greater, demand does not have to be as bad for the incentive to cheat on the cartel to begin to rise as demand drops lower.

Monopoly outputs are supportable when the period benefit from cheating, $f$, is less than the expected net value of future cooperation, $v$, which is independent of $\theta$. When $f$ is a U-shaped function of $\theta$, it directly follows that monopoly outputs can only be sustained for intermediate values of demand whenever cartel members are neither so patient that they can support monopoly outputs in every state, nor so impatient that they can support monopoly outputs in no state (see Figure 2.1). Corollary 1 formalizes the necessary conditions.

**Corollary 1** There exist $\beta$, $\bar{\beta}$ with $\beta < \bar{\beta}$ such that if and only if $\beta \in [\beta, \bar{\beta}]$ the cartel can support monopoly profits only if demand is neither too low nor too high: If and only if $\beta \in [\beta, \bar{\beta}]$, there exist $\theta(\beta), \bar{\theta}(\beta)$ with $a < \theta(\beta) < \bar{\theta}(\beta) < b$ such that monopoly profits can be supported if and only if $\theta \in [\theta(\beta), \bar{\theta}(\beta)]$.

We have shown that there are two forces driving cartel away from supporting monopoly outputs: temptations rooting from the larger potential profit gain when times are good, and desperateness for an extra dollar of profits when times are bad. But, which force is stronger? Proposition 2 shows that the ability to support collusion drops off more quickly when demand falls below $\hat{\theta}$ than when it rises past $\hat{\theta}$, i.e., the left branch of $f$ is steeper than the right branch.

**Proposition 2** Consider any $\theta_1 < \theta_2$ such that $f(\theta_1) = f(\theta_2)$. Then $|f'(\theta_1)| > |f'(\theta_2)|$.

The intuition for Proposition 2 devolves from the increasing desperation implicit in Scherer (1980)'s summary that “Yet it is precisely when business conditions turn sour that price cutting runs most rampant among oligopolistic with high fixed costs.” Proposition 2 goes beyond Theorem 1. Theorem 1 showed that the
Figure 2.1: Incentive to Cheat Function Inside Monopoly Support Region: \( f(\theta; \alpha, c) \) as a function of \( \theta \). Parameters are: \( \alpha = 1/3 \) and \( c = 1/9 \).

incentive to cheat on monopoly output rises not only when demand is larger, but also when market conditions turn sour. Proposition 2 documents an asymmetry in the slopes of the incentive to cheat function, which results in a narrower range for monopoly support in bad times than in good times. Put differently, \( \hat{\theta} \) is always closer to \( \theta \) than to \( \bar{\theta} \).

We now characterize output levels following demand realizations—both high and low—that are sufficiently extreme that the cartel cannot support monopoly outputs. To prevent agents from cheating, cartel output must be increased to a level that makes agents indifferent between cheating and cooperation. More formally, at each \( \theta \in [a, \hat{\theta}] \cup [\bar{\theta}, b] \) incentive compatible quotas, \( q(\theta) \), solve

\[
\left( \frac{(\theta - q(\theta))^2}{4} - c \right)^\alpha - ((\theta - 2q(\theta))q(\theta) - c)^\alpha = \left( \frac{\beta}{1-\beta} \right)(U^C - U^P) \equiv v. \tag{2.6}
\]

We show that the incentive to cheat outside the monopoly support region falls with the level of production relative to the monopoly production level. Define \( z \equiv q(\theta)/\theta \) to be the normalized production level: \( z \) is an index for overproduction relative to monopoly output, as \( 4z = q(\theta)/(\theta/4) = q(\theta)/q^M(\theta) \). When monopoly output can be supported, there is no overproduction, so that \( z = 1/4 \); and when the cartel breaks down and agents revert to Nash outputs, then \( z = 1/3 \). Therefore, outside the monopoly support region the feasible levels of collusion have \( 1/4 < z < 1/3 \), and profits decrease in \( z \).
Rewrite the left-hand side of the IC constraint (2.6) in terms of \( z \) and call it \( H(z, \theta) \):

\[
H(z, \theta) = \left( \theta^2 \frac{(1-z)^2}{4} - c \right)^\alpha - (\theta^2(1-2z)z - c)^\alpha.
\]  

(2.7)

When \( z = 1/4 \), then \( H(1/4, \theta; \alpha, c) \) reduces to the period incentive to cheat on monopoly output, \( f(\theta; \alpha, c) \).

As in Theorem 1, one can show that for every \( z \in (1/4, 1/3) \), the period incentive to cheat, \( H(z, \theta) \), is a U-shaped function of \( \theta \). Proposition 3 shows that when demand realizations make it more attractive to cheat on the cartel, agents must reduce this attraction by increasing output relative to the monopoly level, but that output increases become less and less effective at reducing this incentive. Further, collusion is harder to sustain both for more extreme low demand realizations and for more extreme high demand realizations, requiring greater overproduction:

**Proposition 3** Outside the monopoly support region \([\bar{\theta}, \bar{\theta}]\), the period incentive to cheat is a continuous decreasing, convex function of output relative to monopoly levels: \( \frac{\partial H(z, \theta)}{\partial z} < 0 \), and \( \frac{\partial^2 H(z, \theta)}{\partial z^2} > 0 \). Further, overproduction relative to monopoly output increases when demand is further from the monopoly support region: \( \frac{\partial q(\theta)}{\partial (\theta - \theta')} > 0 \) for \( \theta > \theta' \), and \( \frac{\partial q(\theta)}{\partial (\theta - \bar{\theta})} > 0 \) for \( \theta > \bar{\theta} \).

One might conjecture that when the fixed costs of production, \( c \), are higher, or cartel members are more risk averse (lower \( \alpha \)), it becomes more difficult to support collusion in every demand state. The intuition underlying such a conjecture is that such changes raise the period utility gain from cheating on any given level of output. This conjecture is false. In essence, while the intuition is correct, it is also incomplete. The conjecture that greater fixed costs or increased risk aversion make collusion harder to support would follow directly if the net continuation payoffs from collusion versus punishment did not increase. However, as \( c \) is increased (or as agents become more risk averse), the threat to punish cheating on the cartel by reverting to Nash equilibrium outputs becomes harsher relative to the gain from a given level of cooperation. That is, greater fixed costs or risk aversion can raise the net continuation payoff from collusion rather than punishment. If, indeed, \( v \) rises by enough with greater operating costs or risk aversion to offset the increased period incentive to cheat on the cartel, then greater collusion may be facilitated.

A maintained assumption in the two propositions below is that increases in operating costs \( c \) or in risk aversion (reductions in \( \alpha \)) do not uniformly raise or lower the incentive to cheat on the cartel. That is, when operating costs or risk aversion is higher then collusion is easier to support in some demand states, and harder in others. We then establish a single-crossing property characterizing which states collusion is easier
to support. An exhaustive numerical analysis confirms that this maintained assumption holds whenever
demand is uniformly distributed (and the monopoly support region is interior). Indeed, we have been unable
to construct a counter-example.

**Proposition 4** Consider $\alpha_2 < \alpha_1$. Then outside the monopoly support region, more risk averse agents find
it harder to support collusion in bad times, but easier in good times: For $\alpha_2 < \alpha_1$, there exists a $\theta^*$ such that
for all $\theta \leq \theta^*$, if $z_2(\theta) > 1/4$, then $z_2(\theta) > z_1(\theta)$; and for all $\theta > \theta^*$, if $z_1(\theta) > 1/4$ then $z_2(\theta) < z_1(\theta)$.

**Proposition 5** Consider two levels of fixed costs of production, $c_2 > c_1$. Then outside the monopoly support
region, greater fixed costs make it harder to support collusion in bad times, but easier in good times: For $c_2 > c_1$, there exists a $\theta^*$ such that for all $\theta \leq \theta^*$, if $z_2(\theta) > 1/4$, then $z_2(\theta) > z_1(\theta)$; and for all $\theta > \theta^*$, if $z_1(\theta) > 1/4$ then $z_2(\theta) < z_1(\theta)$.

The key to these proofs is to show that the impact of an increase in $c$ or in risk aversion on the period
gain from cheating, $H(z, \theta)$ falls with $\theta$ for a fixed $z = q(\theta)/\theta$, i.e., that
\[
\frac{\partial^2 H(z, \theta; c)}{\partial c \partial \theta} < 0.
\]
Hence, if we ever have $H_1(z^*, \theta^*) = H_2(z^*, \theta^*)$, then it is unique.

Numerically, we find that whenever demand is uniformly distributed, and agents are sufficiently risk
averse with high enough operating costs that the monopoly support region is interior, then continuation payoffs always rise with c or with risk aversion by amounts that, consistent with these figures and the two propositions, give rise to asymmetric effects on the cartel’s ability to support collusion. That is, we find that the effect of an increase in c or reduction in α on the increased incentive to cheat dominates the impact on net continuation payoffs for sufficiently low demand shocks where agents are especially desperate for another dollar of revenue. However, net continuation payoffs rise with increased operating costs and increased risk aversion, and this effect dominates when demand is sufficiently high, making collusion easier to sustain. These results reflect the induced decreasing relative risk aversion in preferences—the effect of an increase in operating costs or risk aversion on the period utility gain from cheating on a given level of collusion falls as demand, and hence profits, rise.

Asymmetric cartels. Although we do not analyze it formally, Propositions 4 and 5 have suggestive implications for how heterogeneous agents with different levels of fixed operating costs or risk aversion should collude. For example, Saudi Arabia likely has lower fixed operating costs than most other OPEC cartel members. Then the propositions would suggest that Saudi Arabia should have a lower share of output in low demand states (where high operating cost cartel members find collusion more difficult to sustain for a given output share), but a higher share of output when demand is high (and high operating cost cartel members mind ceding share by less, and are willing to do so to in order to obtain a greater share in low demand states).
demand states where they care more about their share). It follows that the low operating cost or less risk averse cartel member’s output should be more sensitive to the level of demand than is the output of higher operating cost or more risk averse cartel members. Thus, Saudi Arabia should be the swing producer, with relatively much lower outputs in bad times, and relatively much higher outputs in good times, so that it would appear as if its output is the primary driver determining cartel outcomes.

2.3 Conclusion

A robust prediction of the theoretical literature on collusion under demand uncertainty when cartel members observe demand and can monitor each other’s actions is that collusion is more difficult when demand is higher. In contrast to this theoretical prediction, most empirical researchers have concluded that price wars are more common when demand is low.

We provide a simple theory of collusion by risk averse agents that face positive fixed operating costs that can reconcile these two literatures by providing conditions under which it is most difficult for cartel members to collusively restrict output when demand is especially low, but that it also becomes difficult to support collusion when demand is high. The idea that cartel members are risk averse captures the observation that many effective cartels are comprised of countries that collusively restrict output of various commodities. Such cartels do not care directly about profits, per se, but rather care about the utility derived by the risk averse citizens who receive those profits. Therefore, we model citizen’s preferences using a power utility function who consume the resulting profits. As a result, the marginal value of a dollar of profit is greater when demand, and hence profits are lower; and this high marginal valuation is further enhanced by the large fixed operating costs that Scherer (1980) cites as playing a vital role in making collusion difficult.

We show that for aggregate cartel output relative to monopoly levels to be a U-shaped function of the level of demand, both ingredients are necessary—cartel members must be risk averse, and operating costs must be positive. We further establish that when cartel members are more risk averse or fixed operating costs are higher, then it becomes more difficult to support collusion in bad demand states, but easier in booms.
2.4 Appendix

**Proof of Proposition 1**: When $\alpha = 1$, it is immediate from Rotemberg and Saloner (1986) that $q(\theta)/q^m(\theta)$ is non-decreasing in $\theta$.

If $c = 0$, and the IC constraint does not bind, then $q^C(\theta)/q^m(\theta) = 1$. Now suppose that the IC constraint binds, and let $\theta_1 < \theta_2$ be two arbitrary values of $\theta$ outside the monopoly support region. Since $q^m(\theta) = \theta/4$, to show that $q(\theta)/q^m(\theta)$ increases in $\theta$ we must show that $q(\theta_2)/\theta_2 > q(\theta_1)/\theta_1$, where $q(\theta_i)/\theta_i \equiv z_i \in (1/4, 1/3)$. To prove that $z_2 > z_1$, suppose instead that $z_1 \geq z_2$. Rewrite the IC constraint in terms of $z_i$ when $c = 0$ as:

$$
\left(\frac{\theta_1^2 (1 - z_1)^2}{4}\right)^\alpha - \left(\theta_1^2 (1 - 2z_1)z_1\right)^\alpha = v.
$$

Since $v$ is independent of $\theta$,

$$
\left(\frac{\theta_1^2 (1 - z_1)^2}{4}\right)^\alpha - \left(\theta_1^2 (1 - 2z_1)z_1\right)^\alpha = \left(\frac{\theta_2^2 (1 - z_2)^2}{4}\right)^\alpha - \left(\theta_2^2 (1 - 2z_2)z_2\right)^\alpha.
$$

Since $\theta_1 < \theta_2$, it follows that

$$
\left(\frac{(1 - z_1)^2}{4}\right)^\alpha - \left((1 - 2z_1)z_1\right)^\alpha > \left(\frac{(1 - z_2)^2}{4}\right)^\alpha - \left((1 - 2z_2)z_2\right)^\alpha.
$$

Calling the four terms in this inequality from left to right as $A, B, C$ and $D$, rewrite the inequality as:

$$
A - B > C - D.\text{ Under the assumption } z_1 > z_2, \text{ and recalling that cooperation profits decrease in } z, \text{ i.e., } (1 - 2z)z \text{ decreases in } z > 1/4, \text{ we have } B/D < 1. \text{ Therefore, } A - B > C - D \text{ implies that } \frac{A - B}{B} > \frac{C - D}{D}. \text{ Therefore, } \frac{A}{B} > \frac{C}{D}, \text{ i.e., }
$$

$$
\left(\frac{(1 - z_1)^2}{4}\right)^\alpha > \left(\frac{(1 - z_2)^2}{4}\right)^\alpha,
$$

or equivalently, $\frac{(1 - z_1)^2}{(1 - 2z_1)z_1} > \frac{(1 - z_2)^2}{(1 - 2z_2)z_2}$ for $z_2 \geq z_1$. But $g(z) = \frac{(1 - z)^2}{(1 - 2z)z}$ is a decreasing function of $z$, i.e., $g'(z) < 0$, a contradiction. □
Proof of Theorem 1: The first-order condition is

\[ f'(\theta) = 2\alpha \theta \left( \frac{9}{64} (\frac{9\theta^2}{64} - c)^{\alpha-1} - \frac{1}{8} (\frac{\theta^2}{8} - c)^{\alpha-1} \right) = 0. \]

Solving yields

\[ \frac{\theta^2}{\frac{9\theta^2}{64}} - c = \left( \frac{8}{9} \right)^{\frac{1}{\alpha}} \Rightarrow \theta = 8 \sqrt{\frac{(1-k)c}{8-9k}}. \]

Notice that \( k = \left( \frac{8}{9} \right)^{\frac{1}{\alpha}} < \frac{8}{9} < 1. \) Clearly, \( \theta < \hat{\theta} \) implies that

\[ \frac{\theta^2}{\frac{9\theta^2}{64}} - c < \left( \frac{8}{9} \right)^{\frac{1}{\alpha}} \Rightarrow 9(\frac{9\theta^2}{64} - c)^{\alpha-1} < 8(\frac{\theta^2}{8} - c)^{\alpha-1}. \]

Therefore, \( f'(\theta) < 0. \) A similar argument holds for \( \theta > \hat{\theta}. \)

The requirement that \( a < \hat{\theta} < b \) imposes bounds on the range of \( \alpha. \) We require

\[ a < \hat{\theta} = 8 \sqrt{\frac{(1-k(\alpha)c}{8-9k(\alpha)}) < b. \]

Solving yields the upper and lower bounds:

\[ \bar{\alpha}(c) = 1 + \frac{\log(9/8)}{\log(\frac{8a^2}{9a^2-64c})} \quad \text{and} \quad \underline{\alpha}(c) = 1 + \frac{\log(9/8)}{\log(\frac{8a^2-64c}{9a^2-64c})}. \]

Since \( \alpha(x; c) = 1 + \log(9/8)/\log(\frac{8x^2-64c}{9x^2-64c}) \) is a decreasing function of \( x, \) with a limit of zero as \( x \) goes to infinity, \( \hat{\theta} \in (a, b) \) exists as long as \( \underline{\alpha}(c) < \alpha < \bar{\alpha}(c). \)

Finally, differentiating \( \hat{\alpha}(c) \) and \( \hat{\theta}(\alpha, c) \) with respect to \( c \) and \( \alpha \) delivers the results:

\[ \frac{\partial \hat{\alpha}(c)}{\partial c} = \frac{8a^2 \log(9/8)}{(9a^2-64c)(a^2-8c) \left( \log(1 - \frac{a^2}{9a^2-64c}) \right)^2} \geq 0, \]
which is non-negative since \( c < \frac{a^2}{b^9} \); and

\[
\frac{\partial \hat{\theta}(\alpha, c)}{\partial \alpha} = -\frac{2^{2n-1} 9^{1/\alpha} \log \left( \frac{9}{8} \right) c}{(\alpha - 1)^2 \left( 2^{\frac{1}{\alpha}} 9^{1/\alpha} - 9 \right)^2} < 0; \quad \frac{\partial \hat{\theta}(\alpha, c)}{\partial c} = \frac{4}{c} \sqrt{\frac{\left( \frac{9}{8} \right)^{1/\alpha} - 1}{8^{1/\alpha} 9^{1/\alpha} - 8}} > 0. \quad \square
\]

**Proof of Corollary 1:** Let \( \theta < \bar{\theta} \) be the two roots of \( f(\theta; \alpha, c) = v \) when it has two roots for \( a \leq \theta \leq b \). Note that \( v \) is independent of \( \theta \). Since \( f(\theta; \alpha, c) \) is a U-shaped function of \( \theta \) (Theorem 1), for intermediate values of \( \theta \) where \( f(\theta; \alpha, c) < v \) the IC constraint (2.3) is slack. Therefore, monopoly profits can be supported for \( \theta \in [\theta, \bar{\theta}] \).

For \( f(\theta; \alpha, c) = v \) to have two roots, \( v \) can be neither too small nor too large. Since \( v = \left( \frac{\beta}{1 - \beta} \right) (U^C - U^P) \), there is one-to-one mapping between \( v \) and \( \beta \). Thus, we must bound \( \beta \) appropriately: \( \beta \) must exceed the \( \tilde{\beta} \) that solves

\[
f(\tilde{\theta}; \alpha, c) = \left( \frac{\beta}{1 - \beta} \right) (U^C - U^P),
\]

and be less than the \( \bar{\beta} \) that solves

\[
\text{Min}\{f(a; \alpha, c), f(b; \alpha, c)\} = \left( \frac{\bar{\beta}}{1 - \bar{\beta}} \right) (U^C - U^P).
\]

Hence, \( f(\theta; \alpha, c) = v \) has two roots for \( \theta \in [a, b] \) if and only if \( \beta \in [\beta, \bar{\beta}] \). \( \square \)

**Proof of Proposition 2:** We must show that \( -f'(\theta_1) > f'(\theta_2) \), i.e.,

\[
-2\alpha_1 \left( \frac{9}{64} \left( \frac{9\theta^2}{64} - c \right)^{a-1} - \frac{1}{8} \left( \frac{\theta^2}{8} - c \right)^{a-1} \right) > 2\alpha_2 \left( \frac{9}{64} \left( \frac{9\theta^2}{64} - c \right)^{a-1} - \frac{1}{8} \left( \frac{\theta^2}{8} - c \right)^{a-1} \right),
\]

or equivalently that

\[
2\alpha_1 \left( -\frac{9}{64} \left( \frac{9\theta^2}{64} - c \right)^{a-1} + \frac{1}{8} \left( \frac{\theta^2}{8} - c \right)^{a-1} \right) + 2\alpha_2 \left( -\frac{9}{64} \left( \frac{9\theta^2}{64} - c \right)^{a-1} + \frac{1}{8} \left( \frac{\theta^2}{8} - c \right)^{a-1} \right) > 0.
\]

53
But follows directly from the concavity of the utility function, \( \alpha < 1: 9/64 > 1/8 \) implies we are subtracting marginal utilities evaluated at higher values, i.e., the marginal utility of cooperative payoffs exceed the marginal utility of their cheat payoff analogues. □

**Proof of Proposition 3:**

\[
\frac{\partial H(z, \theta)}{\partial z} = \frac{1}{2} \alpha \theta^2 \left( 2(4z - 1)(\theta^2(1 - 2z)z - c)^{\alpha - 1} - (1 - z)(\frac{\theta^2}{4}(1 - z)^2 - c)^{\alpha - 1} \right).
\]

To show \( \frac{\partial H(z, \theta)}{\partial z} \) is positive, equivalently we must prove:

\[
\frac{1 - z}{2(4z - 1)} > \left( \frac{\theta^2}{4}(1 - z)^2 - c \right)^{1 - \alpha} \frac{\theta^2 z(1 - 2z) - c}{\theta^2 z(1 - 2z) - c}.
\]

Since the cheat payoff, \((1 - z)^2/4\) always exceeds the cooperation payoff, \(z(1 - 2z)\), the right-hand side exceeds one. Therefore, it suffices to show that

\[
\frac{1 - z}{2(4z - 1)} > \frac{\theta^2}{4}(1 - z)^2 - c \frac{\theta^2 z(1 - 2z) - c}{\theta^2 z(1 - 2z) - c}.
\]

Define \( c' \equiv c/\theta^2 \) and rearrange the above inequality as

\[
\frac{1 - z}{2(4z - 1)} - \frac{\frac{1}{4}(1 - z)^2 - c'}{z(1 - 2z) - c'} = \frac{\frac{1}{2}(3z - 1)(z - 1 + 6c')}{2(4z - 1)(z(1 - 2z) - c')} = \frac{\frac{1}{6}(3z - 1)^2}{2(4z - 1)(z(1 - 2z) - c')} > 0,
\]

for \( 1/4 < z < 1/3 \). The last inequality follows since the above expression decreases in \( c' \) and thus is minimized when \( c' \) equals its upper bound of \( \text{Max}(c/\theta^2) = (a^2/9)/a^2 = 1/9 \), implying that the denominator is positive.

To prove convexity of \( H \), we bound the second derivative of \( H/\alpha \) strictly away from zero (we divide by \( \alpha \) because the derivative of \( H \) goes to zero as \( \alpha \) goes to zero). We also write \( H/\alpha \) in terms of \( c' = c/\theta^2 \in [0, 1/9] \) to make the domain compact:

\[
\frac{1}{\alpha} H(z; \alpha, c') = \frac{1}{\alpha} \left[ \left( \frac{(1 - z)^2}{4} - c' \right)^\alpha - ((1 - 2z)z - c')^\alpha \right],
\]

54
with associated second derivative

\[
\frac{1}{\alpha} \frac{d^2 H}{dz^2} = \frac{1}{4} (1 - \alpha) \left[ 4(4z - 1)^2 ((1 - 2z)z - c')^{\alpha-2} - (1 - z)^2 \left( \frac{1}{4} (1 - z)^2 - c' \right)^{\alpha-2} \right]^{\alpha-2} \\
+ \frac{1}{2} \left[ \left( \frac{1}{4} (1 - z)^2 - c' \right)^{\alpha-1} + 16 ((1 - 2z)z - c')^{\alpha-1} \right].
\] (2.8)

The compact domain has \( z \in [1/4, 1/3] \), \( \alpha \in [0, 1] \) and \( c' \in [0, 1/9] \). Further, \( \frac{1}{\alpha} \frac{d^2 H}{dz^2} \) is continuous and twice differentiable on its domain, with derivatives bounded from below, so that in an \( \epsilon \) ball around any point \((z, \alpha, c')\), \( \frac{1}{\alpha} \frac{d^2 H}{dz^2} \) cannot drop too far below its value at \((z, \alpha, c')\). Therefore, to establish convexity, it suffices to bound \( \frac{1}{\alpha} \frac{d^2 H}{dz^2} \) strictly away from zero on an appropriately fine grid. An exhaustive search on a grid with increments of 0.001 for \( z \), \( \alpha \) and \( c' \) reveals that it achieves a lower bound of 9/2 when \( \alpha = 1 \). See Figure 2.4.

Figure 2.4: 3-dimensional Plot of the Second Derivative of \( H/\alpha \): (a) the left panel shows \( H/\alpha \) as a function of \( \alpha \) and \( c \) at \( z = 1/4 \), and (b) the right panel shows \( H/\alpha \) as a function of \( \alpha \) and \( z \) at \( c = 0 \).

We now establish that over-production relative to monopoly increases in \( \theta - \bar{\theta} \) for \( \theta > \bar{\theta} \); and in \( \bar{\theta} - \theta \), for \( \theta < \bar{\theta} \). First consider any \( \theta_2 > \theta_1 \in (\bar{\theta}, b] \). To establish that \( q(\theta)/q^m(\theta) \) increases in \( (\theta - \bar{\theta}) \), we show that \( z_2 > z_1 \). Suppose instead that \( z_1 > z_2 \). We have:

\[
H(z_i, \theta_i) = \left( \theta_i^2 \frac{(1 - z_i)^2}{4} - c \right)^{\alpha} - \left( \theta_i^2 (1 - 2z_i)z_i - c \right)^{\alpha} = v \quad \text{for} \quad i = 1, 2.
\]

From Theorem 1 for \( \theta > \bar{\theta} \), \( f \) increases in \( \theta \), so

\[
f(\theta_2; \alpha, c) > f(\theta_1; \alpha, c) \Rightarrow H(1/4, \theta_2) > H(1/4, \theta_1).
\]

55
Also from incentive compatibility,

\[ H(z_2, \theta_2) = H(z_1, \theta_1) = v, \]

at the premised \( z_1 > z_2 \), and since \( H(z, \theta) \) is decreasing in \( z \) for any \( \theta \), this implies that \( H(z_2, \theta_1) > H(z_2, \theta_2) \). But if \( H(1/4, \theta_2) > H(1/4, \theta_1) \) and \( H(z_2, \theta_2) < H(z_2, \theta_1) \) then by the intermediate value theorem there exists a \( z' \) with \( 1/4 < z' < z_2 \) such that \( H(z', \theta_2) = H(z', \theta_1) \), a contradiction of \( \theta_2 > \theta_1 \) and \( z' < 1/3 \).

An identical proof by contradiction establishes that if \( \theta_1 < \theta_2 < \theta \), then \( z_1 > z_2 \). That is, \( z_1 < z_2 \) would imply \( H(z_2, \theta_2) = H(z_1, \theta_1) \) at \( z_1 < z_2 \) (by incentive compatibility), and hence \( H(z_1, \theta_2) > H(z_1, \theta_1) \), but here, \( H(1/4, \theta_1) > H(1/4, \theta_2) \) yields a contradiction via the intermediate value theorem. \( \square \)

**Proof of Proposition 4:** Let \( H_1(z, \theta) \equiv H(z, \theta; \alpha_1, c) \) and \( H_2(z, \theta) \equiv H(z, \theta; \alpha_2, c) \) for \( \alpha_1 > \alpha_2 \). We prove that if there exists a \( \theta \) such that \( z_1(\theta) = z_2(\theta) \), then it is unique. Call these values \( \theta^* \) and \( z^* \). To establish this single-crossing result, we prove that for a fixed \( z \), \( H_1 - H_2 \) increases in \( \theta \) by showing that \( \frac{\partial H}{\partial \theta} \) increases in \( \alpha \). Therefore, there exists a neighborhood of \( \theta^* \) and \( z^* \), such that for a fixed \( z \), \( H_1 \) is a steeper function of \( \theta \) than \( H_2 \).

When both IC constraints bind (i.e., \( H_i = v_i \) for \( i = 1, 2 \)) then \( H_1 - H_2 = v_1 - v_2 \) does not vary with \( \theta \), i.e., \( \frac{\partial H}{\partial \theta} = \frac{\partial v}{\partial \theta} = 0 \).

\[
H(z, \theta) = \left( \theta^2 \left(1 - \frac{z}{4}\right)^2 - c \right)^\alpha - \left( \theta^2 (1 - 2z)z - c \right)^\alpha = v(\alpha, c), \tag{2.9}
\]

\[
\frac{\partial H}{\partial \theta} = \frac{1}{2} \theta \alpha \left( 1 - z \right)^2 \left( \theta^2 \left(1 - \frac{z}{4}\right)^2 - c \right)^{\alpha-1} - 4z(1 - 2z) \left( \theta^2 (1 - 2z)z - c \right)^{\alpha-1} = 0.
\]

Defining \( \gamma_F \equiv 4z(1 - 2z) \left( \theta^2 (1 - 2z)z - c \right)^{\alpha-1} \) and \( \gamma_C \equiv 4z(1 - 2z) \left( \theta^2 (1 - 2z)z - c \right)^{\alpha-1} \), we must have \( \gamma_F = \gamma_C \equiv \gamma \). We now prove that \( \frac{\partial H}{\partial \theta} \) increases in \( \alpha \), i.e., \( \frac{\partial^2 H}{\partial \theta \partial \alpha} > 0 \):

\[
\frac{\partial^2 H}{\partial \theta \partial \alpha} = \frac{1}{2} \theta \left( 1 - z \right)^2 \left( \theta^2 \left(1 - \frac{z}{4}\right)^2 - c \right)^{\alpha-1} - 4z(1 - 2z) \left( \theta^2 (1 - 2z)z - c \right)^{\alpha-1}
\]

\[
+ \alpha \left( 1 - z \right)^2 \left( \theta^2 \left(1 - \frac{z}{4}\right)^2 - c \right)^{\alpha-1} \log \left( \theta^2 \left(1 - \frac{z}{4}\right)^2 - c \right) - 4z(1 - 2z) \left( \theta^2 (1 - 2z)z - c \right)^{\alpha-1} \log \left( \theta^2 (1 - 2z)z - c \right).
\]
Substituting $\gamma_F$ and $\gamma_C$, and using $\gamma_F = \gamma_C \equiv \gamma$, rewrite this as:

$$
\frac{\partial^2 H}{\partial \theta \partial c} = \frac{1}{2} \theta \left( \frac{\gamma_F - \gamma_C + \alpha \left( \gamma_F \log \left( \frac{\theta^2 (1-z)^2}{4} - c \right) - \gamma_C \log \left( \frac{\theta^2 (1-2z)z - c}{4} \right) \right) }{\partial \theta} \right) = \frac{1}{2} \theta \alpha \gamma \left( \log \left( \frac{\theta^2 (1-z)^2}{4} - c \right) - \log \left( \frac{\theta^2 (1-2z)z - c}{4} \right) \right) > 0.
$$

The inequality holds since $\frac{(1-z)^2}{4} > (1-2z)z$ for $z \in [1/4, 1/3]$. When monopoly output cannot be supported in both environments, then $\frac{\partial^2 H}{\partial \theta \partial c} > 0$, implies that for $\theta > \theta^*$, we need $z_1(\theta) > z_2(\theta)$ to retrieve $H_1 = v_1$ and $H_2 = v_2$; and $\theta < \theta^*$ demands $z_1(\theta) < z_2(\theta)$.

**Proof of Proposition 5:** Let $H_1(z, \theta) \equiv H(z; \theta; \alpha_1, c)$ and $H_2(z, \theta) \equiv H(z; \theta; \alpha_2, c)$ for $c_1 < c_2$. We prove that if there exists a $\theta$ such that $z_1(\theta) = z_2(\theta)$, then it is unique. Call these values $\theta^*$ and $z^*$. To establish this single-crossing result, we prove that for a fixed $z$, $H_1 - H_2$ increases in $\theta$ by showing that $\frac{\partial H}{\partial \theta}$ decreases in $c$. Therefore, there exists a neighborhood of $\theta^*$ and $z^*$, such that for a fixed $z$, $H_1$ is a steeper function of $\theta$ than $H_2$.

When both IC constraints bind then $\frac{\partial H}{\partial \theta} = \frac{\partial H}{\partial c} = 0$. We have

$$
\frac{\partial H}{\partial \theta} = \frac{1}{2} \alpha (1 - \alpha) \left( (1-z)^2 \left( \frac{\theta^2 (1-z)^2}{4} - c \right)^{\alpha-1} - 4z(1-2z) \left( \frac{\theta^2 (1-2z)z - c}{4} \right)^{\alpha-1} \right) = 0.
$$

We now prove that $\frac{\partial H}{\partial \theta}$ decreases in $c$, i.e., $\frac{\partial^2 H}{\partial \theta \partial c} < 0$:

$$
\frac{\partial^2 H}{\partial \theta \partial c} = \frac{1}{2} \alpha (1 - \alpha) \theta \left( (1-z)^2 \left( \frac{\theta^2 (1-z)^2}{4} - c \right)^{\alpha-2} - 4z(1-2z) \left( \frac{\theta^2 (1-2z)z - c}{4} \right)^{\alpha-2} \right) < 0.
$$

Substituting $\gamma_F$ and $\gamma_C$, and using $\gamma_F = \gamma_C \equiv \gamma$, rewrite this as:

$$
\frac{\partial^2 H}{\partial \theta \partial c} = \frac{1}{2} \alpha (1 - \alpha) \theta \left( \frac{\gamma_F}{\theta^2 (1-z)^2 - c} - \frac{\gamma_C}{\theta^2 (1-2z)z - c} \right) = \frac{1}{2} \alpha (1 - \alpha) \theta \gamma \left( \frac{1}{\theta^2 (1-z)^2 - c} - \frac{1}{\theta^2 (1-2z)z - c} \right) < 0.
$$

The inequality holds since $\frac{(1-z)^2}{4} > (1-2z)z$ for $z \in [1/4, 1/3]$, and hence its reciprocal is smaller. When monopoly output cannot be supported in both environments, then $\frac{\partial^2 H}{\partial \theta \partial c} > 0$, implies that for $\theta > \theta^*$ we need $z_1(\theta) > z_2(\theta)$ to retrieve $H_1 = v_1$ and $H_2 = v_2$, and $\theta < \theta^*$ demands $z_1(\theta) < z_2(\theta)$. □
Chapter 3

Portfolio Allocation for Public Pension Funds

3.1 Introduction

This paper examines the portfolio allocation policies of U.S. state and local government pension funds. It presents a dynamic model of public pension fund investment choice and analyzes how risk-taking behavior may vary with the pension plan’s characteristics. Risk is measured by the volatility of a fund’s asset portfolio rate of return relative to the rate of return on the market value of its liabilities. The model’s implications are examined using annual data on 125 state pension funds over the period from 2000 to 2009.

The asset allocation choice of a public pension fund is critical to understanding the problem of public pension plan under-funding. A public pension fund’s annual investment return is typically much larger in magnitude than its annual employer and employee contributions (Munnell and Soto (2008)). Furthermore, the fund’s portfolio allocation across broad asset classes is the major determinant of its investment return (Bodie (1990) and Brinson et al. (1991)). Thus, portfolio allocation policy has first-order consequences for funding status.

Public pension fund asset allocation also is of interest because, in aggregate, it has changed drastically over time. Figure 3.1 shows that state and local government pension funds invested almost entirely in cash and fixed income during the 1950s, but gradually increased their allocations to equities and, more recently, to other investments (including real estate, private equity, and hedge funds).

A benchmark policy for assessing a pension fund’s investment choice is a portfolio allocation that best hedges or “immunizes” the risk of its liabilities. The value of a pension fund’s liabilities equals the value of the retirement annuities that it is obligated to pay its employees and retirees. These retirement annuities typically are linked to a worker’s wages and years of service, and most often payments are partially indexed

---

1 An alternative version of this chapter is published as Pennacchi, G. and Rastad, M. Portfolio Allocation for Public Pension Funds, Journal of Pension Economics and Finance, April 2011
to inflation. Hence, the value of pension liabilities is exposed to risks from real or nominal interest rate changes and also changes in wage rates. A satisfactory analysis of portfolio choice must account for these risks.

Portfolio allocations that deviate from the benchmark portfolio which best immunizes liabilities introduce what we refer to as “tracking error.” This paper considers how different pension fund objectives influence the choice of tracking error volatility and how this volatility might be influenced by the pension plan’s funding ratio, past returns, and other plan characteristics. Our study is perhaps the first to examine the overall asset-liability risks of a time series – cross section sample of public pension funds. We use data on state and local government wages and various investment classes to compute a single measure of tracking error volatility for each pension fund during each year.

The plan of the paper is as follows. Section II briefly discusses related theoretical and empirical work on the portfolio allocations of public pension funds. Our model is presented in Section III. Section IV describes our data and variable construction, while Section V presents the empirical results. Concluding comments are in Section VI.
3.2 Related Literature on Public Pension Fund Portfolio Allocation

The focus of this paper is a public pension fund’s portfolio allocation relative to a benchmark portfolio that best hedges (immunizes) the fund’s liability risks. To evaluate liability risk, one must first determine a method for valuing pension liabilities since they are not marketable securities. There is disagreement on how this should be done, with the major conflict between the actuarial approach of the Government Accounting Standards Board (GASB) and the market value approach based on finance theory. The GASB actuarial approach discounts a pension plan’s future retirement payments using the expected rate of return on the pension plan’s assets, rather than a discount rate appropriate to the actual risk of the pension plan’s retirement payments.

As pointed out in many papers, most recently by Brown and Wilcox (2009), Lucas and Zeldes (2009), and Novy-Marx and Rauh (2009), valuation using the GASB actuarial standard is inconsistent with basic financial theory and leads to moral hazard incentives in the form of “accounting arbitrage”: a pension plan has the incentive to invest in assets with high systematic risk in order to justify a higher discount rate that will reduce the actuarial valuation of its liabilities. Thus, we take it as a settled question that a financial theory-based market valuation approach more accurately reflects the risk of pension liabilities.

A more subtle issue is whether the market value measure should be the pension fund’s Accumulated Benefit Obligation (ABO) or Projected Benefit Obligation (PBO). Typically, the annual annuity payment paid by a public pension plan to a participant is the product of the participant’s final average salary over the last one to five years, years of credited service, and a benefit multiplier of 1% to 2.5%. The ABO is the present value of these payments based on current years of service and the current average salary, whereas the PBO is the present value based on current years of service and an estimated future average salary just prior to retirement. Thus, for hedging liability risk when the public pension plan is likely to be a continuing concern, the PBO better incorporates future risks. Therefore, we employ it to value our benchmark immunizing portfolio.² The effect of using PBO is to include future wage uncertainty as a component of overall liability risk.

²Bodie (1990) argues that a corporate pension fund’s relevant obligation to be hedged is its ABO because its PBO is not guaranteed by the corporation or by the Pension Benefit Guaranty Corporation (PBGC) should the corporation fail. This reasoning is less relevant for public pension plans. Peng (2008) argues that public plan benefits are relatively more secure because, unlike corporate plans, municipalities typically cannot extinguish their obligations to pay pension benefits, even following bankruptcy. As a consequence, a public-sector worker who continues to be employed is likely to receive her PBO at retirement.
Black (1989) recommends that if a pension fund manager takes a narrow view by hedging the ABO measure of pension liabilities, the pension portfolio should invest almost exclusively in duration-matching bonds. If a broader PBO view is taken, then he recommends some allocation to stocks under the assumptions that stock returns are positively correlated with wage growth. Peskin (2001) supports this view and finds that a 20% to 90% allocation to equities could be optimal depending on the characteristics of a particular public pension fund.

Lucas and Zeldes (2009) come to a similar conclusion from a model where a municipality wishes to minimize tax distortions and pension liabilities are positively correlated with stocks. Their model predicts that pension funds should invest more in stocks if their liabilities are more wage-sensitive, which should be the case if a pension fund has a relatively high ratio of currently-employed pension participants to pension plan retirees. However, they find no empirical evidence for this prediction.

### 3.3 A Public Pension Fund Model

Since a public pension fund’s portfolio choice will derive from its objective function, we begin by considering possible normative and positive objectives of a public pension fund.

#### 3.3.1 The Public Pension Fund’s Investment Objective

As will be discussed, academics and practitioners often propose different investment objectives for a public pension plan. However, if one takes a broad Ricardian (Ricardo (1820))/Modigliani and Miller (1958b) perspective, a municipal pension fund’s objective may be irrelevant. As discussed in Bader and Gold (2007), arguments along the lines of Barro (1974) imply that any balance sheet (including pension fund) decision made by a municipal government could be offset by the savings and portfolio decisions of rational private agents. If private individuals and firms recognize the future tax consequences of a government’s (dis-) savings and portfolio decisions, those public decisions could be over-turned by private portfolio decisions.

However, as summarized in Peskin (2001), the conditions that enable private individuals to fully neutralize government savings and portfolio decisions are unlikely to hold in practice. Heterogeneity amongst individuals, borrowing constraints, tax distortions, and imperfect information regarding government policies imply that public pension policies very likely effect the net tax burdens by individuals and, therefore, have real welfare consequences.
Due to such frictions, Peskin (2001) argues that the risk that a pension fund’s returns fail to match its liabilities imposes intergenerational transfers because future generations are not compensated for the taxes they will pay to cover current pension underfunding. Intergenerational equity is likely to be improved if unfunded pension costs are covered by current, rather than future, generations of taxpayers. Unlike the federal government, municipalities have more limited means to cope with underfunding: they cannot inflate away the value of their liabilities via money creation. Thus, unlike the federal social security program that operates on a pay-as-you-go basis, municipalities may want a funded pension plan to avoid unsustainable fiscal imbalances. Peng (2008) believes this is why almost all state and local pension plans are pre-funded.

Even if pension fund deficits are covered by an immediate rise in taxes paid by the current generation of taxpayers, risk-aversion and intra-generational equity motivate a desire to hedge pension liabilities. Not only might a pension fund’s objective be to fund the present value of pension obligations as they accrue, but also to reduce the uncertainty that investment returns fail to match the change in the present value of obligations due to changing market conditions. A fully-funded pension plan whose investments immunize its liabilities fits this ideal of minimizing tax uncertainty.³

We find this argument compelling, with one caveat. It is unlikely that a municipality’s other expenditures for non-pension benefits are fully matched by contemporaneous tax revenues at each point in time. During economic downturns, tax revenues (excluding pension contributions) typically fail to cover non-pension expenditures. If the municipality wished to hedge the net tax surplus of its aggregate balance sheet, pension investments might be chosen so that their returns outperform those of pension liabilities during economic recessions and underperform them during economic expansions. An implication is that pension plans would optimally choose short positions in procyclical risky assets, such as equities.

In practice, municipal governments typically do not incorporate pension fund investment decisions within a single framework for managing their overall balance sheets. Rather, they delegate the administration of a public pension plan to a Board of Trustees who may be appointed or may be elected by plan participants. The Board’s scope tends to focus narrowly on the pension plan. Typically, the Board sets objectives in the form of allocation guidelines in various asset classes, such as fixed income, equities, and alternative investments. A supporting staff headed by an executive director may implement these objectives and/or delegate them to external money managers whose performance is measured against market benchmarks or peers having similar investment styles.

Peng (2008) notes that asset allocation objectives typically are chosen to meet a numerical goal for ex-

³The Lucas and Zeldes (2009) model with costly tax distortions provides these insights.
post investment returns, rather than a desire to hedge the risk of liabilities. Asset class portfolio weights often are chosen based on mean-variance portfolio efficiency with little regard to the risk of pension liabilities. Consequently, pension funds usually are exposed to significant risk from changes in the market value of their liabilities.

Why might public pension fund investment objectives have minimal connection to liability risk? A potential explanation is that archaic GASB accounting clouds the true market valuation of pension liabilities. Lacking a market value measure of liabilities to benchmark the market value of pension assets may discourage hedging. Moreover, opaque accounting of pension liabilities might affect how the performance of management is measured. Rather than, or in addition to, being judged on how the pension fund’s investments hedge the market risk of its liabilities, the pension Board and staff’s performance may be gauged against the investment performance of similar public pension funds. Peskin (2001) describes such a peer group benchmark as belonging to the “traditional” approach to public pension investment management. Park (2009) argues that such a peer group benchmark is a result of career concerns by the plan’s Board of Trustees and staff and is reinforced by “prudent person” fiduciary standards.

3.3.2 Public Pension Fund Portfolio Choice

Given the previous discussion, we begin with a normative model that assumes a pension fund’s objective is to maximize the utility of a representative taxpayer. We later consider an agency model where the fund’s objective is to maximize the utility of the pension plan’s management.

The model is similar to Chen and Pennacchi (2009), and details regarding its derivation can be found in that paper. Let the initial date be 0, and let the end of the pension fund’s performance horizon be date $T$. The interval from date 0 to $T$ might be interpreted as the municipality’s fiscal year or a longer period over which pension over- or under-funding has tax consequences and thereby affects the wealth of the municipality’s residents. Since our focus is on portfolio allocation given an initial level of funding, we assume that contributions by the pension plan’s government employer and its employees are made just prior to date 0, as are any cash outflows to pay retirement benefits. Thus from date 0 to date $T$, the only changes in the values of pension assets and liabilities derive from their market rates of return.

During the interval from dates 0 to $T$, the pension fund’s benchmark, which for now is assumed to be its liabilities, $L_t$, satisfies

$$\frac{dL_t}{L_t} = \alpha_L dt + \sigma_L dz_L$$

(3.1)
where $dz_L$ is a Brownian motion process. The Appendix in Pennacchi and Rastad (2010) shows how this rate of return process can be derived from the value of individual employees' projected benefits and retirees' annuities. The process depends on risks from changes in wages and changes in the value of nominal or, in the case of Cost of Living Adjustments (COLAs), inflation-indexed (real) bonds.

We assume that the pension fund can invest in a portfolio of securities that perfectly match (immunize) the above rate of return on liabilities.\textsuperscript{4} It can also invest in $n$ “alternative” securities, where security $i$’s rate of return satisfies

$$\frac{dA_{i,t}}{A_{i,t}} = \alpha_i dt + \sigma_i dz_i \quad i = 1, \ldots, n$$

(3.2)

and $dz_i$, $i = 1, \ldots, n$ are other correlated Brownian motions such that $\sigma_L dz_L \sigma_i dz_i = \sigma_{iL} dt$. The $\sigma_i$, $\sigma_L$, and $\sigma_{iL}$ are assumed to be constants. $\alpha_i$ and $\alpha_L$ may be time varying, as would be the case with stochastic interest rates, but their spread, $\alpha_i - \alpha_L$, is assumed constant.

For the model applications that we consider, the pension fund’s optimal investments in the $n$ alternative securities are characterized by constant relative proportions, so that the pension fund’s portfolio choice problem simplifies to one of choosing between the liability immunizing portfolio and a single alternative portfolio invested in the alternative securities with constant portfolio weights $\delta_i$, $i=1,\ldots,n$.\textsuperscript{5} If we let $A_t$ be the date $t$ value of this alternative securities portfolio, then its rate of return is

$$\frac{dA_t}{A_t} = \alpha_A dt + \sigma_A dz_A$$

(3.3)

where $\alpha_A \equiv \sum_{i=1}^n \delta_i \alpha_i$, $\sigma_A^2 \equiv \sum_{i=1}^n \sum_{j=1}^n \delta_i \delta_j \sigma_{ij}$, and $dz_A \equiv \sum_{i=1}^n (\delta_i \sigma_i / \sigma_A) dz_i$.

If at date $t$ the pension fund allocates a portfolio proportion $1 - \omega_t$ to the immunizing portfolio and a proportion $\omega_t$ to this alternative securities portfolio, then the value of the pension fund’s asset portfolio, $V_t$, satisfies

$$\frac{dV_t}{V_t} = (1 - \omega_t) \frac{dL_t}{L_t} + \omega_t \frac{dA_t}{A_t}$$

(3.4)

$$= [(1 - \omega_t)\alpha_L + \omega_t \alpha_A] dt + (1 - \omega_t)\sigma_L dz_L + \omega_t \sigma_A dz_A$$

Whenever $\omega_t \neq 0$, the fund’s return in equation (3.4) deviates from the liability immunizing portfolio’s

\textsuperscript{4}Our empirical work weakens this assumption to allow for imperfect immunization.

\textsuperscript{5}The appendix of Chen and Pennacchi (2009) shows that permitting multiple alternative security choices simplifies to a single alternative portfolio choice problem.
return. Now define \( G_t \equiv V_t / L_t \) to be the pension fund’s date \( t \) funding ratio; that is, value of the fund’s assets relative to that of its liabilities. At date 0, \( G_0 = V_0 / L_0 \) but then evolves over the interval from date 0 to date \( T \) as:

\[
\frac{dG_t}{G_t} = \omega_t (\alpha_A - \alpha_L + \sigma_L^2 - \sigma_A \sigma_L) dt + \omega_t (\sigma_A dz_A - \sigma_L dz_L)
\]

(3.5)

Future over- (under-) funding at date \( T \) is assumed to accrue to (be paid by) the municipality’s taxpayers, which in aggregate equals \( V_T - L_T \). Assuming the population of representative taxpaying individuals at date \( T \) is proportional to the value of date \( T \) liabilities, the representative taxpayer’s after tax wealth date \( T \) is

\[
W_T = W^p_T + \lambda \frac{V_T - L_T}{L_T} = W^p_T + \lambda (G_T - 1),
\]

(3.6)

where \( W^p_T \) is the taxpayer’s date \( t \) personal before-tax wealth and \( \lambda > 0 \) is the ratio of pension liabilities per taxpayer. The larger is \( \lambda \), the more sensitive is the taxpayer’s wealth to the pension plan’s funding status. If taxpayers’ utilities display constant relative risk aversion with coefficient \( (1- \gamma) > 0 \), and pension investment policy has the objective of maximizing their utility, then the pension fund’s asset allocation problem is

\[
Max_{\omega(t) \forall t \in [0,T]} \mathbb{E}_0 \left[ \left( \frac{W^p_T + \lambda (G_T - 1)}{\gamma} \right)^\gamma \right]
\]

(3.7)

subject to equation (3.5).

Consider the case where the taxpayer can invest her before-tax wealth, \( W^P_t \), in the same securities available to the pension fund so that it satisfies

\[
dW^P_t / W^P_t = \sum_{i=1}^{n+1} \omega_{i,t} (\alpha_i dt + \sigma_i dz_i)
\]

(3.8)

where \( \alpha_{n+1} \equiv \alpha_L, \sigma_{n+1} \equiv \sigma_L, dz_{n+1} \equiv dz_L \), and \( \omega_{i,t} \) is the taxpayer’s portfolio weight in security \( i \) at \( t \). Since from equation (3.6) after-tax wealth is linear in \( W^P_T \) and \( G_T \), in a Modigliani-Miller world a given pension fund allocation, \( \omega_t \), determining the process for \( G_t \) in (3.5) could be offset (perfectly hedged) by the taxpayer via her appropriate choice of personal portfolio weights \( \omega^P_{i,t} \) in equation (3.8). In this case, the pension fund’s portfolio choice, \( \omega_t \), becomes irrelevant.

65
However, as discussed earlier it is unrealistic that at each point in time a taxpayer is aware of the pension funds’ investments, liabilities, and, in turn, the extent to which she will need to contribute via future taxation to the fund. Poor pension accounting standards together with a lack of information means that the representative taxpayer is likely unaware of the risk faced by the pension funding ratio, $G_t$.

Consequently, the pension fund’s portfolio choice could expose the taxpayer to unhedged risk that reduces her utility below that of a full information environment. In this more realistic setting, a strong case could be made that the pension fund should be managed to eliminate its tracking error risk so that $\omega_t = 0$ and $dG_t = 0 \forall t \in [0, T]$. If so, $G_T = G_0$ and the taxpayer can choose her personal portfolio weights, $\omega_{t,t}^p$, to maximize (3.7) with no uncertainty regarding the tax to be paid at $T$. Moreover, if the pension fund was fully-funded at date 0, then $G_T = G_0 = 1$ and the taxpayer faces no obligation at date $T$.

Might there be circumstances where pension funding risk could benefit taxpayers? If a significant portion of taxpayers lack access to risky investments, pension tracking error risk could afford them this exposure. Empirical evidence shows that individuals’ holdings of risky investments such as stocks has risen over time, but significant proportions of the population continue to hold almost all of their financial assets in risk-free investments such as bank deposits.\(^6\) If these risk-free investments are not chosen willingly but are due to high costs of accessing risky assets, exposure to pension funding risk could potentially raise utility by generating risk premia for taxpayers.

If we consider this case where $W_T^p$ is risk-free, equal to a constant $\bar{W}^p$, then the pension fund’s utility maximizing proportion invested in the alternative security portfolio for any date $t \in [0, T]$ is

$$\omega_t^* = \frac{\alpha_G}{(1-\gamma)\sigma_G^2} \left( 1 + \frac{\bar{W}^p - \lambda}{\lambda G_t} \right)$$

(3.9)

where $\alpha_G \equiv \alpha_A - \alpha_L + \sigma_L^2 - \sigma_{AL}$ and . Equation (3.9) says that when the risk premium $\alpha_G$ is positive (negative), the pension fund takes a long (short) position in the alternative securities.\(^7\) This deviating position is tempered by the relative volatility of the alternative securities, $\sigma_G$, and the taxpayer’s relative risk aversion, $(1-\gamma)$.

Equation (3.9) indicates that the pension fund should change its allocation in the alternative securities as its funding level, $G_t$, changes

$$\frac{\partial|\omega_t^*|}{\partial G_t} = -\left( \bar{W}^p - \lambda \right) \frac{|\alpha_G|}{(1-\gamma)\lambda G_t^2 \sigma_G^2}$$

(3.10)

\(^6\)See, for example, Mankiw and Zeldes (1991) and Campbell (2006).

\(^7\)It can be shown that the optimal portfolio choice ensures that $\left[ 1 + (\bar{W}^p - \lambda) / (\lambda G_t) \right]$ is always positive.
This derivative is negative whenever $\hat{W}^P - \lambda$ is positive, so that when total taxpayer wealth unrelated to pension funding is larger than total pension liabilities, declines in $G_t$ raise $|\omega_t^*|$. This occurs because with constant relative risk aversion, the individual’s utility is maximized when she holds constant fractions of her wealth in risk-free and risky assets. When pension funding, and therefore after-tax wealth, rises, the pension fund optimally lowers its allocation to the risky alternative securities to maintain a constant share in risk-free assets.

Based on our earlier discussion, the model results for this case come with several caveats. First, the representative taxpayer faces taxation risk not just from the municipality’s pension under-funding but from other deficits/surpluses that may arise from the government’s other activities. Recognizing these other sources of tax uncertainty in the individual’s wealth in equation (3.6) could motivate the pension fund to hedge those risks. Second, the analysis ignores federal personal income taxes. Bader and Gold (2007) show that to minimize federal income taxes, it is preferable for the pension fund to invest in high-taxed bonds and for individuals to hold lower-taxed risky assets, such as equities, in their personal portfolios when possible.

Third, when the pension fund ends with a surplus ($G_T > 1$), it may not accrue to the taxpayer. Political pressure leads to a sharing of the surplus with public employees in the form of a reduction in employee contributions or an increase in pension benefits. Thus, it may be illusory to believe that taxpayers benefit by receiving the risk premia generated by pension investments.

These caveats cast doubt on an “activist” public pension fund investment strategy designed to provide risk premia on behalf of investment-constrained taxpayers. Rather, these qualifications favor an optimal investment policy that passively follows the liability immunizing strategy where $\omega_t = 0 \quad \forall \quad t$. Such a strategy would be transparent to taxpayers, allowing many of them to focus on their individual portfolios. It also avoids generating surpluses that taxpayers would be forced to share with employees. In addition, since this passive policy entails primarily fixed-income investments, it delivers federal tax savings.

Shifting from a normative to a positive theory of public pension investment behavior, note that the practice of delegating pension fund management could lead to agency problems where the Board of Trustees and staff maximize their own utility of wealth rather than that of a representative taxpayer. Since stated objectives guiding pension plan investments often downplay the risk of pension liabilities, the Board and staff may be judged against an alternative benchmark such as the investment performance of peer pension plans. In this light, the wealth in equation (3.6) can be re-interpreted as that of the pension fund management where the process followed by the benchmark $L_t$ in (3.1) may be the average rate of return earned by pension plans.

---

8This point is made by Peskin (2001), Bader and Gold (2007), Peng (2008), among others.
other public pension funds. If explicit or implicit (career concern) compensation is performance-related, the pension Board and staff’s wealth will be linked to future relative performance, \( G_T = \frac{V_T}{L_T} \), measured as the pension plan’s funding ratio or its investment performance relative to its peers.

Assuming that the fund managers’ wealth unrelated their pension performance is invested mainly in risk-free assets and is sufficiently large \( (\bar{W}^p - \lambda > 0) \), the solution for optimal portfolio choice continues to satisfy equations (3.9) and (3.10): pension fund managers will increase the fund’s tracking error risk as their relative performance declines. If their wealth that is unrelated to performance is low \( (\bar{W}^p - \lambda < 0) \), the pension fund’s management will decrease its tracking error risk as its performance declines.

The next sections examine the empirical evidence related to our model based on a time series and cross section of state pension plans. We investigate how a state pension plan’s choice of tracking error volatility relates to its characteristics, including the plan’s funding ratio, its performance relative to its peers, its governance, and participants.

### 3.4 Data and Variable Construction

Our data on state pension funds comes from two sources. The first is Wilshire Associates who generously provided us with an annual time series of investment information on 125 state pension funds over the 2000 to 2009 period.\(^9\) This data includes each fund’s actuarial values of liabilities and actuarial and market values of assets for each of the ten years. For each year, it also gives every fund’s proportions of assets allocated to eight categories: U.S. equities, non-U.S. equities, U.S. fixed income, non-U.S. fixed income, real estate, private equity, hedge funds, and other. In addition, it includes each fund’s assumed rate for discounting liabilities and the total payroll for active participants in the pension fund.

The second data source comes from the Boston PlaceNameCollege PlaceTypeCenter for Retirement Research (CRR). This is publicly-available data on 112 state pension funds for the year 2006.\(^10\) It provides individual pension fund characteristics on governance, the type of plan participants (general employees, teachers, or police and firefighters), and numbers of active members and annuitants.

Comparing the state pension funds in the Wilshire data to those in the CRR data led to 97 matches. We selected from the CRR the following variables: the ratio of a pension fund’s Board members who are plan participants to the total Board members of the pension fund; a dummy variable equaling 1 if the pension

---

\(^9\)Due to late reporting, information on only 50 state funds is available for the year 2009.
\(^10\)This data is at http://crr.bc.edu/frequently_requested_data/state_and_local_pension_data_4.html.
fund had a separate investment council (zero otherwise); and a dummy variable equaling 1 if the contribution rate of the pension fund sponsor was statutorily set (zero otherwise).

As previously mentioned, our measure of a pension fund’s overall asset-liability risk-taking is its tracking error volatility. To translate a given pension fund’s asset-class allocations for a given year into this risk measure, we collected for the period January 1997 to April 2010 monthly time series of asset returns in order to estimate a covariance matrix of returns for seven different asset classes.\(^\text{11}\) The following asset return series were chosen: U.S. equities – Vanguard Total Stock Market Index Institutional Fund; Non-U.S. equities – Vanguard Total International Stock Index Fund; U.S. fixed income – Vanguard Total Bond Market Index Institutional Fund; Non-U.S. fixed income – Barclays Capital Global Majors, Ex. U.S., Fixed-Income Index; Real estate – Wilshire U.S. REIT Index; Hedge funds – Morningstar MSCI Composite Hedge Fund Returns; and Private equity – Cambridge Associates U.S. Private Equity Returns.\(^\text{12}\) In addition, while not an asset class specified in the Wilshire data, a part of our analysis of tracking error volatility will use a Treasury Inflation-Protected Securities (TIPS) return series, which we proxied by the Vanguard Inflation-Protected Securities Institutional Fund.

Data on wage growth and returns on nominal and real (inflation-indexed) bonds were used to estimate the variances of market returns on pension fund liabilities as well as these returns’ covariances with the different asset classes. Wages were proxied by the Bureau of Labor Statistics quarterly Employment Cost Index for State and Local Government Workers. For bonds, a monthly time series of 15-year maturity, zero-coupon bond returns were constructed for nominal Treasuries and Treasury Inflation-Protected Securities (TIPS) using the data of Grkaynak et al. (2007) and Grkaynak et al. (2008).\(^\text{13}\) Nominal and real bonds having a 15-year maturity were chosen because several sources indicate that the typical pension fund’s liabilities have a duration of that length.\(^\text{14}\)

In our model, tracking error volatility was represented by the quantity \(|\omega_t| \sigma_G\), which is mathematically equivalent to the square root of the variance of the difference in the rates of return on pension assets \((V_t)\) and pension liabilities \((L_t)\): \[\sigma^2_V + \sigma^2_L - 2\rho_{VL}\sigma_V\sigma_L.\] We estimated the variance of a pension fund’s assets in a given year as \(\sigma^2_V = \omega'\Omega\omega\), where \(\Omega\) is the estimated covariance matrix of returns for the seven asset classes.

---

\(^{11}\)One of the eight Wilshire asset classes is “Other.” Since it averaged only 1.9%, we ignored it when computing tracking error volatility and proportionally increased the other asset class weights.

\(^{12}\)Each asset return series is monthly except for the Cambridge Associates U.S. Private Equity Returns, which is quarterly. We used the monthly series to estimate the covariance matrix of returns, except for those matrix elements relating to private equity returns, which were estimated based on a quarterly returns.

\(^{13}\)This Treasury and TIPS yield data is at http://www.federalreserve.gov/econresdata/researchdata.htm. We converted the monthly real returns on a 15-year TIPS to nominal ones using the Consumer Price Index.

\(^{14}\)Ryan and Fabozzi (2002) state “an average 15.5 duration should be close to the median or average duration of the pension industry.” Also, MercerLLC (2010) uses a 15-year duration for the average pension plan and states that a plan with a typical mix of active members and retirees has a duration between 13 and 16 years.
and \( \omega \) is the 7×1 vector of the fund’s portfolio weights (allocations) in these seven asset classes.

As shown in the Appendix of Pennacchi and Rastad (2010), pension liabilities under a PBO measure are composed of wage and bond risk, where the bond risk is either nominal or real (in the case of COLA benefits). Thus, the standard deviation of pension liabilities, \( \sigma_L \), can be computed from the estimated standard deviations of wage growth (\( \sigma_{Lw} \)) and nominal or real bond returns \( \sigma_{Lp} \) as well as their correlation \( \rho_{wp} \):

\[
\sigma^2_L = \sigma^2_{Lw} + \sigma^2_{Lp} + 2\rho_{wp}\sigma_{Lw}\sigma_{Lp}.
\]

We calculated this pension liability standard deviation, \( \sigma_L \), assuming that liabilities were either purely nominal or purely real liabilities. As detailed in this Appendix, we also adjusted the standard deviations of wage growth and bond risk (duration) using the fund’s ratio of active (employed) to total plan participants (including retirees). These adjustments led to the funds’ liabilities reflecting different sensitivities to wage risk and different durations ranging from 6 to 20 years.

Because in practice pension funds may not view their liabilities as a benchmark but may benchmark their performance to peer public pension funds, we created an additional tracking error volatility measure where, instead of the liability volatilities just discussed, we set

\[
\sigma^2_L = \omega'_a \Omega \omega_a
\]

where \( \omega_a \) is a 7×1 vector of portfolio allocations that are the averages across the 125 state funds for a particular year.

The last step in constructing tracking error volatility calculates the covariance between a fund’s chosen assets and the selected liability/peer performance for a particular year:

\[
\rho_{VL} \sigma_V \sigma_L.
\]

The final measure of tracking error volatility is the square root of

\[
\sigma^2_V + \sigma^2_L - 2\rho_{VL} \sigma_V \sigma_L.
\]

Table 3.1 gives summary statistics for different asset class rates of return, as well as state and local employee wage growth and returns on 15-year zero-coupon nominal and real bonds. It is the annualized standard deviations and correlations from this table, which were calculated over the 1997 to 2010 period, that we use to construct tracking error volatilities. Of particular interest is the estimated correlation of -0.27 between state and local wage growth and U.S. equity returns. Prior research, including Black (1989), Cardinale (2003), Lucas and Zeldes (2006) and Lucas and Zeldes (2009) advocate pension fund investments in equities as a way to hedge wage uncertainty under the presumption that equities and wages are positively correlated.

A potential criticism of our negative wage-equity correlation estimate is that it is calculated over a quarterly holding period. Recommendations for using stocks to hedge wage growth assume that their correlation is positive over a longer holding period. Models by Lucas and Zeldes (2006), Benzoni et al. (2007), and Geanakoplos and Zeldes (2009) show that wage growth and stock returns can have zero short-run correlation.
### Table 3.1: Summary Statistics of Asset Returns and Wage Growth:

Summary statistics for rates of return on eight different asset categories: U.S. Equities; Non-U.S. Equities; U.S. Fixed Income; Non-U.S. Fixed Income; U.S. Real Estate; Private Equity; Hedge Funds; and Treasury Inflation-Protected Securities (TIPS), as well as State and Local Employee Cost Index wage growth and 15-Year Nominal and Real (TIPS) Bond returns. State and local employee wage growth and private equity returns are calculated based on quarterly data from the second quarter of 1997 to the first quarter of 2010. All other statistics are calculated from monthly return data between February 1997 and April 2010. Means and standard deviations are annualized. Minimums and maximums for wage growth and private equity are one quarter rates while minimums and maximums for the other series are one month rates.

#### Rates of Return:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0533</td>
<td>0.0488</td>
<td>0.0727</td>
<td>0.0529</td>
<td>0.0858</td>
<td>0.1295</td>
<td>0.0815</td>
<td>0.0666</td>
<td>0.0293</td>
<td>0.0795</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.1714</td>
<td>0.1891</td>
<td>0.0885</td>
<td>0.0838</td>
<td>0.2447</td>
<td>0.1232</td>
<td>0.0512</td>
<td>0.0601</td>
<td>0.0107</td>
<td>0.1376</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.1939</td>
<td>-0.2496</td>
<td>-0.0908</td>
<td>-0.0609</td>
<td>-0.3913</td>
<td>-0.1567</td>
<td>-0.0537</td>
<td>-0.0927</td>
<td>0.0010</td>
<td>-0.1466</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.1007</td>
<td>0.1341</td>
<td>0.1072</td>
<td>0.0768</td>
<td>0.2841</td>
<td>0.1466</td>
<td>0.0501</td>
<td>0.0588</td>
<td>0.0196</td>
<td>0.1760</td>
</tr>
</tbody>
</table>

#### Correlation Matrix:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US Equities</td>
<td>1</td>
<td>0.87091</td>
<td>0.06412</td>
<td>0.04328</td>
<td>0.57669</td>
<td>0.79986</td>
<td>0.63777</td>
<td>0.06630</td>
<td>-0.27389</td>
<td>-0.07907</td>
</tr>
<tr>
<td>Non-US Equities</td>
<td>1</td>
<td>1</td>
<td>0.49819</td>
<td>0.11068</td>
<td>0.17921</td>
<td>0.73599</td>
<td>0.69356</td>
<td>0.14086</td>
<td>-0.29040</td>
<td>-0.04836</td>
</tr>
<tr>
<td>US Fixed Income</td>
<td>0.06412</td>
<td>1</td>
<td>1</td>
<td>0.49819</td>
<td>0.17921</td>
<td>0.73599</td>
<td>0.69356</td>
<td>0.14086</td>
<td>-0.29040</td>
<td>-0.04836</td>
</tr>
<tr>
<td>Non-US Fixed Income</td>
<td>0.04328</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>US-Real Estate</td>
<td>0.57669</td>
<td>0.49819</td>
<td>0.17921</td>
<td>0.11068</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Private Equity</td>
<td>0.79986</td>
<td>0.73599</td>
<td>0.17921</td>
<td>0.11068</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>TIPS</td>
<td>0.63777</td>
<td>0.69356</td>
<td>0.12148</td>
<td>0.02980</td>
<td>0.33082</td>
<td>0.74276</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>S&amp;L Wage Growth</td>
<td>-0.27389</td>
<td>-0.29040</td>
<td>0.20711</td>
<td>0.14081</td>
<td>-0.00980</td>
<td>-0.25375</td>
<td>-0.18294</td>
<td>0.14551</td>
<td>0.18128</td>
<td>0.68908</td>
</tr>
<tr>
<td>15-Year Nominal Bond</td>
<td>-0.07907</td>
<td>-0.04836</td>
<td>0.94010</td>
<td>0.46061</td>
<td>0.02485</td>
<td>-0.29535</td>
<td>-0.01632</td>
<td>0.66265</td>
<td>0.25426</td>
<td>1</td>
</tr>
<tr>
<td>15-Year Real Bond</td>
<td>0.04059</td>
<td>0.11604</td>
<td>0.72275</td>
<td>0.49579</td>
<td>0.16940</td>
<td>-0.18294</td>
<td>0.14551</td>
<td>0.95675</td>
<td>0.18128</td>
<td>0.68908</td>
</tr>
</tbody>
</table>

**Observations:** 159 159 159 159 159 52 159 52 159 52 159
but, due to mean reversion, a correlation that approaches 1 over long horizons.\textsuperscript{15}

As a check, we calculated wage-equity correlations, as well as wage-bond correlations, over longer holding periods. Because the BLS state and local wage index only starts in 1981, we use the BLS national wage index which starts in 1952. Equity returns are for the S&P 500 index or a small stock index and bond returns are those for 10-year or 20-year maturity Treasury bonds.\textsuperscript{16}

<table>
<thead>
<tr>
<th>Holding Period</th>
<th>Correlation</th>
<th>Observations</th>
<th>t-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-year</td>
<td>-0.0389</td>
<td>57</td>
<td>-0.2890</td>
<td>0.7737</td>
</tr>
<tr>
<td>five-year</td>
<td>-0.4348</td>
<td>11</td>
<td>-1.4483</td>
<td>0.1815</td>
</tr>
<tr>
<td>nine-year</td>
<td>-0.4406</td>
<td>6</td>
<td>-0.9816</td>
<td>0.3819</td>
</tr>
</tbody>
</table>

Panel A: Correlation between National Wage Index Growth and S&P 500 Return

<table>
<thead>
<tr>
<th>Holding Period</th>
<th>Correlation</th>
<th>Observations</th>
<th>t-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-year</td>
<td>-0.0696</td>
<td>57</td>
<td>-0.4583</td>
<td>0.6543</td>
</tr>
<tr>
<td>five-year</td>
<td>-0.0106</td>
<td>11</td>
<td>-0.0391</td>
<td>0.9753</td>
</tr>
<tr>
<td>nine-year</td>
<td>-0.5626</td>
<td>6</td>
<td>-1.3611</td>
<td>0.2451</td>
</tr>
</tbody>
</table>

Panel B: Correlation between National Wage Index Growth and Small Firm Index Return

<table>
<thead>
<tr>
<th>Holding Period</th>
<th>Correlation</th>
<th>Observations</th>
<th>t-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-year</td>
<td>-0.2163</td>
<td>57</td>
<td>-1.6427</td>
<td>0.1061</td>
</tr>
<tr>
<td>five-year</td>
<td>-0.2645</td>
<td>11</td>
<td>-0.8228</td>
<td>0.4319</td>
</tr>
<tr>
<td>nine-year</td>
<td>0.2449</td>
<td>6</td>
<td>0.5053</td>
<td>0.6399</td>
</tr>
</tbody>
</table>

Panel C: Correlation between National Wage Index Growth and 20-year Treasury bond Return

<table>
<thead>
<tr>
<th>Holding Period</th>
<th>Correlation</th>
<th>Observations</th>
<th>t-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>one-year</td>
<td>-0.1095</td>
<td>57</td>
<td>-0.8171</td>
<td>0.4174</td>
</tr>
<tr>
<td>five-year</td>
<td>-0.0746</td>
<td>11</td>
<td>-0.2245</td>
<td>0.8274</td>
</tr>
<tr>
<td>nine-year</td>
<td>0.4752</td>
<td>6</td>
<td>1.0892</td>
<td>0.3408</td>
</tr>
</tbody>
</table>

Panel D: Correlation between National Wage Index Growth and 10-year Treasury note Return

Table 3.2: Correlations of Equity and Bond Returns with National Wage Growth, 1952-2008

Table 3.2 presents estimated wage – asset return correlations for holding periods from one year to nine years. Panel A indicates little evidence of a positive correlation between wages and large firm stock returns, and the correlation point estimates trend more negative as the holding period increases. Panel B shows the same qualitative correlations hold for small firm stock returns. This pattern of correlations is consistent with Jermann (1999) who estimates wage – stock correlations over the 1929 – 1996 period, finding correlation point estimates that are negative at horizons from 7 to 17 years before turning positive. Since a typical innovations in human capital and financial asset returns that is counter to models such as Benzoni et al. (2007).\textsuperscript{16}

The small stock index was obtained from Professor Kenneth French’s website and represents returns on the smallest 30% of stocks traded on the NYSE, AMEX, and NASDAQ. The S&P500 and bond returns are from the Center for Research in Security Prices (CRSP).
pension fund’s duration of liabilities is about 15 years, stocks may not be the best hedge of wage risk.

Another asset class, possibly bonds, might be a better hedge. Panels C and D of Table 3.2 find a correlation between wage growth and bond returns whose point estimates becomes more positive as the holding period increases. Though none of the correlation estimates in Table 3.2 are statistically significant, the evidence does not justify large equity investments for hedging wage risk at horizons relevant to a typical pension fund.

Along with asset return and wage growth standard deviations and correlations, the final input in constructing tracking error volatilities is each pension fund’s asset allocations for each year. Table 3.3 gives summary statistics of these allocations and the resulting tracking error volatilities calculated for different benchmarks. Average volatilities are lower when the benchmark is real liabilities compared to nominal liabilities. Unsurprisingly, with a benchmark equal to the average of funds’ allocations, average tracking error volatilities are the least.

Table 3.3 also includes yearly averages for two variables that our model predicts may influence tracking error volatility: the plan’s funding ratio and its investment return. The funding ratio is measured as the market value of assets divided by the actuarial value of liabilities. The plan’s investment return in a given year is estimated as the product of its asset-class allocation weights and the returns earned by each asset class. The average funding ratio of these 125 state pension plans was highest at 109% during 2000 prior to the technology stock decline and was at a minimum of 58% during 2009, largely due to an estimated 30% investment loss in 2008.

3.5 Empirical Evidence

3.5.1 Tracking Error Minimizing Allocations and Alternative Portfolio Allocations

As a prelude to examining how funds’ tracking error volatilities vary with their plan characteristics, we first estimate the asset allocations that would minimize the typical fund’s tracking error volatility. We also present estimates of allocations for the “alternative” security portfolio modeled in equation (3.9). These allocations are for a plan having 15-year duration liabilities and a sensitivity to wages reflecting a ratio of
### Table 3.3: Time Series of Average Asset Allocations, Tracking Error Volatilities, and Funding Ratio

Sample average over 125 state pension plans for eight different asset allocations, various tracking error volatilities, funding ratios (market value of assets divided by actuarial value of liabilities), and return on investments. The returns on investments were estimated as the product of their asset-class allocation weights and the returns earned by each of the asset classes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US Equities</td>
<td>0.3996</td>
<td>0.4174</td>
<td>0.3966</td>
<td>0.4261</td>
<td>0.4363</td>
<td>0.4322</td>
<td>0.4172</td>
<td>0.3957</td>
<td>0.3378</td>
<td>0.3299</td>
</tr>
<tr>
<td>Non-US Equities</td>
<td>0.1231</td>
<td>0.1197</td>
<td>0.1303</td>
<td>0.1348</td>
<td>0.1497</td>
<td>0.1534</td>
<td>0.1743</td>
<td>0.1848</td>
<td>0.1802</td>
<td>0.1880</td>
</tr>
<tr>
<td>US Fixed Income</td>
<td>0.2796</td>
<td>0.3258</td>
<td>0.3306</td>
<td>0.3162</td>
<td>0.2769</td>
<td>0.2757</td>
<td>0.2648</td>
<td>0.2629</td>
<td>0.2588</td>
<td>0.2486</td>
</tr>
<tr>
<td>Non-US Fixed Income</td>
<td>0.0156</td>
<td>0.0144</td>
<td>0.0141</td>
<td>0.0122</td>
<td>0.0114</td>
<td>0.0087</td>
<td>0.0085</td>
<td>0.0143</td>
<td>0.0090</td>
<td></td>
</tr>
<tr>
<td>US-Real Estate</td>
<td>0.0314</td>
<td>0.0372</td>
<td>0.0418</td>
<td>0.0405</td>
<td>0.0396</td>
<td>0.0434</td>
<td>0.0520</td>
<td>0.0553</td>
<td>0.0647</td>
<td>0.0524</td>
</tr>
<tr>
<td>Private Equity</td>
<td>0.0378</td>
<td>0.0427</td>
<td>0.0419</td>
<td>0.0456</td>
<td>0.0429</td>
<td>0.0432</td>
<td>0.0454</td>
<td>0.0499</td>
<td>0.0715</td>
<td>0.0659</td>
</tr>
<tr>
<td>Hedge Fund</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0002</td>
<td>0.0013</td>
<td>0.0089</td>
<td>0.0082</td>
<td>0.0115</td>
<td>0.0238</td>
</tr>
<tr>
<td>Other</td>
<td>0.0091</td>
<td>0.0027</td>
<td>0.0037</td>
<td>0.0087</td>
<td>0.0269</td>
<td>0.0235</td>
<td>0.0286</td>
<td>0.0267</td>
<td>0.0342</td>
<td>0.0303</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Liabilities</td>
<td>0.1687</td>
<td>0.1651</td>
<td>0.1634</td>
<td>0.1678</td>
<td>0.1742</td>
<td>0.1746</td>
<td>0.1761</td>
<td>0.1764</td>
<td>0.1732</td>
<td>0.1748</td>
</tr>
<tr>
<td>Real Liabilities</td>
<td>0.1363</td>
<td>0.1329</td>
<td>0.1312</td>
<td>0.1351</td>
<td>0.1408</td>
<td>0.1410</td>
<td>0.1422</td>
<td>0.1423</td>
<td>0.1389</td>
<td>0.1396</td>
</tr>
<tr>
<td>Average Asset Allocation of Peers</td>
<td>0.0195</td>
<td>0.0187</td>
<td>0.0185</td>
<td>0.0173</td>
<td>0.0157</td>
<td>0.0159</td>
<td>0.0163</td>
<td>0.0162</td>
<td>0.0171</td>
<td>0.0143</td>
</tr>
</tbody>
</table>

| Funding Ratio                            | 1.0852| 0.9226| 0.7912| 0.7775| 0.8250| 0.8454| 0.8618| 0.9336| 0.7831| 0.5832|
| Return on Investments                     | -0.0129| -0.0524| -0.0723| 0.2093| 0.1304| 0.0807| 0.1429| 0.0678| -0.3067| 0.1881|
| Growth Rate of Liabilities                | -0.0789| 0.0639| 0.0634| 0.0608| 0.0628| 0.0653| 0.0607| 0.0617| 0.0504|
active participants to total participants of 66.57%, which is our sample average. Calculating allocations for both portfolios employ the estimated covariance matrix for the asset classes’ returns as well as these returns’ covariances with wages and the 15-year maturity bonds. To calculate the alternative portfolio allocations also requires estimates of the assets’ expected returns which are taken from Table 3.1.17

Table 3.4 gives the portfolio allocations that best hedge (immunize) this typical pension fund’s liabilities.18 Columns 1, 2, and 3 show results when the pension fund’s liabilities are purely nominal, having no COLAs. In column 1, the unconstrained allocation that minimizes tracking error volatility calls for a 9% short position in U.S. equities, a 160% allocation to U.S. fixed income, a 24% allocation to private equity, and a 67% short position in hedge funds. The huge allocation to U.S. fixed income is partly explained by our assumption that the pension funds’ fixed-income investments are of lower duration (lower interest rate sensitivity) than the 15-year duration of the pension funds’ nominal liabilities.19 These allocations imply that the fund should borrow via short positions in other asset categories in order to increase its U.S. fixed income investment, thereby raising its asset interest sensitivity. If, instead, the pension fund’s U.S. fixed income portfolio was assumed to take the form of 15-year zero-coupon bonds, then its tracking error minimizing allocation would be approximately 100%, rather than 160%, in U.S. fixed income.20

Because a short position in hedge fund investments is infeasible and a large allocation to private equity may be unrealistic, column 2 of Table 3.4 shows the volatility minimizing allocation without private equity or hedge fund investments. There the allocation to fixed income becomes 136%, with a 13% short position in U.S. equities and a 17% short position in non-U.S. fixed income. Column 3 reports immunizing allocations when short sales are not permitted. Interestingly, the allocation is 100% in fixed-income.

Columns 5 to 9 of Table 3.4 assume that the pension fund’s liabilities are fully inflation-indexed, real liabilities. Given the widespread presence of COLAs in state and local pension benefits, this case may be most realistic. Though the Wilshire data do not specifically identify funds’ allocation to inflation-indexed bonds

---

17 Expected returns were estimated as the means in Table 3.1 except for U.S. Equities and U.S. Fixed Income where we used estimates of 0.0742 and 0.0520, respectively, which are the average of forecasts over the 1997 to 2010 period from the Survey of Professional Forecasters. See http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/. 
18 The tracking error minimizing weights are given by $\omega^* = \Omega^{-1}\Psi - \kappa\Omega^{-1}1$ where $\Omega$ is the $7 \times 7$ covariance matrix of asset returns and $\Psi$ is a $7 \times 2$ vector of covariances between the asset returns and pension liabilities (wage and nominal or real bond), $1$ is a $7 \times 1$ vector of ones, and $\kappa \equiv (1'\Omega^{-1}\Psi - 1) / (1'\Omega^{-1}1)$. 
19 Table 3.1 shows that the annualized standard deviation of U.S. Fixed Income returns is 0.0885 while that of the 15-year nominal bond is 0.1376, implying that the duration of the pension funds’ U.S. fixed-income investments is approximately $15 \times (0.0885/0.1376) = 9.6$ years. Adams and Smith (2009) believe that the typical pension funds’ fixed income investments are of a lower duration than that of their liabilities.
20 This ignores the desire to hedge the wage risk of liabilities. The allocation might exceed 100% given the positive bond return - wage growth correlation.
<table>
<thead>
<tr>
<th></th>
<th>Nominal Liabilities</th>
<th>Real Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variance Minimizing Portfolio</td>
<td>Alternative Portfolio</td>
</tr>
<tr>
<td></td>
<td>No Constraints</td>
<td>No Private Equity &amp; Hedge Funds</td>
</tr>
<tr>
<td>US Equities</td>
<td>-0.0886</td>
<td>-0.1259</td>
</tr>
<tr>
<td>Non-US Equities</td>
<td>-0.0109</td>
<td>-0.0285</td>
</tr>
<tr>
<td>US Fixed Income</td>
<td>1.5991</td>
<td>1.3586</td>
</tr>
<tr>
<td>Non-US Fixed Income</td>
<td>-0.0103</td>
<td>-0.1719</td>
</tr>
<tr>
<td>US-Real Estate</td>
<td>-0.0662</td>
<td>-0.0322</td>
</tr>
<tr>
<td>Private Equity</td>
<td>0.2442</td>
<td>-</td>
</tr>
<tr>
<td>Hedge Fund</td>
<td>-0.6673</td>
<td>-</td>
</tr>
<tr>
<td>TIPS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tracking Error Standard Deviation</td>
<td>0.04054</td>
<td>0.05004</td>
</tr>
<tr>
<td>Expected Return</td>
<td>0.04706</td>
<td>0.04806</td>
</tr>
</tbody>
</table>

Table 3.4: Tracking Error Minimizing Allocations and Alternative Portfolio Allocations: The entries are the pension asset portfolio weights that minimize a typical pension fund’s tracking error volatility.
(they would likely be included in U.S. fixed income), we consider allocations with and without including TIPS. When TIPS are included, column 5 shows that the immunizing allocation is 139% in TIPS with a large short position in private equity. If, as in column 6, private equity and hedge fund investments are disallowed, the TIPS allocation becomes 101%. Column 7 reports allocations when TIPS are excluded. The allocation to U.S. fixed income becomes 60% and the allocation to non-U.S. fixed income is a large 18%. Moreover, the allocation to hedge funds switches signs, to a large 38% allocation financed with a 20% short position in private equity.

In column 8, the allocation that excludes TIPS and short sales calls for U.S. and foreign fixed income allocations of 70% and 21%, respectively, with an 8% allocation to hedge funds. The sizable position in foreign fixed income makes sense since it should hedge U.S. inflation if exchange rates adjust with purchasing power parity. Column 9 permits investments in TIPS but no short positions. The result is a simple 86% allocation to TIPS and a 14% allocation to U.S. fixed-income.

Columns 4 and 10 of Table 3.4 give the “alternative security” portfolio allocations corresponding to the multivariate analog of the quantity $\frac{\alpha_G}{\sigma^2_G}$ in equation (3.9). Unlike the tracking error minimizing portfolio, the alternative security portfolio depends on the asset classes’ expected returns which are difficult to estimate accurately (Merton (1980)). Hence, these estimated alternative portfolio weights should be interpreted cautiously. Column 4 shows that when the pension fund wishes to generate risk premia on behalf of the representative taxpayer and liabilities are nominal, the pension fund will start from its tracking error minimizing portfolio in column (3.1) and move in the direction of eliminating its long position in U.S. fixed income, short foreign equities, and investing primarily in hedge funds, non-U.S. fixed income, private equity. Equation (3.9) indicates the extent of this deviation from the immunizing portfolio will be inversely related to risk-aversion and pension funding.

When liabilities are real, column (3.10) indicates that generating risk premia leads the pension fund to undo its large TIPS position shown in column (3.5) and, similar to the nominal liability case, short foreign equities, and allocate more to private equity, hedge funds, and foreign fixed income. Note from the last row in Table 3.4 that such a deviation generates an excess return of approximately $0.160 - 0.044 = 11.6\%$.

Finally, note from the second to last row of Table 3.4, the minimum tracking error volatilities for both nominal and real liabilities are much lower than the average tracking error volatilities estimated in Table

\[\textit{Chen and Pennacchi (2009).}\]
3.3 for our sample of state pension funds. Figure 3.2 explores this point further by plotting the sample distribution of tracking error volatilities. It graphs, for each fund during each year, the volatility of tracking error relative to the fund’s real liabilities (horizontal axis) versus the fund’s volatility of tracking error relative to its peers (vertical axis). The distribution appears skewed, with only a few pension funds (primarily those with very high allocations to U.S. fixed-income) choosing allocations that immunize real liabilities. The vast majority of pension funds’ allocations result in peer tracking error volatilities below 4%, indicating a tendency to herd far from an immunizing portfolio.

3.5.2 Tracking Error Volatility and Plan Characteristics

Let us now investigate the relationship between a fund’s characteristics and its choice of tracking error volatility based on the following linear regression model:

\[
\sigma_{TE,i,t} = \beta_1 \times (\text{Funding ratio})_{t-1} + \beta_2 \times (\text{Return relative to peers})_{t-1} \\
+ \beta_3 \times (\text{Governance variables}) + \beta_4 \times (\text{Other Control variables}) + \varepsilon_{it}
\]  

(3.11)
where $\sigma_{TE,i,t}$ is the tracking error volatility of fund $i$ in year $t$. Under the normative view that the pension fund should maximize the utility of a representative taxpayer, our model predicts different relationships between tracking error and pension funding under different assumptions. As discussed earlier, if the representative taxpayer is able to invest in the same asset classes available to the pension fund, utility is maximized if the fund simply minimizes tracking error volatility, implying no relationship between tracking error volatility and the plan’s funding ratio ($\beta_1 = 0$). Alternatively, if the representative taxpayer has access only to risk-free investments, utility maximization could imply that tracking error declines with the plan’s funding ratio ($\beta_1 < 0$).

Under the positive theory that the pension fund is managed to maximize the utility of wealth of the pension Board and staff, and their benchmark for compensation includes the “traditional” one of return performance relative to the fund’s peers, then tracking error volatility may vary with the fund’s past investment return. In particular, if the personal wealth of the pension Board and staff is primarily in risk-free investments whose value is large relative to their management compensation ($\bar{W}_p - \lambda > 0$), then tracking error risk should be high following a poor relative return ($\beta_2 < 0$). Conversely, if the compensation of the Board and staff is large relative to their personal wealth ($\bar{W}_p - \lambda < 0$) or, as may seem reasonable, their personal wealth is invested in primarily risky assets, then they may optimally lower tracking error risk following a poor relative return ($\beta_2 > 0$).

The regression includes additional explanatory variables listed in Table 3.5. These include governance-related variables, controls related to the type of participants, and other characteristics such as a proxy for fund size (the natural log of the market value of pension assets), the ratio of payroll to pension liabilities, and the rate chosen by the fund to discount its liabilities. We include the fund’s chosen discount rate, not because it is an exogenous variable, but to explore whether tracking error risk is linked to a possible motive to underreport liabilities via a higher discount rate.

Table 3.6 reports results of equation (3.11). Each regression specification controls for time (year) fixed-effects. In columns 1, 3, 5, 7, and 9, the regressions include time-invariant fund characteristics from the CRR dataset, so our observations drop to 97 pension plans per year. In columns 2, 6, and 10 we exclude the CRR data but control for fund fixed-effects, in which case the observations equal 125 pension plans per year. Columns 4 and 8 run regressions of equation (3.11) in first differences (annual changes) for all of the variables.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tracking Error Volatility Benchmarked on:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal Liabilities</td>
<td>903</td>
<td>0.1712</td>
<td>0.0209</td>
<td>0.0746</td>
<td>0.2163</td>
</tr>
<tr>
<td>Real Liabilities</td>
<td>903</td>
<td>0.1380</td>
<td>0.0154</td>
<td>0.0741</td>
<td>0.1734</td>
</tr>
<tr>
<td>Average Asset Allocation of Peers</td>
<td>1121</td>
<td>0.0171</td>
<td>0.0132</td>
<td>0.0026</td>
<td>0.1294</td>
</tr>
<tr>
<td>Portfolio Share:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>of Total Equity</td>
<td>1121</td>
<td>0.5839</td>
<td>0.0995</td>
<td>0</td>
<td>0.8254</td>
</tr>
<tr>
<td>of US Equity</td>
<td>1121</td>
<td>0.4244</td>
<td>0.1029</td>
<td>0.0741</td>
<td>0.7945</td>
</tr>
<tr>
<td>Not Invested in US Fixed Income</td>
<td>1121</td>
<td>0.6988</td>
<td>0.1131</td>
<td>0</td>
<td>0.8971</td>
</tr>
<tr>
<td><strong>Primary Explanatory Variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding Ratio</td>
<td>1121</td>
<td>0.8528</td>
<td>0.1970</td>
<td>0.1908</td>
<td>1.7520</td>
</tr>
<tr>
<td>Return on Investments</td>
<td>1121</td>
<td>0.0345</td>
<td>0.1463</td>
<td>-0.3891</td>
<td>0.2529</td>
</tr>
<tr>
<td><strong>Governance Variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants to Total Board Members</td>
<td>921</td>
<td>0.5646</td>
<td>0.2227</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy for Separate Investment Council</td>
<td>930</td>
<td>0.3839</td>
<td>0.4866</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy for Legal Restrictions</td>
<td>930</td>
<td>0.3086</td>
<td>0.4622</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Other Control Variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Payroll to Actuarial Liabilities</td>
<td>1121</td>
<td>0.2360</td>
<td>0.2029</td>
<td>0.0085</td>
<td>2.6248</td>
</tr>
<tr>
<td>Discount Rate</td>
<td>1121</td>
<td>0.0801</td>
<td>0.0039</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td>Dummy for Teachers Fund</td>
<td>930</td>
<td>0.5065</td>
<td>0.5002</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy for General State Fund</td>
<td>930</td>
<td>0.6226</td>
<td>0.4850</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy for Police and Fire Fighters Fund</td>
<td>930</td>
<td>0.4441</td>
<td>0.4971</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The results in Table 3.6 do differ little when the dependent variable, tracking error volatility, is based on a nominal or real pension liability benchmark. As shown in columns 1 to 8, three explanatory variables are always statistically significant: the prior year’s investment return relative to peers; the fund’s chosen rate used to discount liabilities; and the proportion of members of the Board of Trustees who are participants. In addition, columns 2 and 6 indicate that the prior year’s funding ratio is significant when the larger 125 plan sample, estimated with fund fixed effects, is used. Notably, prior returns and the discount rate remain significant when the regressions are in changes.

The negative coefficients on the fund’s prior-year return and its funding ratio might be interpreted as risk-shifting behavior counter to a policy of pure immunization but consistent with a policy that generates risk premia for taxpayers as in equation (3.9). However, the positive relationship between a fund’s chosen rate to discount liabilities and its tracking error volatility indicates possible moral hazard on the part of pension fund management.\(^{22}\) According to GASB standards, funds with higher tracking error volatility may justify a higher discount rate because they are investing in assets with higher systematic (priced) risks. Or,

\(^{22}\)That funds choosing higher discount rates also choose higher tracking error volatility is consistent with Park (2009) who finds that public plans selecting higher discount rates are more likely to invest in real estate and alternative investments (including private equity and hedge funds).
### Table 3.6: Regression Results

Numbers in parenthesis are p-values. Each regression includes time (year) fixed effects. In columns (3.1) to (3.4) tracking error volatility is constructed assuming that pension liabilities are in nominal terms while in columns (3.5) to (3.8) tracking error volatility is constructed assuming pension liabilities are fully inflation-indexed real liabilities. For the regressions in columns (3.4) and (3.8), all variables are in annual changes, rather than levels. In columns (3.9) and (3.10) tracking error volatility is constructed assuming a benchmark equal to the average sample asset allocation of peer public pension funds.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Nominal Liabilities</th>
<th>Real Liabilities</th>
<th>Average Asset Allocation of Peers</th>
<th>Total U.S. Non Fixed Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag Funding Ratio</td>
<td>-0.0043 (0.20)</td>
<td>-0.0096 (0.04)</td>
<td>-0.0007 (0.86)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0031 (0.27)</td>
<td>-0.0031 (0.75)</td>
<td>-0.0018 (0.27)</td>
<td></td>
</tr>
<tr>
<td>Lag Return Relative</td>
<td>-0.078 (0.00)</td>
<td>-0.0821 (0.00)</td>
<td>-0.089 (0.00)</td>
<td></td>
</tr>
<tr>
<td>to Peer Average</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
</tr>
<tr>
<td>Lag of Ln Market</td>
<td>0.0008 (0.43)</td>
<td>0.0055 (0.22)</td>
<td>0.0135 (0.83)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0005 (0.08)</td>
<td>0.0003 (0.74)</td>
<td>0.0068 (0.18)</td>
<td></td>
</tr>
<tr>
<td>Value of Assets</td>
<td>0.0004 (0.00)</td>
<td>0.0049 (0.02)</td>
<td>0.0035 (0.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0018 (0.00)</td>
<td>0.0041 (0.00)</td>
<td>0.0008 (0.00)</td>
<td></td>
</tr>
<tr>
<td>Lag Payroll to</td>
<td>0.0006 (0.41)</td>
<td>0.0045 (0.67)</td>
<td>-0.0009 (0.63)</td>
<td></td>
</tr>
<tr>
<td>Actuarial Liabilities</td>
<td>-0.0003 (0.55)</td>
<td>-0.0027 (0.70)</td>
<td>0.0009 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Participants to Total</td>
<td>0.0152 (0.58)</td>
<td>0.0192 (0.50)</td>
<td>-0.0022 (0.77)</td>
<td></td>
</tr>
<tr>
<td>Board Members</td>
<td>0.0012 (0.00)</td>
<td>0.0024 (0.01)</td>
<td>0.0019 (0.00)</td>
<td></td>
</tr>
<tr>
<td>Teachers Fund</td>
<td>0.00012 (0.00)</td>
<td>0.0001 (0.0)</td>
<td>0.0009 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Dummy</td>
<td>0.0005 (0.65)</td>
<td>0.0004 (0.75)</td>
<td>-0.0009 (0.7)</td>
<td></td>
</tr>
<tr>
<td>General State Fund</td>
<td>-0.0005 (0.84)</td>
<td>0.0004 (0.92)</td>
<td>0.0009 (0.77)</td>
<td></td>
</tr>
<tr>
<td>Dummy</td>
<td>-9.4e-6 (0.998)</td>
<td>-0.0009 (0.92)</td>
<td>0.0025 (0.77)</td>
<td></td>
</tr>
<tr>
<td>Police &amp; Fire Fighters</td>
<td>-0.0031 (0.84)</td>
<td>-0.0035 (0.92)</td>
<td>0.0001 (0.92)</td>
<td></td>
</tr>
<tr>
<td>Fund Dummy</td>
<td>-0.0033 (0.84)</td>
<td>-0.0036 (0.84)</td>
<td>-0.0009 (0.84)</td>
<td></td>
</tr>
<tr>
<td>Separate Investment</td>
<td>-0.0002 (0.24)</td>
<td>-0.0003 (0.40)</td>
<td>0.0001 (0.35)</td>
<td></td>
</tr>
<tr>
<td>Council Dummy</td>
<td>-0.0012 (0.40)</td>
<td>-0.0002 (0.35)</td>
<td>-0.0002 (0.35)</td>
<td></td>
</tr>
<tr>
<td>Legal Restrictions</td>
<td>-0.0017 (0.50)</td>
<td>-0.0043 (0.28)</td>
<td>-0.0022 (0.45)</td>
<td></td>
</tr>
<tr>
<td>Dummy</td>
<td>0.00033 (0.28)</td>
<td>0.00028 (0.45)</td>
<td>0.00017 (0.45)</td>
<td></td>
</tr>
<tr>
<td>Fund Fixed Effects</td>
<td>No (787)</td>
<td>Yes (795)</td>
<td>No (787)</td>
<td>No (787)</td>
</tr>
<tr>
<td>N</td>
<td>787</td>
<td>795</td>
<td>787</td>
<td>795</td>
</tr>
<tr>
<td>R²</td>
<td>0.32</td>
<td>0.13</td>
<td>0.20</td>
<td>0.20</td>
</tr>
</tbody>
</table>
it may be that a fund is motivated to select a higher discount rate to understate its liabilities and fictitiously increase its net worth, thereby justifying greater tracking error risk. Since the direction of causality is unclear and the chosen discount rate may be endogenous, regressions reported in columns 3 and 7 exclude it. Doing so does not qualitatively change the results.

The finding that a fund takes more tracking error risk when it has greater participant representation on its Board of Trustees might be explained in a couple of ways. If pension participants are less financially literate than typical Board members, they may be less able to select asset allocations that immunize the plan’s liabilities. Alternatively, participants may intentionally take more tracking error risk to increase the likelihood of a significant pension surplus that will accrue to them in the form of increased benefits or lower employee contribution rates.

Columns 9 and 10 report regression results where the benchmark is the average investment allocation by peer pension plans. The estimated coefficients tend to be qualitatively different from the previous regressions where a liability benchmark was used. The positive and statistically significant coefficients on the plan’s prior-year funding ratio and prior investment return indicate that pension funds deviate more from their peers following better performance. Such behavior would be consistent with the positive theory that the pension fund Board and staff maximize their own utility and their compensation is large relative their personal wealth or their personal wealth is invested in primarily risky assets.

The finding that better-performing pension funds increase their tracking error relative to the average allocation of other funds is not inconsistent with the previous regression results. Figure 3.2 shows that peer and liability benchmarks are quite different, so when better-performing funds deviate more from their peers they move closer to an immunizing portfolio. In other words, a poorly-performing pension fund reduces the hedging of its liabilities and gambles by choosing a riskier portfolio more typical of its peers.

To explore whether our results are sensitive to using tracking error as a measure of a fund’s deviation from an immunization benchmark, we consider three other dependent variables that proxy for deviations from a fixed-income immunization strategy. These are a fund’s allocation to: equities; U.S. equities; or all asset classes except fixed income. The results using these alternative dependent variables are shown in columns 11 to 13 of Table 3.6. Compared to using tracking error benchmarked on nominal or real liabilities, the qualitative results are unchanged.
3.6 Conclusion

This paper introduced a model of public pension fund asset allocation that can be a normative guide for choosing pension investments or a positive theory of pension managerial behavior. It concludes that hedging risk from changes in the market value of the pension fund’s liabilities is likely to be the socially optimal policy. However, career concerns of pension fund managers may conflict with this objective. Our empirical results seem consistent with such managerial agency behavior. We find that a public pension fund’s Board of Trustees and staff tend to allocate assets based on the performance of peer pension funds rather than with the aim of immunizing the plan’s liabilities.

Our empirical analysis of 125 state pension plans over the 2000-2009 period finds that a fund tends to take more asset-liability “tracking error” risk following declines in its relative performance. Tracking error volatility also is higher for pension funds that select a high rate with which to discount their liabilities and pension funds that have a greater proportion of participants on their Boards of Trustees.

The portfolio choices of public pension plans that deviate substantially from the liability immunizing strategies may be encouraged by opaque and misleading accounting standards that are divorced from finance theory. Such standards may lead public plans to follow their “traditional” investment strategies of choosing investments with little regard to their true liability risks. The pension fund asset-liability mismatches resulting from these strategies pose a potential burden to taxpayers that will be realized then economic conditions decline and when losses are most difficult to bear.
References


85


