THESIS DECLARATION

The undersigned

SURNAME	ABDESAKEN
NAME	GERALD MEHRAM
PhD Registration Number	1287598

Thesis title:

THREE ESSAYS IN EMPIRICAL FINANCE

PhD in	FINANCE
Cycle	XXIII
Candidate's tutor	GUIDOLIN, MASSIMO
Year of discussion	2013

DECLARES

Under his responsibility:

1) that, according to the President's decree of 28.12.2000, No. 445, mendacious declarations, falsifying records and the use of false records are punishable under the penal code and special laws, should any of these hypotheses prove true, all benefits included in this declaration and those of the temporary embargo are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree of 30th April 1999 protocol no. 224/1999, to keep copy of the thesis on deposit at the Biblioteche Nazionali Centrali di Roma e Firenze, where consultation is permitted, unless there is a temporary embargo in order to protect the rights of external bodies and industrial/commercial exploitation of the thesis;

3) that the Servizio Biblioteca Bocconi will file the thesis in its 'Archivio istituzionale ad accesso aperto' and will permit on-line consultation of the complete text (except in cases of a temporary embargo);

4) that in order keep the thesis on file at Biblioteca Bocconi, the University requires

that the thesis be delivered by the candidate to Società NORMADEC (acting on behalf of the University) by online procedure the contents of which must be unalterable and that NORMADEC will indicate in each footnote the following information:

- thesis *Three Essays in Empirical Finance*
- by Abdesaken, Gerald
- discussed at Università Commerciale Luigi Bocconi – Milano in 2013
- the thesis is protected by the regulations governing copyright (law of 22 April 1941, no. 633 and successive modifications). The exception is the right of Università Commerciale Luigi Bocconi to reproduce the same for research and teaching purposes, quoting the source;

5) that the copy of the thesis deposited with NORMADEC by online procedure is identical to those handed in/sent to the Examiners and to any other copy deposited in the University offices on paper or electronic copy and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (law of 22 April 1941, no. 633 and successive integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal and shall be exempt from any requests or claims from third parties;

7) that the PhD thesis is not the result of work included in the regulations governing industrial property, it was not produced as part of projects financed by public or private bodies with restrictions on the diffusion of the results; it is not subject to patent or protection registrations, and therefore not subject to an embargo;

Date October 31, 2012

SURNAME	ABDESAKEN	NAME	GERALD MEHRAM

Three Essays in Empirical Finance

Gerald Abdesaken
gerald.abdesaken@phd.unibocconi.it

Thesis submitted to
Università Bocconi

in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in Finance

Department of Finance
Via Röntgen 1
I-20136 Milano, Italy

Submitted on October 31, 2012

für Oma † 19/02/2007

Acknowledgements

There are many people to whom I owe thanks for supporting me throughout my PhD. First and foremost I thank my advisor, Massimo Guidolin, for his advice and guidance. I also thank my tutor and mentor, Paolo Colla, who always provided me with kind words which motivated me to keep working hard throughout my PhD studies. Without Massimo and Paolo, I likely would not have reached this stage of my PhD.

Next, I would like to thank my friends and colleagues from the PhD School at Bocconi. Our great camaraderie helped make the transition to Italian life so much easier for me. Special thanks go to Domenico Tarzia and Daniele Siena, who helped me from the beginning with things like residence per-

mits, apartments, phone calls, and other administrative day to day activities I wasn't able to handle due to the language barrier. Moreover, special thanks go to Antonio Gargano and Yuanji Wen, with whom I shared office 2-C3-05 with for 4 long years (well, Antonio only for 2). There was never a dull moment, whether it was discussions about research, plastic bottles being thrown across the room, Antonio's lunch prep, Yuanji's misunderstanding of the English language or spontaneous outbursts of vulgarity from me and Antonio after unusually long moments of silence.

Moreover, the following deserve special shout-outs for various reasons (in random order): Roberto Steri, Nenad Kos, Mickael Pero, Damijan Cotic, Gunes Gokmen, Andreja Lenarcic, Josh Miller, Christina Procacci, Carlo Chiarella, Alessandro Gobbi, Martina Pasquini, Olaf de Groot, Felix Suntheim, Weiwei Yin, Paola Mascia, Angela Baldassarre, and Daniele Bianchi. There are undoubtedly other friends and colleagues who I have forgotten to thank. If you feel unjustly left out of

the list, I apologize. Please go ahead and add yourself here: [   ].

A huge heartfelt thanks go to Mom, Dad, and Adeline for their love, support and encouragement of my studies in Europe. I'd also like to thank my Aunt Farkhondeh and Uncle Carsten for giving me a home away from home for the past 8 years. Moreover, I owe a debt of gratitude to my grandmother, Oma. Oma pressed me from a young age regarding the importance of pursuing a PhD. My PhD dissertation is dedicated solely to her. It is amazing the influence a loved one can have even 5 and a half years after their passing.

Last, I'd like to thank Marta for being by my side and providing unconditional support during the final stages of my PhD.

Contents

Acknowledgements **iii**

Preface **1**

1 On the Precision of Public Information and Mutual Fund Performance **6**

1.1 Introduction 7

1.2 Theoretical Foundations 17

1.3 Dataset 27

1.3.1 Mutual Fund Data 27

1.3.2 Reliance on Public Information 30

1.4 Empirical Strategy 39

1.4.1 Empirical Predictions 41
1.5 Empirical Analysis and Results 45
1.5.1 Mutual Fund Performance and RPI . . . 48
1.5.2 Mutual Fund Performance and Managerial Reliance on Coefficients of Variation 52
1.6 Conclusions 56

2 Conflicts of Interest in Multi-Fund Management 61

2.1 Introduction 62
2.2 Data 69
2.2.1 Manager Names 69
2.2.2 Opposite Trading and Opp 74
2.2.3 Fund Characteristics 80
2.3 Empirical Strategy 87
2.3.1 Analysis and Results 88
2.3.2 Cross Multi-Fund Subsidization 89
2.3.3 Sorts on YTD Returns 97
2.3.4 Single-fund Managers 100
2.4 Conclusions 107

3 Propensity Score Matched Discretionary Accruals 110

3.1 Introduction 111

3.2 Literature Review 115

3.3 Model Specification 118

3.3.1 Sample Selection 121

3.3.2 Estimation of Discretionary Accruals . . 123

3.3.3 Summary Statistics of Discretionary Accruals 136

3.4 Specification Test 145

3.5 Power Test 151

3.6 Conclusions 164

Bibliography 167

Preface

My thesis comprises of three works in the general area of empirical finance. In the first two chapters of my dissertation, I explore various aspects of mutual fund managerial skill and agency issues in the delegated asset management industry. In my third work, my coauthors and I develop a new measure of discretionary accruals, with the aim of exploring its asset pricing implications in the framework of market-based accounting in future research. All three works explore various distortions evident in financial markets and financial intermediaries.

In Chapter 1, I show that mutual fund managers who adjust their portfolio holdings based on changes in the variation of analyst forecasts achieve significantly lower abnormal returns. Utilizing a partially revealing rational expectations equilibrium setup, I show that an uninformed (hence, unskilled) investor places greater weight on the public signal relative to an informed investor, given an increase in the precision of public information. Following Kacperczyk and Seru (2007), I formulate a new measure of managerial skill based on the standard deviation of analyst recommendations. This measure, dubbed RPI^{σ}, captures the sensitivity of changes in mutual fund portfolio holdings due to changes in the standard deviation of analyst forecasts. Statistical evidence shows that abnormal returns decrease in RPI^{σ}. The results are robust to various conventional performance benchmarks and alternative formulations of RPI^{σ} using analyst price targets, EPS forecasts and coefficients of variation.

Chapter 2 aims to investigate potential agency issues attributable to the existence of multitasking mutual fund managers, i.e. those who manage multiple funds simultaneously. The simultaneous management of multiple mutual funds has the potential to create distortions in delegated portfolio management misaligned with the interests of investors and regulators. Multi-fund managers cross subsidize their top performing funds at the expense of low performing funds. I proxy for cross fund subsidization by measuring the degree in which a fund manager engages in cross trading or opposite trading with other funds that he/she manages simultaneously during each quarter in the sample. Cross trading exhibits a strong positive relationship with alpha for a multi-manager's high performer and a strong negative relationship with alpha for a multi-manager's low performer. Moreover, multi-fund managers promote high year-to-date return funds without sacrificing the performance of low year-to-date funds.

In the final chapter of my thesis, we develop and test an empirical model of discretionary accruals that can control for the simultaneous correlation of multiple determinants with accruals. By estimating logistic regressions of earnings management proxies on conventional determinants of earnings manipulation, we ascertain the propensity to manage earnings for each firm-year in our sample. In order to estimate the discretionary component of accruals, firms are matched according to year, industry and its propensity score. Comparisons with similar models in tests of statistical specification and power indicate that the propensity score matched model further mitigates the probability of producing false positives in earnings management hypothesis testing when the null hypothesis is no earnings management.

Chapter 1

On the Precision of Public Information and Mutual Fund Performance

JEL Classification: G11, G12, G14

1.1 Introduction

Significant theoretical, empirical and methodological progress has been made during the past two decades towards understanding the cross section of mutual fund returns and pinpointing those factors that influence it. For example, works on the persistence of performance and abnormal performance indicators based on passive benchmarks are scattered throughout the economic discourse; see Carhart (1997), Pastor and Stambaugh (2002b), Pastor and Stambaugh (2002a), Bollen and Busse (2004) and Barras et al. (2010), or for conditional based measures Ferson and Schadt (1996) and Ferson and Khang (2002). Mutual fund performance has also been addressed with respect to holdings based measures, for example in Grinblatt and Titman (1993), Daniel et al. (1997) and Wermers (2003). However, surprisingly little work has been done towards understanding the determinants of mutual fund managerial skill. Given the large array of mutual funds available to a potential investor, one may ask how skilled fund managers can be distinguished from the unskilled without looking at conventional performance indicators, which may often simply reflect luck. An answer may lie in the relationship between information acquisition and skill itself. Those who are skilled should have the ability to accumulate private information regarding which stocks to buy or sell, and those who are unskilled may be restricted to the domain of information that is already available to the majority of market participants.

Consider, for example, the domain of public information available to analysts and fund managers regarding a particular stock, S. At any given point in time, analysts assign a recommendation to S which ranges from 1 to 5, corresponding to "strong buy", "buy", "hold", "sell", and "strong sell", respectively. Suppose security S has 2 analysts assigning recommendations for it during 2 time periods ($R_{a,t}$). At $t = 0$, analyst 1 recommends security S a "strong buy" ($R_{1,0} = 1$), and analyst 2 recommends it a "hold" ($R_{2,0} = 3$). At $t = 1$, recommendations change such that analyst 1 recommends security S a "hold" ($R_{1,1} = 3$), and analyst 2 recommends it a "strong sell" ($R_{2,1} = 5$). Consensus forecast changes from 2 to 4 from $t = 0$ to $t = 1$, and an unskilled investor's holdings of S decreases, whereas a skilled investor trades based on superior (private) information (Kacperczyk and Seru, 2007). Note that the standard deviation of the public signal (σ) is preserved. Figure 1.1 presents a graphical representation of this change in analyst recommendations.

Now, suppose that at $t = 0$, analyst 1 recommends security S a "strong buy" ($R_{1,0} = 1$), and analyst 2 recommends it a "hold" ($R_{2,0} = 3$). However, at $t = 1$, analyst 1 changes his recommendation to "buy" ($R_{1,1} = 2$), and analyst 2 revises his to "buy" as well ($R_{2,1} = 2$). The consensus mean remains unchanged at 2, but the standard deviation of the public signal collapses to 0, as shown in Figure 1.2. Do unskilled investors also make portfolio adjustments based on changes in the precision of the public signal? If so, what are the implications in terms of performance?

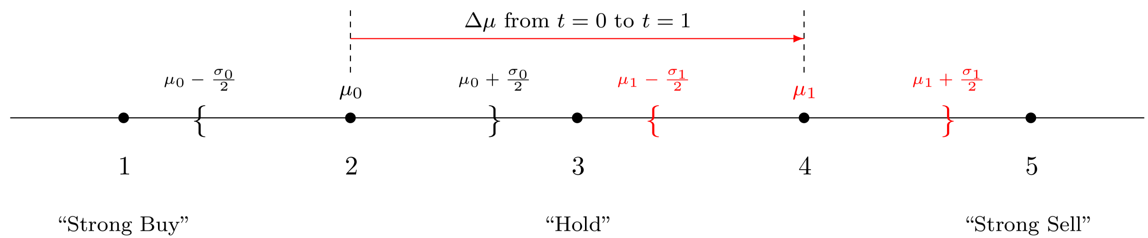

Figure 1.1: A change in analyst recommendations which preserves the standard deviation of the public signal

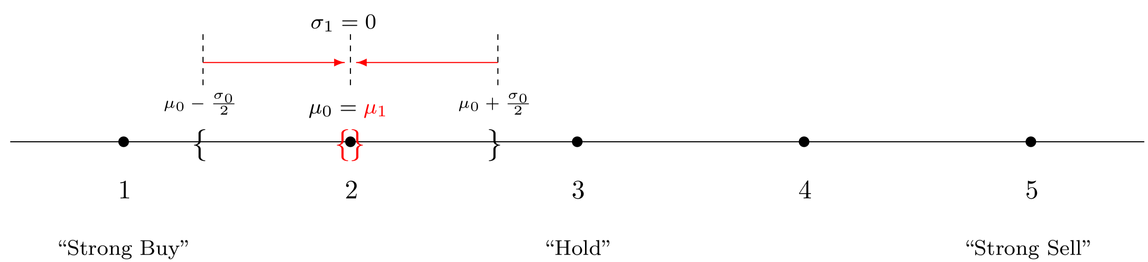

Figure 1.2: A mean preserving change in analyst recommendations

In this article, I introduce a new perspective on mutual fund managerial skill. I propose that fund managers who rebalance their portfolios based on changes in the precision of public information, i.e. the variation in stock analyst recommendations, perform worse than those who trade based on other (private) information. Until now, previous literature has concentrated on portfolio sensitivities with respect to consensus means of analyst forecasts (dubbed as "RPI" in Kacperczyk and Seru (2007), henceforth RPI^{μ}) . The motivation behind this idea is straightforward. Since consensus means have informational content for the unskilled fund manager, then it is reasonable to explore if the variation in analyst recommendations contains informational content as well. Also, a measure based on the precision of public information is worthwhile to compute even though we already have one based on consensus means, because if unskilled investors are trading based on changes in both first and second moments of analyst forecasts, then the effect of increases in portfolio sensitivities to those changes may be over- or underestimated when only considering RPI^{μ}.

The foundation of my empirical predictions stems from a simple theoretical implication based on the partially revealing rational expectations equilibrium model first proposed by Grossman and Stiglitz (1980) and revised by Hellwig (1980), O'Hara (2003) and Kacperczyk and Seru (2007). The model introduces a public signal observable by all, and a private signal observable only to the informed cohort, thus separating agents into informed and uninformed investors. In this context, skill is associated with being informed. It is then shown that the portfolio holdings of uninformed investors are more sensitive to changes in the precision of public information than those of the informed investor. Generally, uninformed investors place more weight on the public signal when the precision of public information increases, i.e. they increase their holdings of the risky asset more so than informed investors do. These theoretical implications provide the basis for my empirical hypotheses.

In order to approach this article from an empirical perspective, I make the following considerations. First, to test the relationship between fund performance indicators and the reliance on variation of public information, a reliable time series of portfolio sensitivities, financials and other characteristics for active strategy equity mutual funds is needed. Using the CRSP Survivorship Bias Free Mutual Funds database including over 1200 mutual funds from 2003 to 2009, I am able to compute portfolio sensitivities with respect to changes in the variation of public information by proposing a regression specification of changes in holdings on lagged changes in the standard deviation of analyst forecasts and defining RPI^{σ} as the R^2 of those regressions. Since these regressions are run using a panel according to each mutual fund's holdings at a particular point in time, the result is a unique time-series of portfolio sensitivities for each mutual fund at the quarterly time frequency. In order to make comparisons regarding performance, a benchmark is necessary, thus I use abnormal performance indicators from the unconditional and conditional CAPM and FF pricing models (Sharpe, 1964; Lintner, 1965; Fama and French, 1993; Ferson and Schadt, 1996).

Assuming that analysts accumulate and process all information available in the public domain, I proxy public information with summary figures communicated by those analysts. Thus, I implement analyst recommendations, price targets, and EPS forecasts as my relevant information variables and formulate separate time series of RPI^{σ} using each.

The main result of the empirical analysis is that RPI^{σ} is negatively related to passive benchmarks. Therefore, fund managers with higher levels of RPI^{σ} perform worse than those who trade based on other information. The results are robust to all three formulations of RPI^{σ} based on alternative information sets. As a second empirical exercise, I also formulate portfolio sensitivities based on contemporaneous reliance on means and variation of public information, which I dub

RPI^{CV}.[1] Results show that also fund managers with higher levels of RPI^{CV} perform worse than managers exhibiting a lower level of RPI^{CV}. It is important to note that the coefficients on portfolio sensitivities with respect to the precision of public information are smaller in magnitude than those based on consensus means, suggesting that previous literature was overestimating the effect of following information in the public domain, since second moments were left unconsidered.

This study provides several relevant contributions to the growing discourse in mutual fund performance. First, since RPI^{μ} overstates the negative effect that increased sensitivity to public information has on abnormal performance, RPI^{σ} and RPI^{CV} are more valid measures of managerial skill. Moreover, RPI^{σ} should also be interesting to those who monitor the investment performance of fund managers from a regulatory perpective. Most importantly, RPI^{σ} provides a relatively unique perspective on pinpointing outperforming mutual funds

[1] CV stands for coefficient of variation.

which has not yet been considered in this discourse.

The paper is organized as follows. In Section 2, I will present the theoretical foundations of the economic problem that I wish to analyze. Section 3 will give a thorough overview of the dataset, including a detailed description regarding the methodology behind the formulation of RPI^{σ}. Section 4 will outline my empirical strategy and formalize my empirical predictions based on the theoretical analysis. Section 5 will present the main results of the empirical analysis. Concluding thoughts and my agenda for further research will follow in Section 6.

1.2 Theoretical Foundations

I will begin by presenting a theoretical model where price plays the role of transmitting information obtained by informed agents to uninformed agents through a noisy rational expectations equilibrium setup, which is based on the framework first introduced by Grossman (1976) and revised by Grossman and Stiglitz (1980) to analyze informational market efficiency. The model stems from those of O'Hara (2003) and Kacperczyk and Seru (2007) (I adopt their notation where possible) and modified to approximate information aggregation (Hellwig, 1980), where the aim will be to analyze how changes in the precision of public information and the number of public signals influence portfolio decision making of uninformed agents relative to those who are informed. The model presented in this paper differs from others in that I allow for multiple public signals in the setup in order to approximate the presence of analysts in the market.

In this model, the timeline consists of two time periods, $t = 0$ and $t = 1$, where portfolios are formulated in the first time period and payoffs realized in the second. Agents trade two assets: a risk-free asset with price equal to 1 & payoff $R_f = 1$, and a risky asset A_1 with price p and random payoff $\tilde{u} \sim \mathcal{N}(\bar{u}, \rho_0^{-1})$ where ρ denotes precision, i.e. the reciprocal of the variance. The market is comprised of a finite number of agents,

N, with $n \in \{1, 2, \ldots, N\}$. Each agent is either informed or uninformed (hence, skilled or unskilled, respectively). The number of skilled agents is set equal to L where $0 < L$ & $N > L$ thus the fraction of skilled investors is denoted as $\mu = \frac{L}{N}$. The fraction of uninformed agents follows as $1 - \mu = \frac{N-L}{N}$. Each agent has an initial endowment, or wealth, denoted by $\bar{m}^n$. Preferences are homogeneous across agents in that all agents are constant absolute risk aversion (CARA) utility maximizers with common coefficient of risk aversion $\lambda > 0$.[2] In order to ensure a noisy rational expectations equilibrium (prices are only partially revealing), a random per capita supply shock t is introduced with $t \sim \mathcal{N}(\bar{t}, \eta^{-1})$.

The information structure of this economy consists of two types of signals. First, the private signal is observable solely to informed investors and is denoted by s_L, where $s_L = \bar{u} + \varepsilon_L$ and $\varepsilon_L \sim \mathcal{N}(0, \rho_L^{-1})$. There are I public signals denoted by s_i with $i = 1, 2, \ldots, I$ observable to informed and unin-

[2]The CARA assumption guarantees that optimal demand will not depend on an agent's wealth.

formed investors, which approximates the presence of analysts. The public signal is determined as $s_i = \bar{u} + \varepsilon_i$, where $\varepsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \rho_N^{-1})$ $\forall i \in I$. Note that the public signals and the private signal share the same fundamental value, $\bar{u}$. Additionally, an orthogonality assumption between ε_L, ε_i, $\bar{u}$ and t is imposed, that is $\varepsilon_L \perp \varepsilon_i \perp \tilde{u} \perp t$ $\forall i \in I$, and the distribution of each signal is assumed to be common knowledge among all market participants. The sample mean of the public signals is given as $\bar{s} = \bar{u} + \frac{1}{I}\sum_{i=1}^{I} \varepsilon_i$ where $\bar{s} \sim \mathcal{N}(\bar{u}, \rho_0^{-1} + \rho_{\bar{N}}^{-1})$ and $\rho_{\bar{N}}^{-1} \equiv \rho_N^{-1}/I$. Following Grossman (1976) and Hellwig (1980), $\bar{s}$ is a sufficient statistic for $\{s_1, s_2, \ldots, s_N\}$, and observing $\bar{s}$ leads to the same conditional distribution of $\tilde{u}$ as the observable vector of public signals. Given the supply, endowments, preferences, and information structure of the economy, the market clearing price of the risky asset is endogenously determined.

The terminal wealth of investor n can be expressed as

$$\tilde{w}^n = \bar{m}^n + (\tilde{u} - p)\alpha^n \tag{1.1}$$

where α^n denotes agent n's holdings of the risky asset. Thus, equation (1.1) represents the initial wealth, $\bar{m}^n$, plus any subsequent capital gains from holding the risky asset, $(\tilde{u} - p)\alpha^n$. Given this constraint, each investor maximizes his expected utility by choosing his optimal holdings of the risky asset, where the objective function follows the conventional mean variance representation, given by $E\left[\tilde{w}^n\right] - \frac{\lambda}{2}Var\left[\tilde{w}^n\right]$. Thus, each agent's optimal demand of the risky asset is given by

$$x^n = \frac{\bar{u}^n - p}{\lambda(\rho^n)^{-1}} \tag{1.2}$$

where $\bar{u}^n$ and ρ^n denote the conditional mean and conditional precision of $\tilde{u}$ respectively, given the information endowment of each agent.

In order to induce learning by observing the market price, uninformed investors conjecture the following price functional

$$p = a\bar{u} + bs_L + c\bar{s} - dt + e\bar{t} \tag{1.3}$$

as in O'Hara (2003). I then define

$$\theta \equiv \frac{p - a\bar{u} - c\bar{s} + \bar{t}(d-e)}{b} = s_L - \frac{d}{b}(t - \bar{t}). \tag{1.4}$$

Hence, θ is a sufficient statistic for s_L and t. Moreover, note that $\theta \sim \mathcal{N}(\bar{u}, \rho_\theta^{-1})$ and $\rho_\theta = \left[(\frac{d}{b})^2 \frac{1}{\eta} + \frac{1}{\rho_L}\right]^{-1}$. Given the price, agents update their posterior distribution of $\tilde{u}$ and revise their demand of the risky asset according to (1.2).[3] Prices are endogenously determined assuming that the market clears, i.e.

[3] The conditional moments of $\tilde{u}$ are derived using the multidimensional projection theorem and as perceived by informed investors the distribution of the risky asset is equal to

$$\tilde{u}|s_L, \bar{s} \sim \mathcal{N}(\frac{\bar{u}\rho_0 + \rho_{\bar{N}}\bar{s} + \rho_L s_L}{\rho_L + \rho_{\bar{N}} + \rho_0}, \rho_L + \rho_{\bar{N}} + \rho_0)$$

while for uninformed investors

$$\tilde{u}|\theta, \bar{s} \sim \mathcal{N}(\frac{\bar{u}\rho_0 + \rho_{\bar{N}}\bar{s} + \rho_\theta \theta}{\rho_\theta + \rho_{\bar{N}} + \rho_0}, \rho_\theta + \rho_{\bar{N}} + \rho_0).$$

the per capita demand for the risky asset equals the per capita supply, or $\mu x_I^n + (1-\mu)x_U^n = t$. I find that there exists a noisy rational expectations equilibrium since the conjectured linear price functional (1.3) is verified with coefficients $a = \frac{\rho_0}{\gamma}$, $b = \frac{\mu\rho_L + (1-\mu)\rho_0}{\gamma}$, $c = \frac{\rho_{\bar{N}}}{\gamma}$, $d = \frac{\lambda(1+\frac{(1-\mu)\rho_\theta}{\mu\rho_L})}{\gamma}$, $e = \frac{(1-\mu)\rho_\theta\lambda}{\mu\rho_L\gamma}$, where $\gamma = \rho_0 + \rho_{\bar{N}} + (1-\mu)\rho_\theta$. The difference between individual demands of the informed and the uninformed investors is given by

$$\Delta \equiv x_I - x_U = \frac{s_L(\rho_L - \rho_\theta) + p(\rho_\theta - \rho_L) + \frac{\rho_\theta\lambda}{\mu\rho_L}(t - \bar{t})}{\lambda}. \quad (1.5)$$

In order to make inferences regarding the sensitivity of the informed investor's portfolio to changes in the public signals (as compared to the uninformed investor's), the partial derivative of (1.5) with respect to $\bar{s}$ is taken and yields

$$\frac{\delta\Delta}{\delta\bar{s}} = \frac{\rho_{\bar{N}} \overbrace{(\rho_\theta - \rho_L)}^{(-)}}{\underbrace{\gamma\lambda}_{(+)}} < 0. \tag{1.6}$$

The result from equation (1.6) highlights that informed investors react less to changes in the public signal than uninformed investors do. Thus, when $\bar{s}$ increases, uninformed investors revise their portfolios by increasing their holdings of the risky asset more so than the informed investors do. Likewise, when $\bar{s}$ decreases, uninformed agents compensate by lowering their holdings of the risky asset more so than the informed investor does. This result is intuitive; the skilled investor also trades based on his privately obtained signal, thus the uninformed investor places greater weight on changes in the public signals than an informed investor would.

Taking the second derivative of (1.5) with respect to $\bar{s}$ and $\rho_{\bar{N}}$ yields

$$\frac{\delta^2 \Delta}{\delta \bar{s} \delta \rho_{\bar{N}}} = \frac{\overbrace{(\rho_\theta - \rho_L)}^{(-)} \overbrace{(\rho_0 + (1-\mu)\rho_\theta)}^{(+)}}{\underbrace{\gamma^2 \lambda}_{(+)}} < 0. \tag{1.7}$$

The result in equation (1.7) tells us that relative to uninformed investors, informed investors place less weight on the public signals, $\bar{s}$, given an increase in the precision of public information, $\rho_{\bar{N}}$. Hence for any given level of $\bar{s}$, uninformed investors increase their holdings of the risky asset more so than an informed investor would when the public signals become finer. The converse is true given a decrease in the precision of public information. Most importantly, the precision of public information has more informational content for the uninformed investor than it does for the informed investor, i.e. the uninformed investor reacts to changes in the variation of the public signal more than the informed investor does. Similar results are obtained from other rational expectation equilibrium models with related asymmetric information setups, such as Brennan and Cao (1997) and Feng and Seasholes (2004).

Given the presence of multiple public signals, the number of analysts could also have implications in terms of information aggregation. In order to understand the effect that the amount of analysts has on the demands of the informed and the uninformed, I take the second derivative of (1.5) with respect to $\bar{s}$ and the number of analysts, I, which yields

$$\frac{\delta \Delta}{\delta \bar{s} \delta I} = \frac{\rho_N \overbrace{(\rho_\theta - \rho_L)}^{(-)} \overbrace{(\rho_0 + (1-\mu)\rho_\theta)}^{(+)}}{\underbrace{\gamma^2 \lambda}_{(+)}} < 0. \tag{1.8}$$

Equation (1.8) tells us that informed investors place less weight on the public signals when the number of analysts increase. Thus, the public signals have higher informational content for uninformed investors when there are more analysts covering the risky asset. Considering the comparative statics presented in equations (1.6), (1.7) and (1.8), I expect the arrival of precise public information to induce uninformed investors to adjust their holdings of the risky asset more aggresively in contrast to informed investors, and I expect this effect to be more pronounced when the number of analysts increases. Therefore, I expect any measure of portfolio sensitivity with respect to changes in variation of information in the public domain and the number of analysts (which I denote with RPI^{σ}, previously coined as "reliance on public information") to proxy for managerial skill, and those managers who demonstrate a high level of RPI^{σ} should perform worse than those who demonstrate low levels of RPI^{σ}.

1.3 Dataset

1.3.1 Mutual Fund Data

Data are accumulated and aggregated from various sources. Mutual fund details and portfolio holdings are obtained from the CRSP Survivorship Bias Free Mutual Fund Database, which contains data regarding fund size, fees, portfolio holdings, objective and performance. Portfolio holdings include both voluntary reports and mandatory SEC filings at the quarterly time frequency, and are available from March 2003 until September 2009. Stock prices are accessed from the CRSP stock files and matched to portfolio holdings.

Analyst forecasts are obtained from the I/B/E/S database at the monthly time frequency and matched with portfolio holdings. Forecasts include analyst recommendations (ranging from 1 to 5 discretely, with 1 corresponding to "strong buy" and 5 to "strong sell"), earnings per share (EPS) forecasts and price targets. Notice that a positive shock to expected future cash flows for any given stock leads to an increase in EPS forecasts and price targets for that security, whereas it has the opposite effect on the value reported under recommendations (positive shocks lead to decreases in consensus analyst recommendations).

For each mutual fund in the sample, unconditional and conditional alphas are computed using a 35-month window following Carhart (1997) and Ferson and Schadt (1996). Risk factors in the unconditional model have been acquired from Kenneth

French's website. Predetermined information variables[4] in the conditional model were acquired from the Federal Reserve.

Mutual funds are screened according to investment objective[5] before they are included in the final dataset. Since the focus of this study is on equity forecasts, I eliminate balanced and bond funds. I also exclude sector funds to avoid industry specific biases, and omit international funds in order to focus strictly on U.S. equity funds. Mutual funds that engage in passive investment strategies, such as index funds, have also been removed.[6] Last, I include multiple share classes only once in the sample. The final dataset includes 1220 equity mutual

[4]Information variables include lagged levels of the 1-month T-bill yield, lagged dividend yields of the CRSP value weighted NYSE and AMEX stock index (computed as the price level at $t-1$ divided by the previous 12 months of dividend payments), lagged levels of the corporate bond yield spread (Moody's BAA corporate bond yield minus Moody's AAA corporate bond yield), lagged levels of the constant maturity 10-year T-bond yield minus 3-month T-bill yields, and a dummy variable for January.

[5]Lipper objective codes are used to identify mutual fund investment strategies.

[6]Regarding mutual funds for which a passive investment strategy is not explicitly specified in the Lipper objective codes, I compute correlation coefficients between fund returns and returns on the S&P 500 stock index, and omit funds with correlations larger than 0.995.

funds with at least 5 consecutive quarters of data from 2003-2009.

1.3.2 Reliance on Public Information

In order to associate fund performance with a manager's reliance on public information, it is necessary to measure the degree in which a manager uses public information variables for portfolio allocation decisions. In order to do so, it is crucial to identify a suitable proxy for the public information domain. I use analyst recommendations, EPS forecasts and price targets as proxies for the public information variable, but in order for analyst forecasts to qualify as a suitable proxy, two assumptions are necessary. First, analysts must have accumulated and incorporated all publicly available information into their recommendations and forecasts. Second, unskilled managers must aggregate all available analyst recommendations in their allocation decisions, thus they do not follow specific analysts and concentrate on consensus forecasts. The latter assumption is reasonable, since the knowledge to follow a specific analyst who can outperform the others could be considered as "skillful."

The empirical methodology to measure the sensitivity of portfolio changes to changes in consensus recommendations directly follows Kacperczyk and Seru (2007), in that the following regression is estimated with OLS using stocks $i = 1, \ldots, n$ in the portfolio of each mutual fund m at each point in time t:

$$\%\Delta Hold_{i,m,t} = \beta_{0,t} + \sum_{p=1}^{4} \beta_{p,t} \Delta R^{\mu}_{i,t-p} + \varepsilon_{m,t} \quad i = 1, \ldots, n \tag{1.9}$$

where $\%\Delta Hold_{i,m,t}$ denotes the percentage change in holding i for mutual fund m between time $t-1$ and t, and $\Delta R^{\mu}_{i,t-p}$ represents the change in mean consensus recommendations for stock i from time $t-p-1$ to $t-p$ for $p = 1, 2, 3, 4$; hence the explanatory variable is lagged four times in equation (1.9). Since the dependent variable is in terms of percent change, I follow previous convention and set the increase to 100% for stocks that enter the portfolio at time t.[7] The reliance on public information using consensus means is then obtained as

$$RPI^{\mu}_{m,t-1} = 1 - \frac{\sigma^2(\varepsilon_{m,t})}{\sigma^2(\%\Delta Hold_{m,t})}. \tag{1.10}$$

Thus, $RPI^{\mu}_{m,t-1}$ is the unadjusted R^2 of (1.9). Since RPI^{μ} represents the amount of variation in changes of portfolio holdings explained by the statistical model, it also represents the sensitivity of portfolio adjustments with respect to changes in consensus means. Capturing RPI^{μ} in this way allows portfolio sensitivities to be accumulated regardless of the direction of trade, however Kacperczyk and Seru (2007) reject the hypothesis that mutual funds in the lowest deciles of RPI^{μ} follow the opposite direction of changes in consensus means, and fail to reject (with 99% confidence) that mutual funds in the highest deciles of RPI^{μ} decrease their holdings given an increase in

[7] Kacperczyk and Seru (2007) set the increase to 100% and find little difference using alternative benchmarks.

consensus recommendations. Hence, it is unlikely that fund managers pursue an "opposites" strategy.

In order to formulate RPI^{μ} using EPS forecasts and price targets, (1.9) requires a slight adjustment. Since changes in EPS forecasts and price targets are not comparable in magnitude across stocks, I take the percentage change instead of the change in levels as the independent variables in the first stage regression, that is:

$$\%\Delta Hold_{i,m,t} = \beta_{0,t} + \sum_{p=1}^{4} \beta_{p,t} \%\Delta F^{\mu}_{i,t-p} + \varepsilon_{m,t} \quad i = 1, \ldots, n \tag{1.11}$$

where $\%\Delta F^{\mu}_{i,t-p}$ denotes the percentage change in forecast (either EPS or price targets) from time $t-p-1$ until $t-p$. Such an adjustment is warranted, since an increase from \$1 to \$2 in earnings per share is not equivalent to an increase from \$50 to \$51, while when working with consensus recommendations the difference between a change from 1 to 2 and 4 to 5 are the same. This rationale applies for price targets as well.

Portfolio sensitivities with respect to changes in the variation of analyst forecasts are formulated in a similar fashion. The following regression is estimated using OLS as before:

$$\%\Delta Hold_{i,m,t} = \beta_{0,t} + \sum_{p=1}^{4} \beta_{p,t} \Delta R^{\sigma}_{i,t-p} + \sum_{p=1}^{3} \gamma_{p,t} \Delta A^{\mu}_{i,t-p} + \varepsilon_{m,t} \qquad i = 1, \ldots, n \tag{1.12}$$

where $\Delta R^{\sigma}_{i,t-p}$ denotes the cross sectional change in the standard deviation of analyst forecasts from time $t-p-1$ to $t-p$ and $\Delta A^{\mu}_{i,t-p}$ represents the average number of analysts from time $t-p-1$ to $t-p$. The reliance on consensus variation, or $RPI^{\sigma}_{m,t-1}$, is defined exactly as in (1.10), hence it is the R^2 of (1.12). The underlying assumption in this empirical setup is that portfolio managers may respond to changes in both cross sectional standard deviations and also the number of analysts, i.e. unskilled managers may put less weight on variation during the portfolio decision process when the amount of analysts is small.[8] Thus, those investors that respond to variation and/or analyst following will be captured either by the β_t's or the γ_t's of (1.12). The formulation of $RPI^{\sigma}_{m,t-1}$ using EPS forecasts and price targets also follows as before, whereby explanatory variables of differences in standard deviations are replaced by percentage changes.

In order to capture the contemporaneous reliance on both means and standard deviations, I compute coefficients of variation with reference to consensus means and standard deviations for each stock at each point in time.[9] Then I estimate

[8]The variation of recommendations for a stock with a large analyst following cannot be compared with the variation of recommendations for a stock with a relatively small analyst following.

[9]The coefficient of variation is defined as $\sigma_{i,t}/\mu_{i,t}$, where $\sigma_{i,t}$ denotes the standard deviation of analyst forecasts and $\mu_{i,t}$ denotes the consensus mean of analyst forecasts for each stock i at time t.

the following regression using OLS:

$$\%\Delta Hold_{i,m,t} = \beta_{0,t} + \sum_{p=1}^{4} \beta_{p,t} \Delta CV_{i,t-p} + \sum_{p=1}^{3} \gamma_{p,t} \Delta A^{\mu}_{i,t-p} + \varepsilon_{m,t}$$

$$i = 1, \ldots, n \tag{1.13}$$

where $\Delta CV_{i,t-p}$ denotes the change in coefficients of variation for consensus analyst recommendations from time $t-p-1$ to $t-p$. Equation (1.13) is also estimated using EPS forecasts and price targets, however $\Delta CV_{i,t-p}$ is replaced with $\%\Delta CV_{i,t-p}$ as before. The reliance on consensus means and standard deviations for each mutual fund at each point in time follows as the R^2 of (1.13) and is denoted as $RPI^{CV}_{m,t-1}$. For EPS forecasts and price targets, decreases in the variation of analyst forecasts and increases in consensus means are both "positives," and cause the coefficient of variation to decrease. For analyst recommendations, an increase in consensus recommendations is a "negative," since 1 corresponds to "strong buy" and 5 corresponds to "strong sell." Therefore, analyst recommendations are corrected when computing coefficients of variations by reversing the numerical values associated with each respective recommendation.[10]

The final dataset consists of a time series of RPI^{μ}, RPI^{σ} and RPI^{CV} formulated using analyst recommendations, price targets and EPS forecasts for each mutual fund in the sample. Summary statistics are provided in Table 1.1, where mutual fund data are separated into deciles according to each fund's average level of RPI^{σ} formulated with analyst recommendations. First, note that as RPI^{σ} increases, as does RPI^{μ} by construction. Moreover, funds with lower RPI^{σ} tend to be younger, adjust their portfolios less often, and charge lower fees. Also, there seems to be strong monotonic increases in TNA and turnover when moving from decile 1 to decile 10 in the table, indicating a positive correlation between RPI^{σ} and

[10] I adjust recommendations by subtracting the consensus mean from 6, thus "strong sell" consequently corresponds to 1 and "strong buy" to 5.

Table 1.1: Summary Statistics

Data are from March 2003 until September 2009. Each mutual fund is assigned to a decile portfolio based on their time series average of RPI_{σ}, where RPI^{μ} and RPI^{σ} represent portfolio sensitivities measured with respect to consensus means and standard deviations of analyst recommendations and are the R^2s of regression specifications (1.9) and (1.12), respectively. TNA represents the total net assets of the mutual fund for month t. Age is the age of the fund, in years. $TurnoverRatio$ is the turnover ratio of each mutual fund, lagged by one year. $ExpenseRatio$ represents the ratio of total investment paid by investors for the mutual fund's operating expenses in the most recently completed fiscal year, and $Returns$ is the monthly return of each mutual fund for month t.

Decile of $RPI\sigma$	Mean						
	$RPI\sigma$	$RPI\mu$	TNA (monthly)	lnAge	Turnover Ratio (%)	Expense Ratio (%)	Returns (monthly)
1	3.0	2.6	493.2	1.1	48.4	0.9	1.9
2	7.4	4.9	49.2	1.4	64.9	1.1	1.5
3	10.8	7.4	197.2	1.4	68.4	1.1	1.3
4	14.3	9.6	137.3	1.9	68.3	1.3	1.8
5	17.3	10.5	1,034.9	1.8	80.3	1.3	1.7
6	20.7	13.2	95.9	1.8	67.1	1.2	1.5
7	24.7	14.4	286.2	1.9	79.1	1.3	1.3
8	30.2	17.3	323.1	2.0	87.0	1.3	1.4
9	38.5	21.6	121,262.3	3.7	91.0	1.4	1.6
10	61.3	33.9	54,432.8	2.9	124.6	1.4	1.6
Total	22.8	13.5	17,814.8	2.0	77.8	1.2	1.6

these covariates. Given the correlation between these variables, a multivariate setting is appropriate in order to make inference regarding the relationship between RPI^{σ} and performance benchmarks.

1.4 Empirical Strategy

In order to show that RPI^{μ}, RPI^{σ} and RPI^{CV} are measures of mutual fund managerial skill, the relationship between RPI and traditional empirical benchmarks must be evaluated. In specific, I estimate the following regression:

$$\alpha_{m,t} = \beta_0 + \beta_1 RPI_{i,t-1} + \gamma' X_{i,t-1} + \varepsilon_{m,t} \tag{1.14}$$

where $\alpha_{m,t}$ denotes the fund-specific performance measurement at time t, $X_{i,t-1}$ a column vector consisting of fund characteristic control variables and γ' a row vector of coefficients. Notice that the specification in (1.14) is forward-looking. RPI has implications regarding future performance, since RPI is measured as the weight that lagged quarterly changes in analyst forecasts until $t-1$ have on portfolio adjustments made from time $t-1$ until time t,[11] and this convention is maintained throughout the empirical analysis. Four specifications for the performance measure $\alpha_{m,t}$ are considered. I follow Carhart (1997) and derive alpha as the intercept of the unconditional CAPM and Fama-French pricing equations using a 35 month rolling window of mutual fund returns, plus the error term at time t (Sharpe, 1964; Lintner, 1965; Fama and French, 1993). Conditional alphas are also computed as in Ferson and Schadt (1996) and Wermers (2003), whereby a vector of predetermined information variables are interacted with the market risk premium and supplemented to the unconditional CAPM and Fama-French specifications.[12] The vector of controls in-

[11] Although RPI is highly persistent with $AR(1)$ coefficients between 0.3 and 0.6 depending on the specification, the forward looking model highlights the timing implications of RPI. Nevertheless, the results remain consistent albeit weaker when contemporaneous variables are inserted.

[12] The vector of predetermined information variables is de-meaned and includes the lagged level of the 30 day annualized T-bill yield, the lagged dividend yield of the CRSP value-weighted NYSE/AMEX stock index,

cludes total net assets (TNA), age, expense ratio, turnover ratio and new money growth.

Regression (1.14) must be corrected for autocorrelation and heteroskedasticity in the panels. Thus, the panel corrected standard errors estimator with panel specific $AR(1)$ structures is used to estimate equation (1.14) with robust standard errors (Beck and Katz, 1995). Note that a correction for contemporaneous correlation is omitted, which is warranted when the panel dataset is highly unbalanced and contains a small time dimension and a relatively large cross sectional dimension (Pesaran, 2004).

1.4.1 Empirical Predictions

In Section 1.2, it was shown that uninformed investors place more weight on the public signal given an increase in the precision of information in the public domain, as opposed to in-

the lagged level of the constant maturity 10-year T-bond yield minus the 3 month T-bill yield, the lagged level of Moody's BAA corporate bond yield minus Moody's AAA corporate bond yield, and a dummy variable for January.

formed investors whose demand for the risky asset is less sensitive to changes in precision. It was also shown that uninformed investors will place more weight on the average signal when the analyst following increases. Given that RPI^{σ} provides a ranking of mutual funds from 0 (portfolios which are least sensitive to precision) to 1 (portfolios which are most sensitive to precision), I expect any formulation of RPI^{σ} to have a negative relationship with fund performance indicators.

> HYPOTHESIS 1: *RPI^{σ} is negatively related to fund performance indicators. Hence, a relationship between mutual fund managerial skill and a fund manager's reliance on the precision of public information and analyst following exists, and those fund managers who rely more heavily on the precision of public information perform worse than those who trade based on other (superior) information.*

Hypothesis 1 can be considered the null hypothesis in this research exercise. The alternative hypothesis would state that no relationship exists between RPI^{σ} and traditional fund performance indicators. That would imply that managers following changes in the precision of public information neither perform worse nor better than those who trade based on other information.

Given the findings of Kacperczyk and Seru (2007) and the implications derived in Section 1.2 regarding the relationship between the reliance on consensus means and mutual fund managerial skill, I would also expect similar results regarding the contemporaneous reliance on both consensus means and the precision of public information, since unskilled managers rely on both in the theoretical framework. Thus, a fund manager who makes portfolio decisions based on simultaneous changes in consensus means and variation in the consensus (hence, the coefficient of variation) should perform worse than those who do not. In this case, the relevant indicator RPI^{CV} should have a negative relationship with fund performance measures, and the effect on those measures should be similar in magnitude as RPI^{σ}.[13]

HYPOTHESIS 2: RPI^{CV} *is negatively related to fund performance indicators. Thus, a relationship between mutual fund managerial skill and a fund manager's reliance on the coefficient of variation of consensus forecasts exists, and those who rely on the coefficient of variation of consensus forecasts and analyst following perform worse than those who do not.*

The latter hypothesis postulates that both the mean and standard deviation of consensus forecasts simultaneously have informational content for the unskilled trader. The alternative

[13] RPI^{σ} and RPI^{CV} should have a similar effect in magnitude because unskilled mutual fund managers require the consensus mean in order to ascertain the implied directionality of their trades, and more precise signals simply determine the weight of the public signal on portfolio decision making for unskilled managers. Since RPI^{σ} is simply a measure of sensitivity between the variation of public information and portfolio allocations, directionality is embedded in it. RPI^{CV} is also embedded with directionality, however more formally with the inclusion of consensus means in its formulation.

hypothesis would assert that no relationship exists between fund performance and RPI^{CV}.

An interesting exercise will be to compare the magnitude in which fund performance decreases relative to RPI^{μ}, RPI^{σ} and RPI^{CV}.

1.5 Empirical Analysis and Results

Here I will show the main empirical results pertaining to the relationship between variation in information from the public domain and mutual fund performance. Subsection 1.5.1 presents evidence linking RPI^{σ} to managerial skill using various information sets, while subsection 1.5.2 will show that managers who rely on coefficients of variation of consensus analyst forecasts to make portfolio decisions perform worse than those who trade based on other information.

Table 1.2: Factor Based Regressions on RPI Formulated using Analyst Recommendations as the Information Set

This table presents the results of regressions based on equation (1.14) corrected for autocorrelation and heteroskedasticity using panel corrected standard errors (Beck and Katz, 1995). Data are from March 2003 until September 2009. The dependent variable in each regression specification is fund specific alpha, $\alpha_{m,t}$. Unconditional and conditional CAPM and Fama-French alphas are considered as performance benchmarks for the analysis (Sharpe, 1964; Fama and French, 1993; Ferson and Schadt, 1996). RPI^{μ}(REC) and RPI^{σ}(REC) represent portfolio sensitivities measured with respect to consensus means and standard deviations of analyst recommendations and are the R^2s of regression specifications (1.9) and (1.12), respectively. $lnTNA$ represents the natural logarithm of each mutual fund's total net assets, lagged one quarter. $lnAge$ is the natural logarithm of mutual fund age, lagged one quarter. $Expenses$ represents the ratio of total investment paid by investors for the mutual fund's operating expenses in the most recently completed fiscal year. $Turnover$ is the turnover ratio of each mutual fund, lagged by one year. NMG is new money growth lagged by one quarter, and is given by $NMG = \frac{TNA_{m,t} - TNA_{m,t-1}(1+R_{m,t})}{TNA_{m,t-1}}$ where $R_{m,t}$ represents fund specific return at time t.

Dependent Variable	Unconditional Alpha				Conditional Alpha			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	α_{CAPM}	α_{FF}	α_{CAPM}	α_{FF}	α_{CAPM}	α_{FF}	α_{CAPM}	α_{FF}
RPI^{μ}_{t-1}(REC)	-0.364***	-0.172*			-0.261***	-0.190**		
	(0.104)	(0.0884)			(0.0959)	(0.0851)		
RPI^{σ}_{t-1}(REC)			-0.129*	-0.131**			-0.134**	-0.155***
			(0.0738)	(0.0619)			(0.0683)	(0.0600)
$lnTNA_{t-1}$	-0.505	0.357	-0.441	0.348	-0.368	0.558	-0.336	0.551
	(1.038)	(0.842)	(1.043)	(0.842)	(0.973)	(0.805)	(0.975)	(0.806)
$lnAge_{t-1}$	-3.048	-0.876	-3.307	-0.896	-4.146**	-1.245	-4.274**	-1.235
	(2.168)	(1.670)	(2.169)	(1.669)	(2.076)	(1.656)	(2.079)	(1.654)
$Expenses_{t-1}$(%)	13.27	7.103	13.06	7.378	7.765	6.482	7.832	6.863
	(10.89)	(9.346)	(10.92)	(9.363)	(10.32)	(9.625)	(10.34)	(9.641)
$Turnover_{t-1}$(%)	0.181***	0.157***	0.178***	0.160***	0.101***	0.117***	0.101***	0.120***
	(0.0262)	(0.0206)	(0.0263)	(0.0207)	(0.0233)	(0.0201)	(0.0234)	(0.0201)
NMG_{t-1}	-0.0340**	-0.0204	-0.0337**	-0.0205	-0.0258*	-0.0140	-0.0257*	-0.0141
	(0.0170)	(0.0207)	(0.0170)	(0.0208)	(0.0153)	(0.0205)	(0.0154)	(0.0207)
N	11624	11624	11624	11624	11624	11624	11624	11624
Panels	1219	1219	1219	1219	1219	1219	1219	1219
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

1.5.1 Mutual Fund Performance and RPI

Table 1.2 presents the results from panel corrected standard error passive factor-based regressions of equation (1.14) using analyst recommendations (denoted with REC) to formulate RPI^{μ} and RPI^{σ}. Quarterly time dummies are included in each regression specification to control for time fixed effects.[14] Models (1) and (2) verify previous literature using unconditional CAPM and Fama-French alphas as dependent variables, and models (5) and (6) verify it using conditional alphas. The coefficient on RPI^{μ}_{t-1} is negative and statistically significant with similar magnitude as in Kacperczyk and Seru (2007) in all four specifications. Models (3), (4), (7) and (8) include RPI^{σ}_{t-1}, the coefficient of which is also negative and statistically significant with respect to conditional and unconditional CAPM and Fama-French alphas. The coefficient on RPI^{μ}_{t-1} varies between -.37% and -.19%, while the coefficient

[14]Each regression model was also estimated using fund specific fixed effects, the output of which is omitted for brevity since the results remain quantitatively similar to the ones presented here, but are available to the reader upon request.

on RPI^{σ}_{t-1}, however, is much more stable between -.12% and -.15%. RPI^{σ}_{t-1} seems to be a more robust indicator of managerial skill with respect to passive benchmarks than RPI^{μ}_{t-1}. Generally, the results corroborate my hypothesis that RPI^{σ} is associated with mutual fund managerial skill when considering passive factor-based portfolios as benchmarks.

For robustness, the analysis was also repeated with alternative formulations of RPI^{σ} using price targets and EPS forecasts instead of analyst recommendations to proxy for information in the public domain. Table 1.3 presents the results of panel corrected standard error estimations of equation (1.14) using unconditional and conditional Fama-French alphas as dependent variables and lagged levels of RPI^{σ} based on price targets and EPS forecasts as independent variables (denoted with PTG and EPS, respectively). The coefficient on RPI^{σ}_{t-1} remains negative in all four model specifications and statistically significant at the 10% level in the unconditional Fama-French specifications and at the 5% level in the conditional

Fama-French specifications.

Table 1.3: Factor Based Regressions with Alternative Formulations of RPI^{σ} using Price Targets and EPS Forecasts
This table presents the results of regressions based on equation (1.14) corrected for autocorrelation and heteroskedasticity using panel corrected standard errors (Beck and Katz, 1995). Data are from March 2003 until September 2009. The dependent variable in each regression specification is fund specific alpha, $\alpha_{m,t}$. Unconditional and conditional Fama-French alphas are considered as performance benchmarks for the analysis (Fama and French, 1993; Ferson and Schadt, 1996). RPI^{σ}(PTG) and RPI^{σ}(EPS) represent portfolio sensitivities measured with respect to standard deviations of analyst price targets and EPS forecasts and are the R^2s of regression specification (1.12) with the explanatory variables of differences in standard deviations replaced by percentage changes. $lnTNA$ represents the natural logarithm of each mutual fund's total net assets, lagged one quarter. $lnAge$ is the natural logarithm of mutual fund age, lagged one quarter. $Expenses$ represents the ratio of total investment paid by investors for the mutual fund's operating expenses in the most recently completed fiscal year. $Turnover$ is the turnover ratio of each mutual fund, lagged by one year. NMG is new money growth lagged by one quarter, and is given by $NMG = \frac{TNA_{m,t} - TNA_{m,t-1}(1+R_{m,t})}{TNA_{m,t-1}}$ where $R_{m,t}$ represents fund specific return at time t.

Dependent Variable:	Unconditional Alpha		Conditional Alpha	
	(1)	(2)	(3)	(4)
	α_{FF}	α_{FF}	α_{FF}	α_{FF}
RPI^{σ}_{t-1}(PTG)	-0.103*		-0.149**	
	(0.0605)		(0.0580)	
RPI^{σ}_{t-1}(EPS)		-0.108*		-0.125**
		(0.0594)		(0.0578)
$lnTNA_{t-1}$	0.307	0.392	0.437	0.561
	(0.843)	(0.843)	(0.809)	(0.808)
$lnAge_{t-1}$	-0.984	-0.986	-1.073	-1.277
	(1.667)	(1.674)	(1.645)	(1.662)
$Expenses_{t-1}$(%)	6.820	7.458	6.505	6.779
	(9.510)	(9.408)	(9.599)	(9.687)
$Turnover_{t-1}$(%)	0.162***	0.160***	0.126***	0.119***
	(0.0209)	(0.0207)	(0.0203)	(0.0201)
NMG_{t-1}	-0.0205	-0.0204	-0.0142	-0.0140
	(0.0207)	(0.0207)	(0.0205)	(0.0205)
N	11618	11616	11618	11616
Panels	1218	1220	1218	1220
Quarterly FE	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

1.5.2 Mutual Fund Performance and Managerial Reliance on Coefficients of Variation

In order to test Hypothesis 2 empirically, equation (1.14) is estimated using RPI formulated with consensus means and standard deviations (hence the coefficient of variation), denoted with RPI^{CV}. Table 1.4 presents the results from passive factor-based regressions of equation (1.14) using the PCSE estimator as before and RPI_{t-1}^{CV} as the RPI measure. Unconditional Fama-French alphas are used as the dependent variable in model specifications (1), (2) and (3), while conditional Fama-French alphas are used in models (4), (5) and (6). Alternative information sets are also considered in the formulation of RPI_{t-1}^{CV}, with columns (1) & (4) pertaining to RPI_{t-1}^{CV} formulated with analyst recommendations (REC), columns (2) & (5) to EPS forecasts (EPS), and (3) & (6) to price targets (PTG). Results show that the coefficient on RPI_{t-1}^{CV} is negative for all six specifications. When considering unconditional Fama French alphas as the performance benchmark, the coefficient on RPI_{t-1}^{CV} formulated with analyst recommendations and price targets are statistically significant at the 10% level. Furthermore, the performance of RPI_{t-1}^{CV} as a predictor of fund performance improves when considering conditional Fama French alphas as the dependent variable, with all three formulations improving in statistical significance, where the coefficient on RPI_{t-1}^{CV} (PTG) is also significant at the 1% level. Moreover, when comparing the results of Table 1.4 with those of Tables 1.2 and 1.3, the magnitude of changes in RPI_{t-1}^{CV} on performance benchmarks is very similar to those of RPI_{t-1}^{σ}. Overall, the results from Table 1.4 support my hypothesis that RPI^{CV} is a sufficient proxy for mutual fund managerial skill. The results are also robust to RPI^{CV} formulated with alternative information sets and passive benchmarks.

Table 1.4: Factor Based Regressions using RPI^{σ} Formulated with Coefficients of Variation

This table presents the results of regressions based on equation (1.14) corrected for autocorrelation and heteroskedasticity using panel corrected standard errors (Beck and Katz, 1995). Data are from March 2003 until September 2009. The dependent variable in each regression specification is fund specific alpha, $\alpha_{m,t}$. Unconditional and conditional Fama-French alphas are considered as performance benchmarks for the analysis (Fama and French, 1993; Ferson and Schadt, 1996). RPI^{CV}(REC), RPI^{CV}(PTG) and RPI^{CV}(EPS) represent portfolio sensitivities measured with respect to the coefficients of variation using analyst recommendations, analyst price targets and EPS forecasts, respectively, and are the R^2s of regression specification (1.13). $lnTNA$ represents the natural logarithm of each mutual fund's total net assets, lagged one quarter. $lnAge$ is the natural logarithm of mutual fund age, lagged one quarter. $Expenses$ represents the ratio of total investment paid by investors for the mutual fund's operating expenses in the most recently completed fiscal year. $Turnover$ is the turnover ratio of each mutual fund, lagged by one year. NMG is new money growth lagged by one quarter, and is given by $NMG = \frac{TNA_{m,t} - TNA_{m,t-1}(1+R_{m,t})}{TNA_{m,t-1}}$ where $R_{m,t}$ represents fund specific return at time t.

Dependent Variable:	Unconditional Alpha			Conditional Alpha		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{FF}	α_{FF}	α_{FF}	α_{FF}	α_{FF}	α_{FF}
RPI^{CV}_{t-1}(REC)	-0.128**			-0.158***		
	(0.0605)			(0.0588)		
RPI^{CV}_{t-1}(EPS)		-0.0754			-0.101*	
		(0.0582)			(0.0564)	
RPI^{CV}_{t-1}(PTG)			-0.108*			-0.147***
			(0.0592)			(0.0567)
$lnTNA_{t-1}$	0.353	0.404	0.377	0.551	0.570	0.540
	(0.841)	(0.846)	(0.841)	(0.804)	(0.810)	(0.808)
$lnAge_{t-1}$	-0.885	-1.051	-1.029	-1.206	-1.323	-1.177
	(1.668)	(1.675)	(1.667)	(1.653)	(1.662)	(1.645)
$Expenses_{t-1}$(%)	7.359	7.267	7.030	6.867	6.624	6.700
	(9.360)	(9.392)	(9.482)	(9.632)	(9.675)	(9.577)
$Turnover_{t-1}$(%)	0.160***	0.158***	0.162***	0.121***	0.118***	0.125***
	(0.0207)	(0.0208)	(0.0209)	(0.0201)	(0.0202)	(0.0203)
NMG_{t-1}	-0.0204	-0.0204	-0.0206	-0.0140	-0.0139	-0.0142
	(0.0208)	(0.0208)	(0.0207)	(0.0207)	(0.0206)	(0.0205)
N	11624	11616	11618	11624	11616	11618
Panels	1219	1220	1218	1219	1220	1218
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

1.6 Conclusions

In this paper, I have argued that fund managers who make portfolio decisions based on the precision of public information perform worse than those who trade based on superior (or private) information, hence those who trade based on the consensus mean and variation of information in the public domain are unskilled. In order to support my claim, I first presented a noisy rational expectations equilibrium model which featured multiple types of signals, and included agents who were either informed or uninformed. The main results of the model showed that portfolio holdings of uninformed investors were more sensitive to changes in the precision of public information than portfolio holdings of the informed investors. The effect is also more pronounced when there is a larger analyst following of the risky asset. This measure of portfolio sensitivity to precision, which I have dubbed as RPI^{σ}, should then be a suitable proxy for mutual fund managerial skill. Managers who feature a high level of RPI^{σ} should perform worse than those who feature relatively low levels.

Consequently, by using a detailed database on U.S. mutual fund holdings, financials and fund characteristics, I was able to construct a measure of RPI^{σ} for each mutual fund on a quarterly basis and found that a negative and statistically significant relationship exists between RPI^{σ} and traditional passive factor-based performance indicators. These results were also robust to alternative information sets produced by analysts. Thus, empirical evidence supports the hypothesis that RPI^{σ} is a measure of mutual fund managerial skill.

This article contributes to the growing literature on the cross section of mutual fund returns. First, RPI^{σ} and RPI^{CV} should complement RPI^{μ} as a proxy for managerial skill. Besides the obvious interest to those who choose to invest in mutual funds, RPI^{σ} and RPI^{CV} would also be interesting to those who monitor the investment performance of mutual fund managers, especially with the growing pressure to regulate financial markets after the subprime crisis. RPI^{σ} also provides a relatively unique perspective on mutual fund performance which has not yet been considered in this discourse, in that the second moments of analyst forecasts and recommendations can also have informational content for portfolio managers.

Interesting questions arise as I consider my agenda for further research. First, can anything be said regarding the dominance of the variation of mutual fund performance over consensus means? Results from Table 1.4 suggest that those who rely on both aspects of public information perform better than those who trade simply based on consensus means. Furthermore, one can take the reverse perspective on measuring mutual fund managerial skill. Skilled portfolio managers should be able to process information faster than analysts and predict the future performance of stocks more quickly and more accurately than analysts do. A similar measure to RPI could be formulated using regressions with changes in recommendations as the dependent variable and lagged changes in portfolio holdings as the independent variable, and extracting the R^2 as a reverse measure of managerial skill. Such a measure would be interesting to compare with RPI^{σ}.

Chapter 2

Conflicts of Interest in Multi-Fund Management

JEL Classification: G11, G23, G30

2.1 Introduction

A commonly overlooked feature of the delegated asset management industry is the prevalence of mutual fund managers who manage multiple funds simultaneously. Multitasking fund managers oversaw a staggering 48% of the mutual funds listed in the CRSP Survivorship Bias Free Mutual Fund database since 2003. Naturally, portfolio managers do not begin their careers as multi-fund managers. They are employed by mutual fund families and compete with one another to climb the corporate ladder. Although younger managers who perform better are more likely to be promoted by their respective fund families, they do so by holding lower levels of unsystematic risk in their portfolios and deviate less from the mean sector weightings of their respective objective groups (Chevalier and Ellison, 1997). Therefore, when star talent is acknowledged, fund families maintain an interest in retaining the best managerial talent by rewarding top performers with higher compensation and more funds. Moreover, there are potential benefits to the fund family from the reorganization of single-fund managers to multi-fund managers. For example, there could be various returns to scale through publicity and fund advertising activities.

Fund families also maintain an interest in promoting high-value funds to increase the revenues of the family. High-value funds charge high fees or attract larger net inflows from superior prior performance. (Gaspar et al., 2006) show that family affiliated mutual funds coordinate their actions in benefit of the family to the point where fiduciary duty to investors is questioned. This activity, dubbed "cross-fund subsidization", is conducted via family-level coordinated mechanisms such as preferential treatment in IPO allocation, or via managerial-level coordinated mechanisms such as cross-trading or opposite-trading.

Regardless of the interests of the family, mutual fund managers compete against each other in tournaments since incentive clauses in compensation contracts are linked to fund performance (Brown et al., 1996; Kempf and Ruenzi, 2008). In the absence of strategic interactions (i.e. in large fund families), "loser" funds tend to increase their levels of risk towards the end of the calendar year in order to increase their chances of jumping in rank. These findings are based on the premonition that managers compete against each other rather than coordinate in the interest of the family. Given the presence of cross-fund subsidization in the mutual fund industry mentioned above, an interesting research question arises: Can opposite trading or cross trading as a mechanism for cross fund subsidization exist in the presence of the fund family tournament?

In this article, I argue that coordinated trades between funds managed by multitasking managers are used as a mechanism for cross fund subsidization. Multi-fund managers circumvent the cooperation problem with other fund managers by coordinating their trades with other funds that he or she manages. The intuition behind this claim is as follows. The relationship between performance and compensation is convex, i.e. managers of funds that perform well are rewarded, while those of poor performing funds are not (Starks, 1987). Hence, a multi-fund manager with two average performing funds increases his compensation if he pushes the performance of one of the funds to the point where an incentive clause in his contract is activated, while he is no worse off due to the poor(er) performance of the other fund.

It is not immediately obvious how opposite trading or cross trading can benefit one fund at the cost of another. First, opposite trading can provide price support for high-value funds that place orders on the open market (Golez and Marin, 2012). Second, cross-trading occurs when funds transfer assets internally between funds. Such transfers of assets are regulated by SEC rule 17a-7 of the Investment Company Act of 1940 and are allowed as long as securities are quoted at the prevailing market price. Sotes-Paladino and Goncalves-Pinto (2011) show, however, that low-value funds can sub-optimally hold illiquid assets and cross-trade those assets at the benefit of high-value funds.

I conduct an empirical examination of opposite trading activity with the underlying alternative hypthesis that multi-fund managers engage in cross fund subsidization via opposite trading and present considerable results in favor of my presumption. First, regressions of risk-adjusted abnormal returns on my measure of cross-fund subsidization for sorts of the highest performing multi-managed funds yield that multi-fund managers increase the performance of top-funds to the extent of approximately 2% in terms of annualized alphas. Moreover, the cost to low performing funds is around the order of -2.5% in annualized alpha. When looking at sorts based on year-to-date returns, I also find some empirical evidence in support of cross fund subsidization by multi-fund managers, indicating that this activity can have an effect on the family tournament. Moreover, I find that multi-fund managers do not exploit the lowest performing funds in terms of year-to-date returns as a tool for opposite trading, reflecting their will to keep their funds and not risk losing any to extreme bad performance (Khorana, 1996).

I can identify multiple strands of literature which relate to this work. First, this work contributes to the growing discourse on multi-fund mutual fund managers. Baks (2003) was the first to compile a database that tracked mutual fund managers throughout their careers and explores the degree in which a fund's performance is attributable solely to the manager. Yadav (2010) explores the performance implications of portfolio matching by multi-fund managers, while Agarwal and Ma (2012) obtain a novel dataset tracking mutual fund managers at the monthly time frequency in order to quantify the effect of switching from singletasking to multitasking. This work also adds to the growing discourse on cross-fund subsidization and agency conflicts in delegated portfolio management. In addition to Gaspar et al. (2006), Guedj and Papastaikoudi (2005) show that larger fund families are more likely to engage in cross-fund subsidization. Bhattacharya et al. (2012) show . Schmidt and Goncalves-Pinto (2012) show that mutual funds coordinate trades in support of funds experience large firesales. Last, it relates to the literature on managerial tournaments in the mutual fund industry. Brown et al. (1996) show that mutual fund managers engage in a segment tournament, where funds compete to be the top performing in their respective investment styles. Busse (2001) revists mutual fund tournaments using daily data to compute measures of volatility. Kempf and Ruenzi (2008) show that mutual fund managers engage in a family tournament. My work is the first that uses mutual fund holdings in order to test the existence of cross fund subsidization between multi-managed funds.

The rest of the paper is organized as follows. In section 2.2, I present the dataset and cleaning procedure to develop my database of mutual fund managers and explain my measure of opposite trading. Section 2.3 follows with my empirical strategy, analysis and results. Section 2.4 concludes.

2.2 Data

The data accumulated for this study are accessed from the CRSP Survivorship Bias Free Mutual Fund Database, from which I initially access data items pertaining to the management of each available mutual fund after year 1999. Namely, I access the CRSP fund and portfolio identifiers (`crsp_fundno` and `crsp_portno`, respectively), the fund name (`fund_name`), the name of the management company (`mgmt_name`), the manager name (`mgr_name`), the beginning date of the manager for each mutual fund (`mgr_dt`) and the lipper objective code (`lipper_obj_cd`).

2.2.1 Manager Names

The names indicated in the `mgr_name` field contain many inconsistencies. Since the aim is to construct a comprehensive database that tracks mutual fund managers and managerial teams throughout the time series, the data undergo a rigorous cleaning process in order to uniquely identify each individual manager and team. First, I exclude all funds for which a unique manager or managerial team is not specified. Hence, I omit funds for which the name indicated is "Team Managed", "Committee", "Multiple Managers" or any other string which would otherwise be associated with a non-specific manager or team.[1] Next, I clean all name prefixes and suffixes a manager may have.[2] Initials and middle names are kept in each string, while all punctuation is removed. I treat managerial teams as single managers if each team member is identified in the `mgr_name` field. Observations for which specific managerial teams are indicated often report the names of each team member in random order, and can include a maximum of 5 team members. Since the order of names can differ throughout the time series and cross-section, I organize each team into alphabetical order according the each team member's last name.

[1] This includes "advisor", "portfolio", and "invest" amongst others.

[2] All prefixes are removed from the name string. Suffixes pertaining to degrees are also dropped from the name string. Managers for which "junior" and "senior" are specified in various forms are indicated with suffixes of "jr" and "sr".

Hence, two team-managed funds consisting of the same managers reported in different order in the same fund family are assumed to have the same manager.

Inconsistencies with name spellings, nicknames and initials are also problematic, since a simple variation of a manager's name can be interpreted as a different manager in the data. In order to mitigate the effect of these inconsistencies, I apply the following filter to equate name strings that are highly similar in the same `crsp_fundno-mgr_dt` panel. For each manager name (by first name then last name) I compute its soundex code, which is a string consisting of a letter followed by three numbers corresponding to the phonetic code of the name (Pfeifer et al., 1996). The letter component corresponds to the first letter in the manager's name and the numbers are assigned according to the subsequent order of consonants in the name string. For example, "John Smith" translates into a soundex code of "J525". I then find the manager name that appears most frequently for each `crsp_fundno-mgr_dt` panel and compute its soundex code. Similar sounding names beginning with the same letter are separated by a few numbers in the numerical component of their soundex codes. I repeat this with last names followed by first names. If the soundex code differs by a value less than 2 (extremely similar sounding name), I interpret this as a typo or inconsistency and replace the name string with the mode for that `crsp_fundno-mgr_dt` panel.

Cleaning manager names using this filter can have numerous drawbacks. First, the `mgr_dt` is an error prone field (Baks, 2003; Agarwal and Ma, 2012). Therefore, the true `crsp_fundno - mgr_dt` panel can be much larger than that reported by CRSP. Additionally, soundex codes are not able to correctly find name matches for misspelled names which contain many consonants. Last, the choice of adjusting manager names within each `crsp_fundno - mgr_dt` panel can be considered conservative and does not capture for example the cases in which a manager's given name is used for one fund and nickname is used for another fund.[3] In order the mitigate the bias that these effects may have on my data, I conduct the following controls. Since I expect the problematic funds to be those with many managerial changes, I manually check those mutual funds which exhibit a high degree of variation in soundex codes. Moreover, I also check and accordingly adjust the remaining name entries manually using random subsamples of the data.

Mutual funds offer different share classes to investors based on their investment horizons and initial investments. Different share classes of the same mutual fund are reported as different funds in the CRSP database, therefore I consolidate multiple share classes in order to prevent the double reporting of funds. Share classes are indicated after a semi-colon in the `fund_name` field in the CRSP database. In order to determine which share classes belong to each fund, I parse each fund

[3] A more aggressive approach would be to adjust names by the `mgmt_name` - date panel, but the number of false adjustments made were too high.

name into two components using the semi-colon as a separator and match the funds by name. I also merge share classes with the `crsp_fundno` and `crsp_portno` mapping to be sure that each share class is coupled with the same holdings information.

After merging multiple share classes into one fund, I match funds in the same fund family according to the manager's name in order to determine the number of funds each manager is responsible for in the dataset. Table (2.1) presents the frequency distribution of multi-fund management in my sample at this point of the cleaning process. Approximately 48% of the observations are attributable to single-fund managers and 52% of the observations pertain to multi-fund managers of 2 or more funds.

2.2.2 Opposite Trading and Opp

I document the degree in which multi-fund managers engage in opposite trading of equities by checking the security hold-

Table 2.1: Multi-fund managing
This table presents the frequency distribution of funds managed for all observations in the dataset. Data included are from January 2003 until December 2011. I compute the number of simultaneous funds managed by matching manager names in each family-fund calendar date panel and counting the respective number of funds run by that manager. Frequencies are computed prior to the style exclusions for actively managed equity mutual funds.

Funds Managed	Percent of Observations	Cumulative (%)
1	48.04	48.04
2	18.44	66.48
3	8.53	75.01
4	4.49	79.50
5	2.40	81.90
6	2.03	83.93
7	0.82	84.75
8	0.99	85.73
9	1.32	87.05
10	0.65	87.70

Results reported are cut off at 10 simultaneously managed funds.
Data are fund-quarter observations

ings in each cross section of funds managed. As mentioned above, I exploit the degree in which a multi-fund manager conducts opposite trades as my mechanism for cross fund subsidization. In order to correctly measure opposite trades as a mechanism for cross fund subsidization, I use all funds managed by a multi-fund manager which contain equity positions (according to PERMNO) regardless of fund objective. Mutual fund holdings are obtained from CRSP and appear in my sample at the quarterly time frequency. For each multi-fund manager $n \in N$ at each point in time t, I check the change in holding $s \in S_n$ for each fund $f, f' \in F_n$ from time $t-1$ to time t, where F_n denotes the funds run by manager n and S_n the entire set of holdings held by each fund managed by manager n. I denote the change in holdings of security s as $\Delta H_{s,t} = H_{s,t} - H_{s,t-1}$. In order to accumulate the number of opposite trades executed by fund f, I first define the following indicator function for buy-side opposite trades executed by f:

$$1^{Buy}_{s,f,f'} = \begin{cases} 1 \text{ if } \Delta H_{s,f} > 0 \text{ and } \Delta H_{s,f'} < 0 \text{ for } f \neq f' \\ 0 \text{ otherwise} \end{cases} \tag{2.1}$$

thus, $1^{Buy}_{s,f,f'} = 1$ when an increase in holding s of fund f is accompanied by a decrease in holding s for another fund f' run by the same manager in period t. Moreover, I capture possible opposite trading opportunities by defining

$$1^{Poss}_{s,f,f'} = \begin{cases} 1 \text{ if } H_{s,f} > 0 \text{ or } |\Delta H_{s,f}| > 0 \\ \text{ and } H_{s,f'} > 0 \text{ or } |\Delta H_{s,f'}| > 0 \text{ for } f \neq f' \\ 0 \text{ otherwise.} \end{cases} \tag{2.2}$$

Therefore, 1^{Buy} captures a buy-side opposite trade while 1^{Poss} captures an opposite trading opportunity.[4] This procedure is repeated for each time period that the multi-fund manager remains in the sample as long as the holdings information are

[4]Sell side opposite trading opportunities essentially equate to the number of securities held by fund f, thus are negated.

available. I define the proportion of opposite trading for each fund f of manager n in each time period as

$$Opp^{+}_{f,n} = \frac{\sum_{s \in S_n} \sum_{f' \neq f \in F_n} 1^{Buy}_{s,f,f'}}{\sum_{s \in S_n} \sum_{f' \neq f \in F_n} 1^{Poss}_{s,f,f'}} \tag{2.3}$$

where the plus superscript indicates the proportion with respect to buy-side opposite trades conducted by fund f. In order to capture the proportion of both buy and sell side opposite trades conducted by fund f, I define

$$1^{BuySell}_{s,f,f'} = \begin{cases} 1 \text{ if } \Delta H_{s,f} > 0 \text{ and } \Delta H_{s,f'} < 0 \text{ for } f \neq f' \\ 1 \text{ if } \Delta H_{s,f} < 0 \text{ and } \Delta H_{s,f'} > 0 \text{ for } f \neq f' \\ 0 \text{ otherwise} \end{cases} \tag{2.4}$$

and derive the proportion of buy and sell side opposite trades conducted by fund f as

$$Opp^{\pm}_{f,n} = \frac{\sum_{s \in S_n} \sum_{f' \neq f \in F_n} 1^{BuySell}_{s,f,f'}}{\sum_{s \in S_n} \sum_{f' \neq f \in F_n} 1^{Poss}_{s,f,f'}} \tag{2.5}$$

where *Opp* now receives the superscript $\pm$. The proportions in (2.3) and (2.5) are computed across time for each multi-manager mutual fund at the quarterly time frequency and are bounded between 0 and 1 by construction. Note that strategic cross trading can include a strategic sell as well as a strategic buy, hence the distinction between $Opp^{\pm}_{f,n}$ and $Opp^{+}_{f,n}$. Moreover, I implicitly assume that multi-fund managers engage in strategic opposite trading only with other funds that he/she manages; hence, there is no coordination problem with other managers. Last, *Opp* contains opposite trades relative to all funds managed by the same fund manager whereas Gaspar et al. (2006) only consider pairwise measures of opposite trading. Measuring cross-fund subsidization with *Opp* is also fund specific as opposed to manager specific and allows me to draw inference between strategic opposite trading and fund-level attributes or benchmarks.

I also derive a similar measure of opposite trading for single-fund managers. In order to do so, I assume that single-fund managers engage in opposite trading activity only with other single-fund managers in the same fund family. I accumulate the proportion of opposite trading similar as in (2.3) and (2.5) except now the set of mutual funds in F contain single-managed funds for a fund family and S contains the entire set of security holdings contained in F. I denote the proportion of buy-side opposite trades and buy- and sell-side opposite trades for single-fund managers as $Opp^{+}_{f,s}$ and $Opp^{\pm}_{f,s}$, respectively.

2.2.3 Fund Characteristics

I obtain monthly data on returns and total net assets for each `crsp_fundno` and define portfolio returns as the weighted average of the share class returns according to total net assets under management. Expenses, fees and turnover are computed similarly. New money growth (flows) is defined as

$$NMG = \frac{TNA_{m,t} - TNA_{m,t-1}(1 + R_{m,t})}{TNA_{m,t-1}} \qquad (2.6)$$

where $R_{m,t}$ represents share class specific return at time t. Aggregate proportional flows are taken as the weighted average of NMG according to total net assets under management across share classes for the previous three months.

I follow Gaspar et al. (2006) and include only funds which adhere to an investment objective reflective of an actively managed domestic equity mutual fund.[5] Mutual funds that engage in passive investment strategies, such as index funds, have also been removed.[6]

For my measures of performance, I follow Carhart (1997), Wermers (2003) and Kacperczyk and Seru (2007) and compute unconditional alphas using pricing models which include a single market factor (Sharpe, 1964; Lintner, 1965), HML and SMB factors (Fama and French, 1993), and a momen-

[5]Lipper objective codes which are included in my analysis include: CA CG CS G GI ELCC EMN FS H I ID LCCE LCGE LCVE LSE MATD MATE MATH MATI MTAG MCCE MCGE MCVE MLCE MLGE MLVE RE S SCCE SCGE SCVE SG TK TL UT.

[6]Regarding mutual funds for which a passive investment strategy is not explicitly specified in the Lipper objective codes, I compute correlation coefficients between fund returns and returns on the S&P 500 stock index, and omit funds with correlations larger than 0.995.

tum factor (Carhart, 1997). Risk factors in the unconditional model have been acquired from Kenneth French's website. To determine alphas for month t, I run the following regression

$$\begin{aligned} R^e_{i,T} = \alpha_i + \beta_{i,MKT} * R^e_{M,T} + \beta_{i,SMB} * SMB_T \\ + \beta_{i,HML} * HML_T + \beta_{i,UMD} * UMD_T + \varepsilon_{i,T} \end{aligned} \tag{2.7}$$

for $T = t - 24, \ldots, t - 1$ where R^e indicates excess returns over a risk-free benchmark for the fund returns and market returns, respectively. Equation (2.7) is estimated using the previous 24 months of returns in order to mitigate look-ahead bias (Carhart, 1997). Four-factor alphas for time t are derived by taking the factor loadings from (2.7) and determining the fitted value of the regression using risk premia in time t, i.e.

$$\begin{aligned} \alpha_{i,t} = R^e_{i,t} - \hat{\beta}_{i,MKT} * R^e_{M,t} - \hat{\beta}_{i,SMB} * SMB_t \\ - \hat{\beta}_{i,HML} * HML_t - \hat{\beta}_{i,UMD} * UMD_t \end{aligned} \tag{2.8}$$

Alphas for the Fama-French and CAPM specifications are computed using the same procedure in equations (2.7) and (2.8) with the momentum factor excluded and the SMB & HML factors omitted, respectively. Computing alphas using a 24-month window imposes the requirement that funds be in the sample for at least 2 years, otherwise they are excluded from the analysis.

Finally, I exclude all observations for which alphas, holdings, fund size, age, turnover, expenses and flows are missing. I also exclude fund family quarters for which there aren't at least 7 unique mutual fund managers identifiable in the data in order to mitigate cases where families simply report the CIO or other senior manager as the head of each fund. The final sample size includes 1467 mutual funds from September 2003 until December 2011. Summary statistics are separated by multi-fund and single fund managers, and are reported in tables (2.2) and (2.3) respectively. Before computing these statistics, I first find the time-series mean of each data item for each mutual fund in order to mitigate any possible biases attributable to funds which have been in the sample for a longer period of time.

A quick comparison of tables (2.2) and (2.3) shows that mutual funds run by multi-fund managers yield higher raw and risk adjusted returns than single-managed funds. The t-test for difference in means of raw returns between multi-fund and single fund managers yields a t-statistic of 5.02 and a p-value of 0.001%. Funds run by single-fund managers tend to be over 3 years older on average. Moreover, flows to multi-managed funds are almost one percentage point higher, hence these funds receive larger positive cash inflows on average, likely due to performance chasing as documented in Sirri and Tufano (1998).

Tables (2.2) and (2.3) also provide a brief look into the opposite trading activity of mutual funds in the sample. On average, multi-managed funds execute approximately 3.5% of their buy and sell side opposite trades and 1.6% of their buy-side opposite trades (according to $Opp^{\pm}$ and Opp^{+}, respectively). Moreover, the maximum of $Opp^{\pm}$ is 48.7%, i.e. the

Table 2.2: Summary statistics for funds run by multi-fund managers
Data are obtained from the CRSP Survivorship Bias Free Mutual Fund Database and include multi-managed funds from January 2003 until December 2011. TNA are total net assets and are measured in millions of USD. New money growth indicates the proportion of fund inflows with respect to the previous 3 month level of total net assets. Raw returns are indicated with r_m. Expense ratio is the previous year's expense ratio as reported by CRSP, in percent. Age indicates the number of years since the fund was initiated. Opp^+ is the proportion of buy-side opposite trades conducted by a multi-managed mutual fund and indicates the number of buy-side opposite trades conducted with other funds run by the same multi-fund manager over the total number of opposite trading opportunities with other funds run by the same multi-fund manager throughout the last quarter of reported holdings. $Opp^{\pm}$ is the proportion of buy and sell side opposite trades executed by multi-managed mutual funds throught the last quarter of reported holdings. α represented risk adjusted abnormal returns using CAPM, Fama-French, and 4-factor loadings, respectively. The time series mean for each data item is taken before computing the belowmentioned statistics.

	(1) Mean	Standard Dev.	Minimum	Maximum
$TNA(millions)$	1099.7	5235.9	0.100	129651.8
$NMG(\%)$	2.212	14.46	-46.78	89.49
$r_m(\%)$	2.203	3.251	-12.43	18.22
$ExpRatio(\%)$	0.974	0.504	0	2.687
$Age(years)$	11.48	11.49	1.708	82.75
Opp^+	0.0159	0.0460	0	0.321
$Opp^{\pm}$	0.0357	0.0860	0	0.487
$\alpha_{CAPM}(\%)$	0.154	1.323	-8.000	7.543
$\alpha_{FF}(\%)$	0.0194	1.299	-7.774	12.29
$\alpha_{4factor}(\%)$	-0.0290	1.332	-11.01	11.47
Number of funds:	730			

Table 2.3: Summary statistics for funds run by single-fund managers Data are obtained from the CRSP Survivorship Bias Free Mutual Fund Database and include multi-managed funds from January 2003 until December 2011. TNA are total net assets and are measured in millions of USD. New money growth indicates the proportion of fund inflows with respect to the previous 3 month level of total net assets. Raw returns are indicated with r_m. Expense ratio is the previous year's expense ratio as reported by CRSP, in percent. Age indicates the number of years since the fund was initiated. Opp^+ is the proportion of buy-side opposite trades conducted by a single-managed mutual fund and indicates the number of buy-side opposite trades conducted with other single-managed funds in the same fund family over the total number of opposite trading opportunities with other single-fund managed families throughout the last quarter of reported holdings. $Opp^{\pm}$ is the proportion of buy and sell side opposite trades executed by single-managed mutual funds throught the last quarter of reported holdings. α represented risk adjusted abnormal returns using CAPM, Fama-French, and 4-factor loadings, respectively. The time series mean for each data item is taken before computing the belowmentioned statistics.

	(1) Mean	Standard Dev.	Minimum	Maximum
$TNA(millions)$	1101.0	3099.3	0.400	52433.1
$NMG(\%)$	1.300	16.61	-46.78	177.5
$r_m(\%)$	1.408	2.784	-15.22	12.55
ExpRatio(%)	1.210	0.395	0.01000	3.033
$Age(years)$	14.76	13.66	1.250	83.88
Opp_s^+	0.00905	0.0274	0	0.354
$Opp_s^{\pm}$	0.0191	0.0466	0	0.396
$\alpha_{CAPM}(\%)$	0.0672	1.211	-10.39	4.664
$\alpha_{FF}(\%)$	-0.0764	1.156	-9.049	5.574
$\alpha_{4factor}(\%)$	-0.144	1.111	-9.040	5.268
Number of funds:	736			

fund that engages in the highest degree of opposite trading activity executes half of their possible opposite trading opportunities on average. When considering only buy-side opposite trades, this percentage decreases to 32.1% of opposite trading opportunities. The proportion of opposite trades conducted by single-fund managers is smaller on average than when compared with multi-fund managers. For buy-side only trades, the proportion is very close to zero on average, and $Opp_s^{\pm}$ is equal to almost 2% on average.

2.3 Empirical Strategy

My conjecture is that mutual fund managers who manage more than one fund are inclined to subsidize top performing mutual funds at the expense of low performing mutual funds. This incentive to support top performing funds can be attributed to the positive relationship between fund performance and subsequent investor inflows (Chevalier and Ellison, 1997; Sirri and

Tufano, 1998).

2.3.1 Analysis and Results

In order to estimate the effect of opposite trading on performance benchmarks, I estimate the following empirical model

$$\alpha_{f,t} = \beta_0 + \beta_1 * Opp_{f,t} + \beta' * Controls + \varepsilon_{f,t}. \tag{2.9}$$

It is important to note that the relationship between $Opp_{f,t}$, and performance indicators is ambiguous when considering (2.9) with respect to the entire sample. If $Opp_{f,t}$ is used as a mechanism for cross-fund subsidization, then we should observe a widening of the difference in returns between high-value and low-value funds (Gaspar et al., 2006). Subsequently, $Opp_{f,t}$ should take a higher value for both high- and low-value funds. In order to test for any non-linearities between $Opp_{f,t}$ and the dependent variable, I estimate (2.9) for subsamples based on the highest and lowest performing funds as well as the full sample of multi-fund managers. In order to correctly make statistical inference, (2.9) is estimated using the panel corrected standard errors (PCSE) approach proposed by Beck and Katz (1995). The PCSE approach controls for contemporaneous correlation in the cross section and heteroskedasticity/autocorrelation in the error term. Control variables are included in every specification and include the natural log of total net assets, the natural log of age, the expense ratio of the fund, the turnover ratio, and the lagged level of NMG. Quarterly fixed effects are also included for each regression.

2.3.2 Cross Multi-Fund Subsidization

Under the null hypothesis of no cross-fund subsidization, I should observe $\beta_1 = 0$ from equation (2.9) for all subsamples of data. Such a null hypothesis implies that multi-fund managers do not coordinate their trading strategies between the funds they simultaneously manage. The alternative hypothesis is one of coordinated cross-fund subsidization between the high-value and low-value funds. Since $Opp_{f,t}$ should help high-value funds at the cost of low-value funds, I expect $\beta_1 > 0$ for high value funds and $\beta_1 < 0$ for low-value funds under the alternative hypothesis. Moreover, $\beta_1 = 0$ for the full sample can be considered a necessary but not sufficient condition in favor of the alternative hypothesis.

Table (2.4) presents the results of my estimation of (2.9) using the full sample of data. As expected and in support of the alternative hypothesis, I observe coefficients which are statistically indistinguishable from zero for the first three specifications based on $Opp_{f,t^{(\pm)}}$. Moreover, $Opp^{+}_{f,t}$ exhibits a statistically insignificant relationship with Fama-French and 4-factor alphas, but has a marginally significant relationship with the CAPM alpha at the 90% confidence level.

Results from my first direct test of cross fund subsidization between multi-managed funds is presented in table (2.5). Columns 1 through 3 for each panel include tests on the subsample of data including only the upper benchmark percentile

Table 2.4: I estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp_{f,t} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for the full sample of multi-fund managers. Standard errors are corrected for contemporaneous correlation, heteroskedasticity and autocorrelation in the error term. Specifications (1) through (3) include the proportion of buy and sell-side opposite trades conducted by fund f with other funds run by the same multi-fund manager and columns (4) through (6) include the proportion of buy-side opposite trades as regressors. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

	Full Sample					
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp_t^{\pm}$	0.392	0.0431	-0.164			
	(0.309)	(0.273)	(0.279)			
Opp_t^{+}				0.826*	0.291	0.103
				(0.466)	(0.439)	(0.423)
$lnTNA_t$	-0.0382**	0.0101	-0.00693	-0.0383**	0.00974	-0.00737
	(0.0185)	(0.0165)	(0.0162)	(0.0185)	(0.0165)	(0.0162)
$lnAge_t$	0.0879*	-0.00803	0.0641	0.0904*	-0.00543	0.0686
	(0.0477)	(0.0427)	(0.0416)	(0.0480)	(0.0429)	(0.0417)
$Expenses_t(\%)$	-6.372	-11.43**	-18.80***	-6.417	-11.09**	-17.84***
	(5.496)	(5.039)	(4.982)	(5.497)	(5.025)	(4.978)
$Turnover_t(\%)$	-0.00470	0.0775**	0.116***	-0.00466	0.0778**	0.117***
	(0.0408)	(0.0372)	(0.0412)	(0.0407)	(0.0371)	(0.0411)
NMG_{t-1}	-4.551	-4.899	3.124	-4.578	-4.915	3.091
	(7.890)	(7.092)	(6.916)	(7.893)	(7.094)	(6.932)
N	4420	4420	4420	4420	4420	4420
Panels	730	730	730	730	730	730
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

Table 2.5: I estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp^{\pm}_{f,t} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for sorts of multi-fund managers based on contemporaneous raw returns. Specifications (1) through (3) include the proportion of buy and sell-side opposite trades conducted by fund f with other funds run by the same multi-fund manager for the high return sorted funds for each multi-fund manager and columns (4) through (6) include the low sorts. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

Panel A:	$H_{95} = 1$			$L_5 = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_t$	5.445***	2.234**	1.301*	-4.639***	-2.564**	-3.678***
	(1.634)	(0.954)	(0.695)	(1.074)	(1.082)	(1.181)
N	1445	1445	1445	1444	1444	1444
Panels	464	464	464	460	460	460
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:	$H_{80} = 1$			$L_{20} = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_t$	4.873***	2.088***	2.106***	-2.217***	-1.365***	-2.308***
	(0.997)	(0.764)	(0.609)	(0.461)	(0.453)	(0.557)
N	1604	1604	1604	1621	1621	1621
Panels	532	532	532	534	534	534
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel C:	$H_{50} = 1$			$L_{50} = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_t$	2.335***	0.899*	0.989**	-1.210***	-0.915***	-1.146***
	(0.497)	(0.465)	(0.402)	(0.273)	(0.332)	(0.333)
N	2051	2051	2051	2369	2369	2369
Panels	624	624	624	648	648	648
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

of funds of each multi-fund manager sorted by returns, and columns 4 through 6 include tests on the subsample of data including only the lower benchmark percentiles of funds of each multi-fund manager sorted by returns. Panel A includes benchmarks for the upper 95th and lower 5th performing percentile of funds for each multi-fund manager. Rows 1 through 3 indicate the results on the subsample of funds in the upper 20th percentile of returns for each multi-fund manager. The coefficient on $Opp_t^{\pm}$ is highly positive and significant at the 99% significance level for all three specifications, indicating that the managers who buy securities in their top funds and sell them in other funds increase the performance of their top funds. Moreover, a one standard deviation increase in $Opp_t^{\pm}$ translates to an economically significant 2% increase in annualized Fama-French alpha (.08*2.074*12) for the very high performing funds in Panel A. Rows 1 through 3 in panels B and C represent upper percentile cutoffs of 80% and 50%, respectively. The coefficient on $Opp_t^{\pm}$ remains positive and highly statistically significant for all regression specifications. Note as well that the magnitude on $Opp_t^{\pm}$ decreases as we reduce the percentile benchmark for returns. This indicates that strategic cross-fund opposite trading is aimed more towards pushing up the performance of the best performing funds in each given quarter.

Columns 4 through 6 from table (2.5) display the results based on subsamples of funds which perform the worst for the multi-fund managers. The coefficient on $Opp_t^{\pm}$ for these funds is highly negative and statistically significant. For example, a one standard deviation increase in $Opp_t^{\pm}$ translates to a -2.66% decrease in annualized Fama-French alpha. As I increase the lower benchmark from 5% to 20% and 50% in Panels B and C, the magnitude of the coefficient on $Opp_t^{\pm}$ decreases similarly as it did in models 1 through 3, hence $Opp_t^{\pm}$ has its largest effect on the lowest performing funds of each multi-fund manager in the sample. Overall, I take the results in table (2.5) as strong evidence in favor of opposite trading as a mechanism for cross fund subsidization for multi-fund managers.

Table (2.6) presents the results of my tests for cross fund subsidization using the proportion of buy-side opposite trades (Opp_t^+) as my measure of opposite trading. The results remain robust, albeit a bit weaker in terms of statistical significance when comparing the coefficients in table (2.6) to those of table (2.5). Close consideration of these results allow me to draw two more interesting conclusions. First, if sell-side opposite trades are excluded from my measure of cross-fund subsidization (as they are with Opp_t^+), the relationship between Opp and fund alphas weakens for the high percentile sort. This implies that multi-fund managers rely on sell-side trading opportunities in order to improve the performance of their high-value funds. Moreover, the coefficients on Opp_t^+ in columns 4 through 6 are larger in magnitude than those of the low percentile sorts for $Opp_t^{\pm}$ in table (2.5). Hence, when considering only buy-side opposite trades as a mechanism for strategic cross-fund subsidization, the effect of Opp on performance benchmarks is

Table 2.6: estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp^{+}_{f,t} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for sorts of multi-fund managers based on contemporaneous raw returns. Specifications (1) through (3) include the proportion of buy-side opposite trades conducted by fund f with other funds run by the same multi-fund manager for the high return sorted funds for each multi-fund manager and columns (4) through (6) include the low sorts. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

Panel A:	$H_{95} = 1$			$L_5 = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
Opp^{+}_{t}	5.243**	1.078	0.402	-4.845***	-3.781***	-5.030***
	(2.429)	(1.491)	(1.102)	(1.422)	(1.318)	(1.695)
N	1445	1445	1445	1444	1444	1444
Panels	464	464	464	460	460	460
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:	$H_{80} = 1$			$L_{20} = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
Opp^{+}_{t}	4.639***	0.930	1.538	-2.526***	-1.558*	-3.409***
	(1.382)	(1.285)	(1.020)	(0.912)	(0.822)	(1.079)
N	1604	1604	1604	1621	1621	1621
Panels	532	532	532	534	534	534
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel C:	$H_{50} = 1$			$L_{50} = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
Opp^{+}_{t}	2.309***	0.532	1.366**	-1.376***	-0.721	-1.648***
	(0.633)	(0.763)	(0.576)	(0.471)	(0.606)	(0.609)
N	2051	2051	2051	2369	2369	2369
Panels	624	624	624	648	648	648
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

worse. This implies that an important driver in the wedge between the top performing funds and the low performing funds for multi-fund managers is the sell side opposite trade from the high to the low percentile sorts.

2.3.3 Sorts on YTD Returns

Multi-fund managers cross-subsidize their top performing funds at the expense of the low performing funds, but will only do so until the threshold where a low-performing fund would get reassigned to another manager, since losing a fund impacts the manager's compensation. Khorana (1996) documents an inverse relationship between the probability of managerial replacement and previous period's performance. Hence, if a manager systematically lowers the performance of low year-to-date funds, he risks losing that fund. In this section, I test for this by repeating the analysis using sorts based on a multi-fund manager's high and low YTD return funds.

Table (2.7) presents the results of running regression (2.9)

Table 2.7: I estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp^{\pm}_{f,t} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for sorts of multi-fund managers based on contemporaneous year-to-date returns. Specifications (1) through (3) include the proportion of buy and sell-side opposite trades conducted by fund f with other funds run by the same multi-fund manager for the high YTD return sorted funds for each multi-fund manager and columns (4) through (6) include the low YTD return sorts. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

Panel A:	$H_{95}(YTD) = 1$			$L_5(YTD) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_t$	0.833	-1.004	-0.618	0.489	-0.212	-1.285***
	(0.951)	(0.849)	(0.720)	(0.566)	(0.553)	(0.496)
N	1417	1417	1417	1536	1536	1536
Panels	433	433	433	416	416	416
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:	$H_{80}(YTD) = 1$			$L_{20}(YTD) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_t$	1.321**	1.177**	1.286**	-0.326	-0.203	-1.111**
	(0.587)	(0.580)	(0.525)	(0.624)	(0.596)	(0.546)
N	1561	1561	1561	1708	1708	1708
Panels	493	493	493	475	475	475
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel C:	$H_{50}(YTD) = 1$			$L_{50}(YTD) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_t$	0.149	0.0630	0.0493	0.827*	0.116	-0.595
	(0.396)	(0.363)	(0.358)	(0.457)	(0.457)	(0.404)
N	1985	1985	1985	2435	2435	2435
Panels	572	572	572	577	577	577
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

on sorts of high and low YTD percentiles. Panel A includes sorts on the 95th percentile and higher performing year to date funds and lowest 5th percetile funds, Panel B sorts on benchmarks of 80% and 20%, while the set of multi-managed funds is separated by the median in Panel C. First, notice that in columns 4 through 6, there is no systematic decrease in performance based on changes in $Opp_t^{\pm}$. Only for the 4-factor specification in column 6 is the relationship between *Opp* and performance negative and significant. More importantly, note that there is no significant relationship between *Opp* and alphas for the high percentile YTD return sort in Panel A. This result indicates that multi-fund managers do not systematically cross subsidize their best performing funds. This could be due to various reasons. First, those funds may have already been previous "winners", i.e. are high-value funds for their respective fund families and have already received relatively higher advertising exposure (Kempf and Ruenzi, 2008) or preferential IPO allocations (Gaspar et al., 2006). Therefore, the incentive to further increase the performance of these funds may be marginal. Panel B, however, shows that multi-fund managers cross-fund subsidize their upper 20th percentile of funds they simultaneously manage. The degree of subsidization is also economically significant. A one standard deviation increase in $Opp_t^{\pm}$ results in an increase of 1.23% of annualized 4-factor alpha (.08*1.286*12). The coefficient on $Opp_t^{\pm}$ is also statistically different from 0 at the 95% confidence level for all three specifications. Last, the YTD return sorts based on upper and lower 50th percentiles of funds managed by each manager yields a relationship with *Opp* that is unclear.

2.3.4 Single-fund Managers

In this section, I analyze the opposite trading activity of single-fund managers. Since single-fund managers do not have multiple portfolios between which they can engage in strategic cross fund subsidization as the multi-fund managers do, I assume that single fund managers coordinate opposite trades with other single-fund managers in the same fund family. However, coordination with other fund managers comes with an important caveat. Single-fund managers compete in a family tournament just as multi-fund managers do. Hence, opposite trading with other single-fund managers should be coupled with high coordination costs.

I begin by analyzing the relationship between the single-fund Opp and performance benchmarks using the full sample. As in the previous section, any effect that Opp could have on high performing funds should have the opposite effect on low performing funds, therefore there should be no relationship between measures of cross trading and fund performance. Specifications 1 through 3 of table (2.8) include $Opp^{\pm}_{t,s}$ as an explanatory variable and columns 4 through 6 include $Opp^{+}_{t,s}$ as the measure of cross trading. The coefficient on Opp_s for all 6 models is statistically indistinguishable from zero.

I test for fund subsidization between single fund managers by conducting sorts on returns as was done in tables (2.5)

Table 2.8: I estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp_{f,t,s} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for the full sample of single-fund managers. Specifications (1) through (3) include the proportion of buy and sell-side opposite trades conducted by fund f with other single-managed funds in the same fund family and columns (4) through (6) include the proportion of buy-side opposite trades as regressors. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

	Full Sample					
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_{t,s}$	0.00287	0.0574	0.110			
	(0.320)	(0.301)	(0.274)			
$Opp^{+}_{t,s}$				0.116	0.191	0.265
				(0.533)	(0.505)	(0.476)
$lnTNA_t$	0.00106	0.0185	-0.00282	0.000899	0.0182	-0.00297
	(0.0177)	(0.0172)	(0.0153)	(0.0177)	(0.0171)	(0.0152)
$lnAge_t$	0.0481	-0.0143	0.0184	0.0485	-0.0137	0.0191
	(0.0411)	(0.0400)	(0.0371)	(0.0411)	(0.0400)	(0.0370)
$Expenses_t(\%)$	0.0268	-3.585	-8.771*	0.0835	-3.556	-8.793*
	(5.233)	(4.921)	(4.734)	(5.210)	(4.906)	(4.722)
$Turnover_t(\%)$	0.0960***	0.0870***	0.0169	0.0960***	0.0868***	0.0167
	(0.0316)	(0.0304)	(0.0264)	(0.0316)	(0.0304)	(0.0264)
NMG_{t-1}	0.172	-0.205	0.807**	0.172	-0.203	0.809**
	(0.449)	(0.428)	(0.368)	(0.449)	(0.428)	(0.368)
N	5261	5261	5261	5261	5261	5261
Panels	736	736	736	736	736	736
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

and (2.6), with one main difference. The sorts are based on the upper and lower percentile of returns for single-managed funds in each fund family. The results of the analysis based on the proportion of buy and sell side opposite trades conducted are presented in table (2.9).

Panel A of table (2.9) includes sorts on funds in the upper 80th and lower 20th percentiles of single-managed funds based on returns. The results indicate that Opp_s exhibits radically different behavior with respect to performance benchmarks than *Opp* did for multi-fund managers. Columns 1 through 3 of Panel A indicate that there is a negative and statistically significant coefficient on $Opp^{\pm}_{t,s}$ where in columns 4 through 6 there is a large and statistically positive effect on the low performing funds in the sample. Panel B contains very similar results with sorts conducted at the 50th percentile. These results are in support of Kempf and Ruenzi (2008) who document that high performing funds tend to purchase more risky assets in a fund tournament in order to "lock in" their posi-

Table 2.9: I estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp^{\pm}_{f,t,s} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for sorts of multi-fund managers based on contemporaneous raw returns. Specifications (1) through (3) include the proportion of buy and sell-side opposite trades conducted by fund f with other single-managed funds in the same fund family for the high return sorted funds for each single-fund manager and columns (4) through (6) include the low sorts. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

Panel A:	$H_{80}(Single) = 1$			$L_{20}(Single) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_{t,s}$	-1.389**	-1.448**	-1.244**	1.330*	1.076	0.393
	(0.613)	(0.684)	(0.603)	(0.764)	(0.806)	(0.786)
N	1183	1183	1183	1160	1160	1160
Panels	475	475	475	514	514	514
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:	$H_{50}(Single) = 1$			$L_{50}(Single) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{\pm}_{t,s}$	-1.181***	-1.138***	-0.963***	1.666***	1.516***	1.048***
	(0.331)	(0.326)	(0.314)	(0.436)	(0.434)	(0.400)
N	2508	2508	2508	2753	2753	2753
Panels	630	630	630	673	673	673
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

tion in the fund family when strategic interactions are present. Conversely, low-performing funds are forced to engage in an opposite strategy in order to have a chance to catch up to the winners. Therefore, the existence of strategic interaction in this context (since single-fund managers are competing with one another) appears to have a strong influence on the results.[7]

Table (2.10) presents results using $Opp^{+}_{t,s}$ as an explanatory variable, and largely resemble the results obtained in table (2.9). The relationship between $Opp^{+}_{t,s}$ and fund alpha is negative and significant for the high performing funds and positive and significant for the low performing funds. The magnitudes on $Opp^{+}_{t,s}$ seem to be much higher, indicating that buy-side opposite trades seem to be the driver of extreme performance for single-managed funds.

The results in tables (2.9) and (2.10) are not indicative of cross-fund subsidization. Moreover, it seems indicative of the strategic interaction hypothesized by Kempf and Ruenzi

[7]Moreover, the lower bound requirement of 7 fund managers includes family sizes where such strategic interactions take place (small families).

Table 2.10: I estimate the following regression:
$\alpha_{f,t} = \beta_0 + \beta_1 * Opp^{+}_{f,t,s} + \beta' * Controls + \varepsilon_{f,t}$
using Panel Corrected Standard Errors (Beck and Katz, 1995) for sorts of multi-fund managers based on contemporaneous raw returns. Specifications (1) through (3) include the proportion of buy-side opposite trades conducted by fund f with other single-managed funds in the same fund family for the high return sorted funds for each single-fund manager and columns (4) through (6) include the low sorts. Control variables include the log level of total net assets, the log level of fund age, the percentage expense ratio from the last reported fiscal year, the turnover ratio, and lagged levels of net money growth.

Panel A:	$H_{80}(Single) = 1$			$L_{20}(Single) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{+}_{t,s}$	-2.924**	-3.292**	-2.650**	3.190**	2.756**	1.205
	(1.238)	(1.400)	(1.238)	(1.327)	(1.203)	(1.656)
N	1183	1183	1183	1160	1160	1160
Panels	475	475	475	514	514	514
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B:	$H_{50}(Single) = 1$			$L_{50}(Single) = 1$		
	(1)	(2)	(3)	(4)	(5)	(6)
	α_{CAPM}	α_{FF}	$\alpha_{4factor}$	α_{CAPM}	α_{FF}	$\alpha_{4factor}$
$Opp^{+}_{t,s}$	-1.712***	-1.414***	-1.245***	3.267***	3.246***	1.996**
	(0.487)	(0.470)	(0.480)	(0.750)	(0.738)	(0.783)
N	2508	2508	2508	2753	2753	2753
Panels	630	630	630	673	673	673
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Quarterly FE	Yes	Yes	Yes	Yes	Yes	Yes

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

(2008). However, it is important to point out that in the absence of strategic in interaction, such as for multi-fund managers, opposite trading can be implemented as a mechanism for cross fund subsidization. Therefore, multi-fund managers have the ability to manipulate their relative rankings in a family fund tournament, whereas this mechanism is not privileged to managers who run only one fund.

2.4 Conclusions

I have shown considerable evidence in support of opposite trading used as a mechanism for cross fund subsidization amongst multi-fund managers. This effect is not only highly statistically significant, but also very economically signiicant. Single fund managers that engage in opposite trading strategies have been shown to have an opposite effect. Moreover, multi-fund managers do not sacrifice the extreme low performing mutual funds in order to cross fund subsidize top performing funds, likely due to the fear that they will lose their fund after extreme bad performance.

My results have serious regulatory implications. Until now, fund families need only to indicate in a fund's perspectus if that manager runs other funds. Moreover, SEC rule 17a-7 of the Investment Company Act of 1940 should be sternly revisited in order to take into account the abiguity of the fair market price for illiquid securities. Clearly, investors in the low-end of the cross fund subsidization are losing and managers could be found to be in breach of fiduciary duty.

Chapter 3

Propensity Score Matched Discretionary Accruals

with Pureum Kim and Roberto Steri

JEL Classification: M41, C12, C15

3.1 Introduction

Conventional accounting standards allow a certain degree of flexibility for managers to report earnings. The channel through which earnings manipulation is possible is through accruals, which can be decomposed into a discretionary and non- discretionary component. The estimation of the discretionary component, otherwise known as earnings management, is of great interest to researchers and regulators, because it sheds light on the direct manipulation of earnings by upper management. Moreover, the existence of a high or low discretionary accruals component has asset pricing implications. Low discretionary accrual firms earn positive abnormal returns over a given window of time, while high discretionary accruals firms earn negative abnormal returns (Xie, 2001).

Many models exist in the earnings management literature to estimate the discretionary component of accruals. Healy (1985) and DeAngelo (1986) look at changes in accrual levels as proxies for earnings management, while more complicated statistical models also attempt to estimate earnings manipulation (Jones, 1991; Dechow et al., 1995; Kothari et al., 2005). In order to assess the appropriateness of these models, specification and power tests with the null hypothesis of no earnings management are done to measure the likelihood of type-I and type-II errors, respectively. Unfortunately, most of these models suffer from a certain degree of statistical misspecification, which means that the likelihood of making false positives in earnings management hypothesis testing is still too high. Moreover, the high degree of statistical misspecification can be attributed to omitted variables that have been shown to be correlated with accruals. Hence, a model which incorporates the combined effect of these omitted variables could greatly improve the degree of statistical specification, but until now models have been limited in terms of the number of covariates they can incorporate into their models.

Employing a propensity score matching technique, we investigate how conventional models of estimating discretionary accruals are affected in terms of statistical specification and power. The propensity score approach aims to match firms by a function of covariates rather than by each covariate individually. This characteristic enables researchers to match firms by a number of covariates to effectively control for contemporaneous correlation and any nonlinearities with accruals. Due to these advantages, the usage of propensity score matching has started to increase in accounting and finance (Tucker, 2012). Based on empirical testing, we form a choice model and then test against the traditional accruals models on earnings management detection using simulations. Overall, our findings show that propensity score matching is the most well defined discretionary accruals model with respect to mitigating false positives with the null of no earnings management, and does not sacrifice statistical power when compared to other matching models.

Power and specification exhibit a concession which prevents excessively improving one without worsening the other.

Because of this trade-off, the suitability of one model over another relies on the underlying economic consequences of their implementation. For example, if the cost of a type II error is higher than the cost of type I error, one should prefer a model that rejects more frequently the null hypothesis, resulting in a consequent risk of increasing the number of times in which the null hypothesis is wrongly rejected even though it is actually true. Because of the lack of information about the general cost of these two types of error without a specific case as a frame of mind, we try to resolve this trade-off by jointly evaluating the specification and power of our accruals measure with respect to other measures (Dechow et al., 1995; Kothari et al., 2005).

This paper is organized as follows. Section 2 reviews the relevant methodological papers in earnings management research. Section 3 motivates the determinants used in our model and presents the dataset. Section 4 reports the results of the simulation analysis where we examine the performance of propensity score matching to other earnings management models in a specification test. Section 5 follows with a test of statistical power for each earnings management model. Section 6 concludes.

3.2 Literature Review

We begin by reviewing the most commonly employed accruals models for measuring earnings management attempt to identify the discretionary portion of the accruals. Healy (1985) and DeAngelo (1986) first developed an earnings management detection model by examining the level and change of total accruals with respect to changes from the previous time period and the time series average of total accruals, respectively. These models suffered from an assumption that nondiscretionary accruals were time invariant. Therefore, in order to relaxing the assumption of time invariance of nondiscretionary accruals, Dechow and Sloan (1991) develop an industry model where nondiscretionary accruals are regressed on the two-digit SIC industry median total accruals in order to more accurately filer out the non-discretionary component of accruals. Jones (1991) also relax the assumption of constant nondiscretionary accruals and attempts to control for the change in the firm's business model by specifying nondiscretionary accruals as a function of change in revenues and plant, property and equipment. Improving upon the Jones Model, Dechow et al. (1995) adjusted the change in revenues by subtracting the change in receivables and denote this estimator the modified Jones model.

Despite such improvements, the modified Jones model suffers from omitted variable bias. Kothari et al. (2005) find one omitted variable in firm performance, and control for the effect of firm performance on accruals by comparing discretionary accruals of firms in the same two-digit SIC code in the same year and with similar performance. They improve the performance accruals model by controlling for the effect of growth, which may lead to misdiagnosis of discretionary accruals, as growth firms are more likely to have more volatile accruals. Revisiting performance matched accruals, Ecker et al. (2012) examine whether a firm's size may be a better control for firm characteristics as firm size is correlated with important firm attributes including business model, monitoring and information environment.

The development of these models have resulted in a stepwise improvement of discretionary accruals estimators, and subsequently, the ability for researchers to identify and filter out the non-discretionary portion of accruals. Nevertheless, these models still suffer from a degree of misspecification due to correlated omitted variables in samples with extreme financial performance. Other studies have attempted to match based on multiple determinants. For example, Collins et al. (2011) match firms by both firm performance and sales growth. Their results show that matching both by performance and sales growth can significantly improve the statistical power of discretionary accruals estimates.

Despite the advancement of model specification through matching, it is difficult to match on numerous firm characteristics (Tucker, 2012). Especially, given that earnings management is a function of multiple factors including firm size, monitoring and auditor quality, it would be very difficult to construct matched samples using matching alone. Propensity score matching overcomes this issue through a two stage process. In the first stage a selection model is specified by using a logit regression to regress an outcome dummy on a set of firm characteristics. In the second stage, the scores generated from the logit regression prediction is used to match firms.

3.3 Model Specification

To implement propensity score matching, we first build a choice model by considering firm characteristics that have been shown to affect earnings management in the prior literature. From this set of variables, we select those that have been shown to be correlated with accruals in the literature. The list of covariates that we consider is as follows: industry, year, ROA, age, size, book-to-market, leverage, institutional ownership, accounts receivable and Olhson's O-score.

We match firm years in the same two-digit SIC code, since industry is a proxy for a firm's business model. Beneish (1997) and Bagnoli and Watts (2000) show that firms are more likely to manipulate earnings when they expect their competitors to manage earnings. Matching firms in the same year controls for any time trends and common shocks. We also include a measurement of performance (return on assets) as it is correlated with accruals and omitting it would cause a misspecification (Kothari et al., 2005). In addition, Skinner and Sloan (2002) suggest that fast growing firms are more likely to manage earnings in order to avoid being penalized by missing expectations. Firm age reflects the firm's life cycle. For instance, firms have more incentives to manage earnings higher before an IPO (Teoh et al., 1998). Size has been shown to be another determinant of earnings manipulation due to the reduction in the variability of accruals with an increase in size (Rajgopal and Venkatachalam, 2011; Gu et al., 2005; Watts and Zimmerman, 1986), since larger firms tend to be older and established with decreased operational volatility. Furthermore, Ecker et al. (2012) show that size matched accruals quality lead to increased statistical power to detect earnings management. Moreover, the book to market ratio controls for firms which have additional growth opportunities and typically exhibit a higher degree of earnings manipulation (Teoh et al., 1998). Firms with higher leverage are also more likely to engage in income increasing accounting procedures (Watts and Zimmerman, 1986; DeFond and Jiambalvo, 1994). According to Nelson et al. (2003), firms are likely to manage earnings through accounts receivable, as management has substantial discretion over these accruals. Hence, larger accounts receivables are likely to be correlated with higher likelihood of earnings management. We also include firm riskiness approximated using the Ohlson's O-Score (Griffin and Lemmon, 2002).

3.3.1 Sample Selection

Data are collected from the COMPUSTAT Industrial Annual Research tapes and initially include all observations between 1962 and 2008. More specifically, we include all data items needed to compute total accruals and the determinants by which we define sorted subsets of accruals. Accordingly, we exclude all observations which lack any of the data items necessary to compute total accruals or estimate discretionary accruals using the Jones or modified Jones models. We also drop all observations that are missing any of the components to calculate the determinants by which we compute subsets of the accruals measures. In order to maintain sufficient degrees of freedom when estimating discretionary accruals using the Jones and modified Jones model, we also omit all observations for which there aren't at least 10 total firm-year observations in the same two-digit SIC code (Kothari et al., 2005). Any observation for which its lagged value is missing is also omitted from the dataset. As a last exclusion, we require the absolute value of total accruals scaled by total assets to be greater than one. The final dataset includes 193,532 observations. Total accruals are winsorized at the 1% and 99% level, and discretionary accruals measures are expressed as a proportion of lagged total assets.

In the following analyses, we report simulation results for 1500 random samples of 100 firms each[1] drawn without replacement from the full sample or from sorted subsamples based on our determinants. By sorting observations into subsamples based on extreme accounting positions, we can compare the degree of misspecification exhibited by our model to other models of discretionary accruals (Dechow et al., 1995; Kothari et al., 2005). The stratified subsets we consider for the specification and power tests are constructed by ranking

[1] We conduct our simulation tests with a much larger number of random samples (1500) than Kothari et al. (2005) (250) due to our much larger sample size (193,532 as opposed to approximately 123,000 observations, respectively.

all observations based on each determinant[2]. We then retain the upper and lower quartiles based on each determinant from each year and pool the annual fourth quartiles and the annual first quartiles together to form the subsamples.

3.3.2 Estimation of Discretionary Accruals

In order to separate total accruals into discretionary and non-discretionary components, we first define total accruals as

$$TA = \frac{\Delta NCCA - \Delta CL + \Delta CLD - DP}{AT_{t-1}} \tag{3.1}$$

where TA denotes total accruals, $\Delta NCCA$ is the change in non-cash current assets, ΔCL is the change in current liabilities, ΔCLD is the change in the current portion of long-term debt, DP is depreciation and amortization, and AT_{t-1} is

[2]The subset determinants considered include those proposed by Kothari et al. (2005) (i.e. book to market ratio, change in sales, E/P ratio and market capitalization), the O-score, the debt to total assets ratio (defined as COMPUSTAT items $\frac{DLTT+DLC}{AT}$), accounts receivable scaled by total assets, and property plant and equipment scaled by lagged total assets (items $\frac{PPENT}{AT_{t-1}}$).

lagged total assets. As our baseline benchmarks, we consider the Jones and the modified Jones models to estimate our discretionary accruals components. In order to estimate discretionary accruals using the Jones model, we run the following regression cross sectionally for firm years in each two-digit SIC code

$$TA = \beta_0 + \beta_1(\frac{1}{AT_{i,t-1}}) + \beta_2(\frac{\Delta SALES_{i,t}}{AT_{i,t-1}}) + \beta_3(\frac{PPE_{i,t}}{AT_{i,t-1}}) + \varepsilon_{i,t} \tag{3.2}$$

where $AT_{i,t-1}$ is lagged total assets, $\frac{\Delta SALES_{i,t}}{AT_{i,t-1}}$ is the change in total revenue scaled by lagged total assets, $\frac{PPE_{i,t}}{AT_{i,t-1}}$ is property, plant and equipment scaled by lagged total assets, and $\varepsilon_{i,t}$ is a normally distributed disturbance term. The non-discretionary component of total accruals is defined as

$$NDA = \hat{\beta}_0 + \hat{\beta}_1(\frac{1}{AT_{i,t-1}}) + \hat{\beta}_2(\frac{\Delta SALES_{i,t}}{AT_{i,t-1}}) + \hat{\beta}_3(\frac{PPE_{i,t}}{AT_{i,t-1}}) \tag{3.3}$$

where the $\hat{\beta}$s are the estimated coefficients from the regressions in equation (3.2) (Jones, 1991). Thus, discretionary accruals are defined as

$$DA = TA - NDA \tag{3.4}$$

or the residuals ($\varepsilon_{i,t}$) from the regressions in (3.2).

The modified version of the Jones model proposed by Dechow et al. (1995) includes an adjustment to the change in total revenues by accounting for the change in accounts receivable, due to the relative ease in managing earnings in the credit account rather than the cash account. Therefore, the non-discretionary component (and subsequent discretionary component) of total accruals in the modified Jones model is estimating by subtracting the change in accounts receivable from the change in total revenue scaled by lagged total assets (i.e. $\frac{\Delta SALES_{i,t} - \Delta AR_{i,t}}{AT_{i,t-1}}$).

Performance Matching and Matching on other Determinants

In order to take into account the non-linearity of firm profitability as a determinant in earnings management, Kothari et al. (2005) proposed a matching procedure based on the Jones model to determine another measure of discretionary accruals. The performance matched discretionary accruals measure is better able to mitigate type I errors when compared to the Jones and modified Jones models as well as variants that included ROA_t and ROA_{t-1} as regressors in (3.2). In order to estimate discretionary accruals using performance matching, we match each firm-year observation with another firm-year in the same two-digit SIC code that exhibits the closest level of profitability (ROA_t), hence a nearest neighbor approach is employed. The Jones model performance matched discretionary accrual for each firm-year is then defined as the difference between the Jones model discretionary accruals component of the observation and the Jones model discretionary accruals component of its matched observation. Matching observations based on performance has been shown to have desirable properties in earnings management research, as it has been shown to reduce the probably of making false positives in hypothesis testing for cases where the researchers variable of interest is correlated with performance.

In order to complement matching on the basis of performance, we also compute similar measures matched based on other determinants using the same procedure mentioned above. The additional determinants we matched by include the book to market ratio, Ohlson's O-score, the debt to total assets ratio and market capitalization. Our additional discretionary accruals measures are meant to assess the performance of models matched with other determinants which have been shown to be correlated with earnings management. Also, employing a matching technique will improve the quality of our discretionary accruals measures because it addresses any possible non-linearities our determinants may have with accruals that a linear regression framework may not be able to capture.

Propensity Score Matching

One downside to matching based on single determinants is that the resulting discretionary accruals measures do not take into account the combined effect that multiple determinants may have. Although the performance matched discretionary accrual is matched based on performance, year and industry, a matching procedure based on the total effect of performance with other variables shown to be correlated with earnings management is needed. However, multiple simultaneous matching can be cumbersome when the dimension of determinants increases, and finding a match may become noisy or even futile due to the lack of observations (Li and Prabhala, 2005).

In order to consider the combined effect of multiple determinants, we propose a matching approach based on the propensity to manage earnings using conventional proxies for earnings management. Our discretionary accruals measure is estimated in two stages - estimation and matching. First, we estimate logistic regressions run on 100 randomized subsamples of 1600 observations each from our full sample without replacement. We believe that this resampling technique improves the consistency of our estimates for the propensity score approach. More specifically, we run the following logistic regression for each subsample:

$$EM_i = \beta_0 + \sum_{n=1}^{k} \beta_n X_{i,n} + \varepsilon_i \tag{3.5}$$

where EM_i is a dummy variable equaling 0 if observation i does not meet our managed earnings criteria and 1 if it does, $X_{i,n}$ indicates determinant n for observation i and ε_i is a logistically distributed disturbance term. In order for an observation to be tagged as an EM observation, either ROA must be greater than or equal to 0% and less than or equal to 0.05% in a given firm-year ($0\% \leq ROA_t \leq 0.05\%$), or the change in ROA must be equal to or greater than 0% or less than or equal to 0.05% in a given firm-year ($0\% \leq \Delta ROA_t \leq 0.05\%$) or a firm was required to restate its earnings in a given year as indicated in the Government Accountability Office (GAO) database or Audit Analytics database (AA). The choice set of determinants include the O-Score, the debt to total assets ratio, ROA, the book to market ratio, operating cash flows, size measured by market capitalization, accounts receivable and the E/P ratio.

By estimating (3.5) and averaging out the coefficients from the 100 iterations, we obtain the conditional probability (a.k.a. propensity score) that each firm-year engages in earnings management based on our selected determinants. For each observation in the full sample and stratified subsamples, we compute the propensity to manage earnings. We then match each observation by propensity score with its nearest neighbor in the same year and two digit SIC industry code. The propensity score matched discretionary accrual for firm-year i is defined as its Jones model estimated discretionary accrual minus the matched Jones model estimated discretionary accrual.

Matching based on the propensity to manage earnings has several advantages over matching based on single determinants. First, we overcome the multidimensional matching problem pointed out above by exploiting conventional proxies for earnings management in a binomial dependent variable regression setting. Thus, we are always able to find a match since our propensity score is, technically, one-dimensional. Our regression framework is also able to accommodate for cross correlations amongst determinants. However, one drawback could be the inability of the propensity score to address any potential joint non-linearities between the determinants and accruals.

Table 3.1 presents the results of the logistic regressions estimated in equation (3.5). We estimate the logistic regressions twice with two different vectors of determinants. In the first specification, we include all of the determinants in our choice set. The results (presented in column 1 under "All Determinants") indicate that the O-score, debt to assets ratio, ROA and book to market ratio have a significant effect on the probability of managing earnings, and that the combined effects of cash flows, firm size, accounts receivable and the E/P ratio are confounded by the significant regressors. Hence, our findings in the first logistic regressions suggest that we specify a more prudential model based on the vector of significant regressors, the results of which are indicated in column 2, labeled "Propensity Score."

As can be seen, the coefficient on the debt to assets ratio and the book to market ratio strongly resemble those we measured in the first regression, while the coefficients on ROA and the O-score slightly change and gain in statistical significance. This suggests that much of the covariation between our insignificant regressors and proxies for earnings management from the first model specification is obfuscated by the O-score and return on assets. In order to be certain that our model is not too parsimonious, we also estimated additional specifications including the regressors from column 2 and the insignificant regressors one at a time to be sure the correlation between those determinants and our earnings management proxies are absorbed by the O-score, debt to assets ratio, ROA and book

Table 3.1: Propensity Score - Logistic Regressions

This table presents the results of logistic regressions on various determinants correlated with discretionary accruals measures. For each specification, we estimate $EM_i = \beta_0 + \sum_{n=1}^{k} \beta_n X_{i,n} + \varepsilon_i$ on 100 randomized subsamples of 1600 observations each without replacement from our full sample, where EM_i is a dummy variable equaling 1 if observation i meets our managed earnings criteria, $X_{i,n}$ indicates determinant n for observation i and ε_i is a logistically distributed disturbance term. Average coefficients, t-statistics and p-values are reported. The number of observations in our full sample equals 193,532 and are obtained from the COMPUSTAT Industrial Annual tapes. An observation is tagged as an EM observation if it meets one of the following criteria: either ROA must be greater than or equal to 0% and less than or equal to 0.05% in a given firm-year ($0\% \leq ROA_t \leq 0.05\%$), or the change in ROA must be equal to or greater than 0% or less than or equal to 0.05% in a given firm-year ($0\% \leq \Delta ROA_t \leq 0.05\%$) or a firm was required to restate its earnings in a given year as indicated in the Government Accountability Office (GAO) database or Audit Analytics database (AA).

	All Determinants (1)	Propensity Score (2)
$OSCR$	-1.24***	-1.35***
	(-5.05)	(-5.66)
	[0.00]	[0.00]
DtA	1.57***	1.56***
	(5.99)	(6.12)
	[0.00]	[0.00]
ROA	2.61***	2.85***
	(5.52)	(6.17)
	[0.00]	[0.00]
BtM	.09*	.09*
	(1.88)	(1.90)
	[0.06]	[0.06]
$CASHFLOW$	.018	
	(0.52)	
	[0.60]	
$SIZE$	0.00	
	(0.07)	
	[0.95]	
AR	.03	
	(0.07)	
	[0.95]	
EP	.53	
	(1.36)	
	[0.17]	
$Constant$	0.00	0.06
	(0.07)	(0.65)
	[0.94]	[0.52]

T-statistics in parentheses

P-values in brackets - * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

to market ratio. Our findings indicate that our model in column 2 is well defined. Therefore, we define the propensity to manage earnings for each observation as

$$PS_i = .06 - 1.35 * OScore_i + 1.56 * DtA_i + 2.85 * ROA_i + 0.086 * BtM_i. \tag{3.6}$$

The negative coefficient on the O-score is in line with DeAngelo et al. (1994) who show that firms approaching default are more closely monitored by creditors and auditors, reducing their capacity to manage earnings. The positive coefficient on ROA reflects the additional capacity to manage earnings when firms are more profitable (McNichols, 2000; Bowen et al., 2008). Moreover, the positive coefficient on book to market indicate that managers with more growth potential are likely to manipulate earnings (Teoh et al., 1998) while the positive coefficient on debt to assets exhibits that firms which are more leveraged have more incentive to manage earnings to avoid debt covenant violations (Bowen et al., 2008) and to reduce debt-related costs (Minton and Schrand, 1999).

3.3.3 Summary Statistics of Discretionary Accruals

In table 3.2 below, we report the mean, median, standard deviation, and upper/lower quartiles of total accruals and discretionary accruals estimated using the Jones model, modified Jones model, matching based on ROA, book to market, the O-score, debt to assets, and market capitalization as well as propensity score matched discretionary accruals. First, it is noteworthy to point out that all of our accruals measures (except for total accruals scaled by total assets) exhibit mean levels close to zero. This is due to the fact that discretionary accruals are measured as the residuals from regression (3.5). Moreover, note that the standard deviations of our accruals measures increase with matching. This is expected if we assume that each observation's discretionary accruals component is uncorrelated with its match on average, thus the standard deviation would have to increase by construction.

Table 3.3 presents the means of the abovementioned accruals measures for each of our subsamples. As we can see, many of the accruals measures exhibit non-zero means when applied to subsamples based on our selected determinants. However, it is noteworthy that the means of propensity score matched discretionary accruals are consistently closer to zero when compared with the other accruals measures. Hence, it seems that the bias of non-zero accruals is mitigated with the propensity score matched method, meaning we can expect our measure to perform well in the specification test conducted below.

Total accruals themselves are likely to exhibit a degree of correlation which can lead to serially correlated estimates of discretionary accruals. The serial correlation in total accruals is attributable to the reverting nature of the accruals account following management decisions. Discretionary accruals estimates that exhibit a zero mean and no serial correlation are

Table 3.2: Descriptive Statistics for Discretionary Accrual Measures

This table reports the mean, median, standard deviation, and upper/lower quartiles of total accruals and discretionary accruals estimated using the Jones model, modified Jones model, matching models based on ROA, book to market, the O-score, debt to assets, and market capitalization as well as propensity score matched discretionary accruals. Data are collected from the COMPUSTAT Industrial Annual Research tapes and include observations between 1962 and 2008. Accordingly, we exclude all observations which lack any of the data items necessary to compute total accruals, estimate discretionary accruals using the Jones or modified Jones models, or calculate our determinants. We also omit all observations for which there aren't at least 10 total firm-year observations in the same two-digit SIC code (Kothari et al., 2005). Any observation for which its lagged value is missing is also omitted from the dataset. We exclude observations for which the absolute value of total accruals scaled by total assets is less than one. The final dataset includes 193,532 observations. Total accruals are winsorized at the 1% and 99% level, and discretionary accruals measures are expressed as a proportion of lagged total assets.

	Mean	Standard Deviation	Lower Quartile	Median	Upper Quartile
Total Accruals	-0.03516	0.12248	-0.08749	-0.03735	0.01342
Jones Model	-0.00019	0.09927	-0.04387	0.00113	0.04437
Modified Jones Model	-0.00016	0.10261	-0.04563	0.00028	0.04488
Jones matched on ROA	0.00009	0.13630	-0.06787	0.00000	0.06778
Propensity Score matched Jones	-0.00021	0.13726	-0.06876	0.00000	0.06833
Jones matched on BtM	-0.00009	0.14002	-0.07006	0.00000	0.07001
Jones matched on OSCR	-0.00028	0.13841	-0.06884	0.00000	0.06864
Jones matched on DtA	-0.00135	0.13949	-0.07097	0.00000	0.06853
Jones matched on Size	0.00010	0.14136	-0.06942	0.00000	0.06942

Table 3.3: Means of Discretionary Accrual Measures

This table reports the mean of total accruals and discretionary accruals estimated using the Jones model, modified Jones model, matching models based on ROA, book to market, the O-score, debt to assets, and market capitalization as well as propensity score matched discretionary accruals sorted based on stratified subsets of the full sample. Data are collected from the COMPUSTAT Industrial Annual Research tapes and include observations between 1962 and 2008. Accordingly, we exclude all observations which lack any of the data items necessary to compute total accruals, estimate discretionary accruals using the Jones or modified Jones models, or calculate our determinants. We also omit all observations for which there aren't at least 10 total firm-year observations in the same two-digit SIC code (Kothari et al., 2005). Any observation for which its lagged value is missing is also omitted from the dataset. We exclude observations for which the absolute value of total accruals scaled by total assets is less than one. The final dataset includes 193,532 observations. Total accruals are winsorized at the 1% and 99% level, and discretionary accruals measures are expressed as a proportion of lagged total assets.

	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size	Low Size
Total Accruals	-0.0374	-0.0478	0.0146	-0.0773	-0.0150	-0.0898	-0.0393	-0.0518
Jones Model	-0.0015	-0.0096	0.0078	-0.0101	0.0049	-0.0301	0.0012	-0.0111
Modified Jones Model	-0.0037	-0.0079	0.0182	-0.0184	0.0066	-0.0336	0.0017	-0.0135
Jones matched on ROA	-0.0021	0.0010	0.0030	0.0006	-0.0032	-0.0015	-0.0072	0.0007
Propensity Score matched Jones	-0.0033	0.0003	0.0039	-0.0005	-0.0029	-0.0048	-0.0087	0.0009
Jones matched on BtM	0.0002	-0.0005	0.0056	-0.0052	0.0034	-0.0219	-0.0022	-0.0040
Jones matched on OSCR	-0.0039	0.0007	0.0045	-0.0029	-0.0000	-0.0108	-0.0058	-0.0004
Jones matched on DtA	-0.0034	-0.0078	0.0058	-0.0097	0.0040	-0.0290	-0.0006	-0.0091
Jones matched on Size	0.0010	-0.0085	0.0065	-0.0061	0.0045	-0.0252	-0.0003	-0.0002

	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Total Accruals	-0.0775	-0.0180	-0.0455	-0.0357	0.0015	-0.0643	-0.0574	-0.0332
Jones Model	-0.0244	0.0100	-0.0017	-0.0007	0.0140	-0.0076	-0.0003	-0.0001
Modified Jones Model	-0.0269	0.0114	-0.0028	0.0001	0.0177	-0.0092	-0.0006	-0.0000
Jones matched on ROA	0.0013	-0.0017	0.0051	-0.0016	0.0125	-0.0034	-0.0027	0.0072
Propensity Score matched Jones	0.0012	-0.0034	0.0028	-0.0009	0.0144	-0.0043	-0.0027	0.0070
Jones matched on BtM	-0.0150	0.0059	0.0052	-0.0033	0.0149	-0.0066	-0.0011	0.0025
Jones matched on OSCR	-0.0012	0.0003	0.0076	-0.0056	0.0142	-0.0049	-0.0009	0.0036
Jones matched on DtA	-0.0225	0.0093	-0.0008	-0.0047	0.0128	-0.0078	-0.0012	-0.0034
Jones matched on Size	-0.0176	0.0082	0.0007	-0.0003	0.0144	-0.0071	-0.0004	0.0022

considered to be well defined due to their ability to filter out non-discretionary accruals. Therefore, accruals measures that exhibit a lower degree of serial correlation are also expected to reject the null hypothesis of no earnings management less often than other models. In order to estimate the degree in which each accruals measure is serially correlated, we run regressions of each accruals measure on its lagged value for each year and average out the standard errors of our estimated coefficients as proposed by Fama and MacBeth (1973) and include a Newey and West (1987) correction for autocorrelation and heteroskedasticity of the error term. The degree of serial correlation is equivalent to the average estimated coefficient on the lagged regressor. Table 3.4 presents serial correlations for the full sample and stratified subsets according to our declared determinants.

Comparing the serial correlation of the discretionary accruals estimates with those of total accruals scaled by total assets, we see that our matchings models show a much lower

Table 3.4: Serial Correlation of Accruals and Discretionary Accruals Measures

The table reports estimates of the serial correlation of total accruals and our discretionary accruals measures. To obtain an estimates of serial correlation, we run regressions of each accruals measure on its lagged value for each year and average out the standard errors of our estimated coefficients as proposed by Fama and MacBeth (1973) and include a Newey and West (1987) correction for autocorrelation and heteroskedasticity of the error term. The degree of serial correlation is equivalent to the average estimated coefficient on the lagged regressor. We present serial correlations for the full sample and stratified subsets according to our declared determinants. Data are collected from the COMPUSTAT Industrial Annual Research tapes and include observations between 1962 and 2008. Accordingly, we exclude all observations which lack any of the data items necessary to compute total accruals, estimate discretionary accruals using the Jones or modified Jones models, or calculate our determinants. We also omit all observations for which there aren't at least 10 total firm-year observations in the same two-digit SIC code (Kothari et al., 2005). Any observation for which its lagged value is missing is also omitted from the dataset. We exclude observations for which the absolute value of total accruals scaled by total assets is less than one. The final dataset includes 193,532 observations. Total accruals are winsorized at the 1% and 99% level, and discretionary accruals measures are expressed as a proportion of lagged total assets.

	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size	Low Size
Total Accruals	0.1550	0.0582	0.1749	0.2162	-0.0273	0.1429	0.0256	0.3292	0.0581
Jones Model	0.0087	-0.0440	0.0256	0.0650	-0.0871	-0.0029	-0.0581	0.1538	-0.0490
Modified Jones Model	0.0285	-0.0312	0.0529	0.0738	-0.0684	0.0095	-0.0382	0.1716	-0.0318
Jones matched on ROA	0.0005	-0.0189	0.0051	0.0343	-0.0494	-0.0007	-0.0305	0.0624	-0.0392
Propensity Score matched Jones	-0.0057	-0.0319	0.0135	0.0342	-0.0612	-0.0036	-0.0288	0.0695	-0.0443
Jones matched on BtM	-0.0030	-0.0256	-0.0019	0.0360	-0.0540	-0.0026	-0.0429	0.0496	-0.0358
Jones matched on OSCR	-0.0009	-0.0279	0.0020	0.0298	-0.0535	-0.0016	-0.0254	0.0586	-0.0286
Jones matched on DtA	0.0053	-0.0103	0.0175	0.0385	-0.0409	-0.0010	-0.0247	0.0571	-0.0273
Jones matched on Size	-0.0031	-0.0049	0.0076	0.0318	-0.0547	-0.0158	-0.0472	0.0821	-0.0291

	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Total Accruals	0.0630	0.2160	0.1754	0.0924	0.1271	0.1270	0.2557	0.0791
Jones Model	-0.0437	0.0526	0.0246	-0.0247	0.0062	0.0286	0.1019	-0.0308
Modified Jones Model	-0.0319	0.0785	0.0452	-0.0083	0.0294	0.0374	0.1102	-0.0040
Jones matched on ROA	-0.0238	0.0251	0.0021	-0.0070	-0.0108	0.0122	0.0440	-0.0299
Propensity Score matched Jones	-0.0359	0.0217	-0.0082	-0.0202	-0.0089	0.0027	0.0473	-0.0376
Jones matched on BtM	-0.0256	0.0122	0.0026	-0.0216	-0.0026	0.0035	0.0513	-0.0330
Jones matched on OSCR	-0.0345	0.0226	0.0009	-0.0201	-0.0092	0.0132	0.0526	-0.0244
Jones matched on DtA	-0.0316	0.0289	0.0058	-0.0150	0.0048	0.0149	0.0443	-0.0162
Jones matched on Size	-0.0282	0.0202	0.0035	-0.0198	-0.0091	0.0126	0.0525	-0.0481

degree of serial correlation with respect to the full sample and the stratified subsets. Moreover, the propensity score matched discretionary accrual consistently exhibits a degree of serial correlation close to zero.

3.4 Specification Test

In order to determine if our measure of discretionary accruals over- or underrejects the null hypothesis of no earnings management, we run simulations using 1500 random samples of 100 firm-years each for the full sample and each stratified subsample and conduct a t-test on the mean level of discretionary accruals for each model. The specification test is designed to evaluate the percentage of random samples for which each respective discretionary accruals measure detects a false positive on the null hypothesis of no earnings management (type I error).

For our two panels of specification tests in tables 3.5 and

3.6, we present the results of specification tests with alternative hypotheses of negative and positive discretionary accruals, respectively. Hence, rejection rates with respect to the one-sided t-test for negative and postive discretionary accruals at the 5% confidence level are presented (i.e. we reject the null of no earnings management if the p-value computed for any given subsample does not exceed 5%. Moreover, since we are drawing our samples randomly we expect average discretionary accruals to be equal to zero, since we do not expect a random draw to systematically exhibit active earnings management (Kothari et al., 2005). Moreover, given the large amount of iterations, we are able to increase the probability that the null hypothesis of no earnings management (i.e. our results are unlikely to be spurious).

Following (Kothari et al., 2005), we assume that rejection rates in the interval between 2% and 8% correspond to a well-specified model. More specifically, we interpret a rejection rate below 2% as biased too greatly in favor of the null hypothe-

Table 3.5: Specification Test - H_a: Discretionary Accruals Less Than 0

This table presents rejection rates with respect to the one-sided t-test for negative discretionary accruals at the 5% confidence level (i.e. we reject the null of no earnings management if the p-value computed for any given subsample does not exceed 5%. We run simulations using 1500 random samples of 100 firm-years each without resampling for the full sample and each stratified subsample and conduct a one-sided t-test on the mean level of discretionary accruals for each model. We interpret a rejection rate below 2% as biased too greatly in favor of the null hypothesis, whereas a rejection rate over 8% indicates that the null hypothesis of no earnings management is rejected too often, and define a failure as the number of times for which the rejection rate of each accruals model falls below 2% or above 8% for the full sample and each stratified subset we define in our tests.

	Failure	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	10	5.1	5.9	19.4	0.6	25.2	0.9	76.8	2.7
Modified Jones Model	11	5.3	9.5	15.6	0.1	50.7	0.8	83.3	2.5
Jones matched on ROA	5	5.2	5.5	5.1	3.1	4.5	8.5	6.3	17.2
Propensity Score matched Jones	5	4.7	7.7	5.7	2.3	5.0	8.6	8.4	20.6
Jones matched on BtM	8	4.8	4.5	6.2	2.2	9.1	2.0	36.4	7.1
Jones matched on OSCR	7	5.1	8.5	5.3	2.1	7.2	5.4	14.8	14.1
Jones matched on DtA	12	6.6	8.9	13.1	1.1	15.2	1.9	52.5	5.1
Jones matched on Size	8	5.4	3.9	13.3	1.5	9.7	2.5	40.6	5.0

	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	22.7	57.9	0.5	6.3	6.1	0.1	20.9	4.7	5.5
Modified Jones Model	27.3	62.6	0.4	7.3	5.2	0.0	26.7	5.7	5.6
Jones matched on ROA	4.9	3.9	7.0	1.7	6.7	0.7	7.4	7.6	1.7
Propensity Score matched Jones	4.8	4.1	7.3	3.3	5.8	0.7	8.9	7.8	2.0
Jones matched on BtM	8.9	20.0	1.4	1.6	7.7	0.3	14.0	6.5	3.5
Jones matched on OSCR	5.1	5.7	4.7	1.4	9.5	0.6	9.7	6.3	2.5
Jones matched on DtA	13.4	35.7	0.6	4.8	9.7	0.6	15.5	5.6	6.9
Jones matched on Size	5.4	24.9	1.1	4.7	5.7	0.3	13.3	5.5	3.9

Table 3.6: Specification Test - H_a: Discretionary Accruals Greater Than 0

This table presents rejection rates with respect to the one-sided t-test for positive discretionary accruals at the 5% confidence level (i.e. we reject the null of no earnings management if the p-value computed for any given subsample does not exceed 5%. We run simulations using 1500 random samples of 100 firm-years each without resampling for the full sample and each stratified subsample and conduct a one-sided t-test on the mean level of discretionary accruals for each model. We interpret a rejection rate below 2% as biased too greatly in favor of the null hypothesis, whereas a rejection rate over 8% indicates that the null hypothesis of no earnings management is rejected too often, and define a failure as the number of times for which the rejection rate of each accruals model falls below 2% or above 8% for the full sample and each stratified subset we define in our tests.

	Failure	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	10	3.6	3.7	0.8	15.9	0.3	13.0	0.0	6.4
Modified Jones Model	10	3.6	2.3	1.2	41.3	0.1	16.5	0.0	6.3
Jones matched on ROA	4	4.2	3.3	4.3	6.2	5.1	3.3	3.5	0.7
Propensity Score matched Jones	3	4.7	2.9	4.5	7.7	4.3	3.1	3.5	0.4
Jones matched on BtM	7	4.8	5.2	4.5	8.7	2.5	7.9	0.1	3.0
Jones matched on OSCR	5	4.1	2.5	4.8	8.3	3.3	5.9	1.2	1.4
Jones matched on DtA	10	3.7	2.5	1.4	9.6	0.9	9.2	0.0	5.2
Jones matched on Size	9	4.5	5.9	0.7	10.8	1.9	10.4	0.3	3.8

	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	0.5	0.0	31.8	3.9	4.3	32.6	0.6	5.2	5.5
Modified Jones Model	0.4	0.0	35.4	3.0	5.2	41.3	0.5	4.5	5.3
Jones matched on ROA	5.6	6.2	3.3	10.3	3.6	20.1	3.2	2.2	11.2
Propensity Score matched Jones	4.9	6.2	2.6	6.3	3.8	23.9	2.2	3.4	9.9
Jones matched on BtM	3.2	0.5	12.7	9.3	3.5	23.3	1.5	4.1	7.4
Jones matched on OSCR	4.9	5.3	4.6	13.0	2.2	22.7	2.3	4.5	7.9
Jones matched on DtA	1.1	0.3	16.5	5.0	2.5	21.7	1.2	4.7	3.5
Jones matched on Size	3.7	0.5	16.6	5.6	4.3	22.3	1.4	4.2	6.1

sis, whereas a rejection rate over 8% indicates that the null hypothesis of no earnings management is rejected too often. In addition to presenting the rejection rates for each subsample, we also provide summary information that could be useful for an immediate comparison between the performance of the models in our specification test. In particular, we define a *failure* as the number of times for which the rejection rate of each accruals model falls below 2% or above 8% for the full sample and each stratified subset we define in our tests. For example, the number of failures of the Jones model matched by size in table 3.5 is 8 out of 17.

In table 3.5, it is first noteworthy that the number of failures is high for the non-matched discretionary accruals measures. With the exception of the Jones model matched on book to market (12 failures), the Jones and modified Jones models produce 10 and 11 failures out of 17, respectively. Moreover, the performance matched and propensity score matched Jones models produce the least amount of failures, at 5 each for all of our subsamples. Jones matching based on book to market (8 failures), O-score (7 failures) and size (8 failures) all under- and overreject the null hypothesis of no earnings management when the alternative is an earnings decreasing activity. In this case, the propensity score matched accrual measure performs just as well as the performance matched accrual.

Table 3.6 presents results of our specification tests where the alternative hypothesis is that the mean discretionary accrual is significantly greater than zero for each subset. Remarkably, the propensity score matched discretionary accruals model fails only 3 out of 17 times. Performance matching fails 4 times, while matching based on other determinants all produce a higher number of failures. The Jones and the modified Jones models are again the least efficient in this part of our specification test, producing 10 failures each.

Our results indicate that our propensity score matched discretionary accruals model is better able to maintain the null hypothesis of no earnings management when the alternative hypothesis is an earnings increasing activity. Moreover, it is also as well defined as the performance matched Jones model when the alternative hypothesis is negative discretionary accruals. Overall, our statistical evidence shows that our model is the best specified. Hence, the earnings management researcher should implement the propensity score matching approach in order to mitigate type I errors or false positives. An added feature of our estimator is its ability to address the correlation between many determinants and accruals, making it appropriate for a wide array of earnings management tests.

3.5 Power Test

In this section, we test the statistical power of our earnings management models and present the results in tables 3.7 and 3.9. In simulation tests of statistical power, we aim to measure the percentage of times in which the null hypothesis of no earnings management is rejected when actually false (hence, we test for type II errors). As it has been done for the specification, we conduct simulations on 1500 random samples of 100 firms each. Similarly for the test on specification, the power is computed by performing a t-test for each random sample of mean discretionary accruals and counting the number of times each respective p-value falls below the significance threshold of 5%. Likewise to the specification test, discretionary accruals are expected to be zero when we draw random samples. Therefore, in order to be sure that the null hypothesis of no earnings management is false, we artificially introduce earnings manipulation by adding seeds of $\pm1\%, \pm2\%, \pm4\%$ and $\pm10\%$ to total accruals before computing discretionary accruals with each model. We also follow Kothari et al. (2005) and assume that earnings management is 50% revenue based, so we also add half of the seed to the difference in sales and the difference in accounts receivable in the Jones model before estimating each model. Thus, it is certain that the randomly selected firm years serially manage earnings, and we can measure the ability of our models to detect this by conducting one sided t-tests for each random subsample. Adding seeds of both signs to total accruals allows us to test the alternative hypotheses of negative and positive earnings manipulation, respectively.

Tables 3.7 and 3.8 report the results of the power test with the alternative hypothesis of negative discretionary accruals. First, we immediately notice that the power of the Jones model and the modified Jones model is much higher for the full sample and all stratified subsets. For example with a -1% seed, the mean rate of rejection of the null hypothesis is 37.3% and 38.9%, respectively. The rejection rates for the matched accruals models average between 22-27% for the 1% seed. Decreasing accruals with a 2% and 4% seed further reflects the pattern exhibited for the -1% seed, albeit with higher rejection rates. All accruals models reject the null hypothesis of no earnings management with 100% probability on average when the seed is decreased by 10%. Also, note that the performance matched Jones model rejects the null hypothesis less

Table 3.7: Power Test - H_a: Discretionary Accruals Less Than 0 (-1% and -2% seeds)

In this table, we measure the percentage of times in which the null hypothesis of no earnings management is rejected when firms have engaged in revenue decreasing activity (hence, we test for type II errors). We conduct simulations on 1500 random samples of 100 firms each without resampling. The power is computed by performing a one-sided t-test for each random sample of mean discretionary accruals and counting the number of times each respective p-value falls below the significance threshold of 5%. In order to simulate the earnings decreasing activity, we artificially introduce earnings manipulation by subtracting seeds of 1%, 2%, 4% and 10% from total accruals before computing discretionary accruals with each model. We also follow Kothari et al. (2005) and assume that earnings management is 50% revenue based, so we also subtract half of the seed to the difference in sales and the difference in accounts receivable in the Jones model before estimating each model.

Panel A: Seed = -1%

	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	37.36	27.1	39.9	43.1	6.6	57.4	13.5	93.9	35.0
Modified Jones Model	38.97	26.1	49.3	36.4	1.3	80.3	10.1	96.1	32.3
Jones matched on ROA	18.90	15.9	26.7	13.1	11.3	14.5	27.2	16.9	48.9
Prop Score matched Jones	20.58	18.8	29.4	13.1	10.7	15.5	26.8	21.6	53.9
Jones matched on BtM	21.61	17.6	21.0	13.3	7.7	25.7	12.7	58.9	26.9
Jones matched on OSCR	20.29	18.5	30.5	14.0	9.5	20.5	21.9	33.7	41.4
Jones matched on DtA	27.24	21.0	27.5	26.3	7.7	36.3	11.7	73.9	23.7
Jones matched on Size	22.16	17.9	17.5	28.5	7.2	27.9	10.5	63.2	28.7

	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	51.3	84.6	5.6	30.5	27.7	1.7	61.3	37.1	18.9
Modified Jones Model	57.0	86.8	3.9	33.7	24.9	0.7	67.3	38.0	18.3
Jones matched on ROA	14.0	13.6	23.6	10.1	19.4	3.5	26.6	29.5	6.8
Prop Score matched Jones	14.5	12.4	29.0	14.1	20.0	2.2	29.2	30.9	8.0
Jones matched on BtM	22.4	42.2	9.7	10.6	23.5	2.6	35.5	26.3	10.9
Jones matched on OSCR	14.3	16.8	19.3	7.3	30.4	3.7	30.4	23.0	9.7
Jones matched on DtA	32.4	62.2	5.7	19.3	28.1	2.5	39.1	25.7	20.1
Jones matched on Size	13.7	45.3	6.8	16.3	18.5	2.5	36.1	24.8	11.3

Panel B: Seed = -2%	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	67.02	65.5	79.3	74.5	25.9	85.3	53.3	98.4	87.9
Modified Jones Model	64.77	62.3	83.1	68.0	6.3	96.4	45.1	99.1	83.9
Jones matched on ROA	41.65	41.7	54.9	31.7	27.1	36.0	58.1	37.3	82.2
Prop Score matched Jones	43.40	42.4	57.2	31.1	28.4	39.7	60.8	41.7	86.1
Jones matched on BtM	43.63	41.1	48.0	32.1	22.3	49.4	38.6	77.7	63.3
Jones matched on OSCR	43.05	41.9	60.1	29.8	25.6	45.2	51.0	51.9	79.8
Jones matched on DtA	49.93	44.3	58.7	49.9	21.7	63.2	36.7	88.6	59.5
Jones matched on Size	43.68	40.0	39.5	52.9	21.2	50.7	32.8	79.9	69.3
	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	81.5	96.5	31.1	63.5	61.1	12.4	92.2	80.3	50.5
Modified Jones Model	84.4	97.4	24.9	65.3	55.7	7.4	94.1	79.9	47.9
Jones matched on ROA	33.0	28.9	53.7	28.1	41.2	13.5	55.9	65.3	19.5
Prop Score matched Jones	33.1	31.3	61.1	30.7	40.3	10.3	58.5	65.5	19.5
Jones matched on BtM	44.7	65.2	27.9	25.5	44.9	10.3	65.6	55.3	29.7
Jones matched on OSCR	35.7	34.9	48.0	20.9	53.5	10.5	59.7	58.1	25.4
Jones matched on DtA	55.3	79.0	20.8	39.5	49.6	11.3	69.7	58.9	42.1
Jones matched on Size	32.4	70.1	24.4	34.9	37.2	8.9	66.3	53.8	28.2

Table 3.8: Power Test - H_a: Discretionary Accruals Less Than 0 (-4% and -10% seeds)

In this table, we measure the percentage of times in which the null hypothesis of no earnings management is rejected when firms have engaged in revenue decreasing activity (hence, we test for type II errors). We conduct simulations on 1500 random samples of 100 firms each without resampling. The power is computed by performing a one-sided t-test for each random sample of mean discretionary accruals and counting the number of times each respective p-value falls below the significance threshold of 5%. In order to simulate the earnings decreasing activity, we artificially introduce earnings manipulation by subtracting seeds of 4% and 10% from total accruals before computing discretionary accruals with each model. We also follow Kothari et al. (2005) and assume that earnings management is 50% revenue based, so we also subtract half of the seed to the difference in sales and the difference in accounts receivable in the Jones model before estimating each model.

Panel A: Seed = -4%	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	96.26	99.0	99.9	98.4	83.3	99.9	98.3	100.0	100.0
Modified Jones Model	93.09	98.4	99.9	96.7	53.8	99.9	96.3	100.0	100.0
Jones matched on ROA	83.42	89.7	96.1	72.4	75.5	82.5	96.2	80.2	99.5
Prop Score matched Jones	83.71	88.8	95.5	74.5	75.5	85.5	95.8	82.6	99.8
Jones matched on BtM	84.54	88.8	94.3	71.2	69.5	89.6	88.1	97.1	96.9
Jones matched on OSCR	84.00	88.5	96.9	74.1	74.5	86.8	93.9	89.5	99.4
Jones matched on DtA	87.93	89.8	96.0	87.1	71.4	94.3	88.7	99.0	96.8
Jones matched on Size	84.34	87.3	88.5	88.8	66.3	89.9	85.9	97.9	98.9

	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	99.5	99.9	95.2	98.5	99.0	69.9	99.9	99.9	95.9
Modified Jones Model	99.5	99.9	91.8	98.4	98.3	55.5	100.0	99.9	94.3
Jones matched on ROA	80.3	73.1	94.3	77.9	86.8	56.0	94.8	98.7	63.9
Prop Score matched Jones	78.1	73.3	96.8	82.2	86.1	49.5	95.3	98.1	65.7
Jones matched on BtM	87.2	93.7	82.0	74.1	89.1	47.1	97.5	96.9	74.2
Jones matched on OSCR	80.3	74.7	93.9	69.5	90.9	50.5	95.5	97.6	71.5
Jones matched on DtA	92.1	98.1	77.6	85.3	91.2	49.1	97.7	96.4	84.1
Jones matched on Size	75.7	93.6	81.7	84.3	82.7	45.6	97.2	95.8	73.7

Panel B: Seed = -10%

	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Modified Jones Model	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on ROA	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Prop Score matched Jones	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on BtM	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on OSCR	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on DtA	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on Size	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	Low Size	**High O-Score**	**Low O-Score**	**High DtA**	**Low DtA**	**High AR**	**Low AR**	**High PPE**	**Low PPE**
Jones Model	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Modified Jones Model	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on ROA	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Prop Score matched Jones	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on BtM	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on OSCR	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on DtA	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on Size	100.0	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0

frequently than the propensity score matched Jones model for the -1%, -2% and -4% seeds. Moreover, the propensity score matched Jones also exhibits similar rejection rates compared to the other matched models on average.

We also report the results of the power test with the alternative hypothesis of positive discretionary accruals in tables 3.9 and 3.10. The Jones model and the modified Jones model still reject the null hypothesis more often than the other models for the full sample and all stratified subsets. With a 1% seed, the mean rate of rejection of the null hypothesis is 27.1% and 28.38%, respectively. The rejection rates for the matched accruals models average between 16-18% for the 1% seed. Increasing the seed also increases the rejection rates for every model similarly as in table 3.7. Most of the models also converge to a 100% rejection rate when the seed is increased to 10%. Moreover, the propensity score matched model underrejects the null hypothesis when compared to the performance matched model for each seed, indicating that the performance matched model performs better in power tests with the alternative hypothesis of positive discretionary accruals.

All of our accruals measures exhibit the tradeoff behavior noted earlier when we consider the power test versus the specification test. The Jones and the modified Jones models overwhelmingly outperform the other models for both alternative hypotheses in terms of statistical power. When comparing the performance matched Jones with the propensity score matched model, we document an interesting tradeoff. The propensity score matched model mitigates type II errors when the alternative hypothesis is negative mean discretionary accruals, and the performance matched model rejects the null more often when the alternative hypothesis is positive discretionary accruals.

Table 3.9: Power Test - H_a: Discretionary Accruals Greater Than 0 (1% and 2% seeds)

In this table, we measure the percentage of times in which the null hypothesis of no earnings management is rejected when firms have engaged in revenue increasing activity (hence, we test for type II errors). We conduct simulations on 1500 random samples of 100 firms each without resampling. The power is computed by performing a one-sided t-test for each random sample of mean discretionary accruals and counting the number of times each respective p-value falls below the significance threshold of 5%. In order to simulate the earnings increasing activity, we artificially introduce earnings manipulation by adding seeds of 1%, 2%, 4% and 10% from total accruals before computing discretionary accruals with each model. We also follow Kothari et al. (2005) and assume that earnings management is 50% revenue based, so we add half of the seed to the difference in sales and the difference in accounts receivable in the Jones model before estimating each model.

Panel A: Seed = 1%

	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	27.10	24.7	27.3	5.8	42.6	4.5	50.3	0.0	50.5
Modified Jones Model	28.38	24.3	18.7	6.7	72.5	0.9	55.4	0.0	50.9
Jones matched on ROA	18.55	17.3	15.3	15.2	21.9	17.2	13.7	13.9	9.3
Prop Score matched Jones	17.93	17.1	14.1	13.9	24.2	16.1	16.1	9.1	6.9
Jones matched on BtM	18.07	17.7	21.1	12.5	26.2	7.9	26.5	0.7	16.7
Jones matched on OSCR	17.80	17.4	12.1	14.1	24.2	11.9	18.8	5.3	10.9
Jones matched on DtA	16.05	16.3	12.9	5.9	27.7	4.5	29.4	0.3	20.3
Jones matched on Size	18.76	18.0	21.6	5.3	26.9	8.1	29.1	0.4	27.1

	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	4.5	0.2	76.3	19.7	23.2	66.1	9.1	34.9	20.9
Modified Jones Model	2.5	0.1	78.1	15.9	24.0	73.7	7.1	31.5	20.3
Jones matched on ROA	17.3	15.7	17.7	27.5	12.5	43.3	13.9	16.5	27.3
Prop Score matched Jones	17.1	15.3	14.3	22.2	15.5	47.3	10.8	17.1	27.8
Jones matched on BtM	10.9	2.7	34.6	26.7	11.2	46.9	8.1	18.9	17.9
Jones matched on OSCR	14.7	12.9	21.4	32.0	8.5	46.5	10.9	20.3	20.6
Jones matched on DtA	5.4	1.0	44.8	16.3	10.7	40.3	6.7	20.3	10.2
Jones matched on Size	14.7	1.6	43.5	17.7	14.9	43.6	8.0	21.0	17.4

Panel B: Seed = 2%

	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	53.69	64.3	71.7	20.8	72.7	24.1	87.9	0.5	92.0
Modified Jones Model	52.07	61.5	58.7	23.9	91.1	7.6	89.7	0.1	91.8
Jones matched on ROA	41.89	44.5	44.1	33.9	44.3	40.3	39.7	28.5	33.3
Prop Score matched Jones	39.87	44.0	39.1	33.1	45.5	36.7	37.8	23.1	28.7
Jones matched on BtM	39.28	41.8	51.6	27.5	48.2	25.0	60.0	2.7	46.8
Jones matched on OSCR	40.23	41.7	38.1	33.0	46.4	31.7	50.5	13.3	39.5
Jones matched on DtA	34.60	37.6	36.3	17.9	49.2	14.7	58.7	1.5	52.5
Jones matched on Size	40.25	44.1	46.0	16.9	50.9	23.9	61.0	2.7	64.4

	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	18.8	2.3	96.5	52.0	53.9	90.5	38.9	77.3	48.5
Modified Jones Model	13.5	1.7	97.5	46.1	55.2	93.6	32.1	74.5	46.5
Jones matched on ROA	37.0	32.1	43.2	56.8	32.6	69.2	36.0	45.2	51.4
Prop Score matched Jones	35.7	31.4	40.9	47.9	33.0	73.2	32.5	44.9	50.3
Jones matched on BtM	25.3	9.3	64.7	50.9	29.9	71.1	27.1	48.9	36.9
Jones matched on OSCR	32.1	26.6	52.8	59.9	25.5	69.4	31.9	48.7	42.9
Jones matched on DtA	16.4	4.3	72.4	34.0	26.0	67.0	23.8	48.1	27.7
Jones matched on Size	31.1	7.1	73.7	41.1	36.5	69.3	26.0	51.6	38.0

Table 3.10: Power Test - H_a: Discretionary Accruals Greater Than 0 (4% and 10% seeds)

In this table, we measure the percentage of times in which the null hypothesis of no earnings management is rejected when firms have engaged in revenue increasing activity (hence, we test for type II errors). We conduct simulations on 1500 random samples of 100 firms each without resampling. The power is computed by performing a one-sided t-test for each random sample of mean discretionary accruals and counting the number of times each respective p-value falls below the significance threshold of 5%. In order to simulate the earnings increasing activity, we artificially introduce earnings manipulation by adding seeds of 1%, 2%, 4% and 10% from total accruals before computing discretionary accruals with each model. We also follow Kothari et al. (2005) and assume that earnings management is 50% revenue based, so we add half of the seed to the difference in sales and the difference in accounts receivable in the Jones model before estimating each model.

Panel A: Seed = 4%	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	86.33	98.8	99.8	76.3	98.3	84.6	100.0	17.6	100.0
Modified Jones Model	83.34	98.7	99.4	76.8	99.9	60.1	100.0	11.6	100.0
Jones matched on ROA	87.05	88.3	92.1	78.1	87.3	84.9	91.2	72.1	91.0
Prop Score matched Jones	85.33	87.2	90.0	75.2	86.9	81.9	90.5	62.1	89.6
Jones matched on BtM	80.56	87.7	96.1	70.7	89.0	72.5	96.4	26.3	94.4
Jones matched on OSCR	83.96	87.7	90.3	75.3	88.7	77.8	94.9	49.5	94.7
Jones matched on DtA	75.27	86.5	88.6	59.0	90.1	63.3	96.3	16.2	97.1
Jones matched on Size	78.98	87.4	92.6	59.9	89.7	70.3	96.5	20.7	99.1
	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	75.7	32.6	100.0	97.1	97.2	99.9	95.7	100.0	93.9
Modified Jones Model	65.8	25.1	100.0	95.8	97.0	100.0	93.7	100.0	92.8
Jones matched on ROA	80.6	78.9	91.5	94.2	81.3	96.5	86.3	95.9	89.8
Prop Score matched Jones	80.4	77.3	90.5	91.2	83.6	97.3	83.2	95.0	88.7
Jones matched on BtM	70.9	41.3	97.1	92.3	80.8	96.6	80.5	95.5	81.5
Jones matched on OSCR	76.6	69.5	94.5	94.9	75.2	96.5	81.8	96.7	82.8
Jones matched on DtA	58.6	27.0	98.5	82.6	77.5	94.7	78.9	95.0	69.6
Jones matched on Size	73.9	35.5	98.7	88.1	82.7	95.6	77.7	95.3	79.1

Panel B: Seed = 10%	Mean	All Firms	High BtM	Low BtM	High SALES	Low SALES	High EP	Low EP	High Size
Jones Model	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Modified Jones Model	100.00	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0
Jones matched on ROA	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Prop Score matched Jones	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on BtM	99.98	100.0	100.0	100.0	100.0	100.0	100.0	99.7	100.0
Jones matched on OSCR	100.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on DtA	99.96	100.0	100.0	100.0	100.0	100.0	100.0	99.6	100.0
Jones matched on Size	99.98	100.0	100.0	100.0	100.0	100.0	100.0	99.7	100.0
	Low Size	High O-Score	Low O-Score	High DtA	Low DtA	High AR	Low AR	High PPE	Low PPE
Jones Model	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Modified Jones Model	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on ROA	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Prop Score matched Jones	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on BtM	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on OSCR	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on DtA	100.0	99.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Jones matched on Size	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0

3.6 Conclusions

In this study, we propose a new procedure to compute the discretionary accruals component of total accruals. Our methodology is able to control for the correlation of more than one determinant with accruals, while previous research has only been able to match on one or two variables in the cross section. We match firm industry years based on their propensity to manage earnings, where the propensity score is calibrated from a logistic regression of earnings management proxies on our determinants of earnings management. Then, tests of model specification and power are conducted in order to investigate the statistical properties of our matched discretionary accruals measure compared to conventional models of earnings management and models matched on other determinants.

Our main finding is that propensity score matching mitigates the probability of commiting a type I error when the null hypothesis is no earnings management. Given our criteria for a well defined test, the propensity score matched model produces the least amount of failures when we conduct sorts based on subsets of various determinants. In doing so, our model has shown to be more robust to subsets which exhibit extreme financial levels. In statistical tests of power, propensity score matching does not sacrifice its rejection rate with respect to performance matching in order to gain in the specification test.

Since models has been shown to be either statistically well specified with low power, or perform poorly in tests of specification and exhibit high power, the researcher needs to weigh the relative costs of performing type I versus type II errors before choosing a model to estimate discretionary accruals. We feel that it is more important for researchers, practitioners and regulators to be interested in taking a conservative approach to earnings management detection and opt to mitigate false positives, making propensity score matching the appropriate model to select in earnings management research. This, of course, is at the discretion of the researcher.

Our agenda for further research is to understand the relationship between our measure of discretionary accruals and abnormal equity returns. Given the asset pricing implications involving abnormal accruals pointed out by Xie (2001), it is of great interest to us to revisit this line of research with a measure for earnings manipulation estimated with a small degree of statistical misspecification.

Bibliography

Agarwal, V. and L. Ma (2012): "Managerial Multitasking in the Mutual Fund Industry," *AFA 2011 Denver Meetings Paper.*

Bagnoli, M. and S. G. Watts (2000): "The effect of relative performance evaluation on earnings management: a game-theoretic approach," *Journal of Accounting and Public Policy*, 19, 377 – 397.

Baks, K. P. (2003): "On the Performance of Mutual Fund Managers," *Working Paper - Emory University.*

Barras, L., O. Scaillet, and R. Wermers (2010):

"False Discoveries in Mutual Fund Performance: Measuring Luck in Estimating Alphas," *The Journal of Finance*, 65, 179–216.

Beck, N. and J. N. Katz (1995): "What to do (and not to do) with Time-Series Cross-Section Data," *The American Political Science Review*, 89, pp. 634–647.

Beneish, M. D. (1997): "Detecting GAAP violation: implications for assessing earnings management among firms with extreme financial performance," *Journal of Accounting and Public Policy*, 16, 271 – 309.

Bhattacharya, U., J. Lee, and V. Pool (2012): "Conflicting Family Values in Mutual Fund Families," *The Journal of Finance*, Forthcoming.

Bollen, N. and J. Busse (2004): "Short-Term Persistence in Mutual Fund Performance," *Review of Financial Studies*, 18, 569–597.

Bowen, R. M., S. Rajgopal, and M. Venkatachalam (2008): "Accounting Discretion, Corporate Governance, and Firm Performance*," *Contemporary Accounting Research*, 25, 351–405.

Brennan, M. J. and H. H. Cao (1997): "International Portfolio Investment Flows," *The Journal of Finance*, 52, pp. 1851–1880.

Brown, K. C., W. V. Harlow, and L. T. Starks (1996): "Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry," *The Journal of Finance*, 51, pp. 85–110.

Busse, J. A. (2001): "Another Look at Mutual Fund Tournaments," *Journal of Financial and Quantitative Analysis*, 36, 53–73.

Carhart, M. (1997): "On Persistence in Mutual Fund Performance," *The Journal of Finance*, 52, 57–82.

Chevalier, J. A. and G. D. Ellison (1997): "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105, 1167–1200.

Collins, D. W., R. S. Pungaliya, and A. M. Vijh (2011): "The Effects of Firm Growth and Model Specification Choices on Tests of Earnings Management," *Working Paper*.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997): "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks," *The Journal of Finance*, 52, 1035–1058.

DeAngelo, H., L. DeAngelo, and D. J. Skinner (1994): "Accounting choice in troubled companies," *Journal of Accounting and Economics*, 17, 113 – 143.

DeAngelo, L. E. (1986): "Accounting Numbers as Market Valuation Substitutes: A Study of Management Buyouts of

Public Stockholders," *The Accounting Review*, 61, pp. 400–420.

Dechow, P. M. and R. G. Sloan (1991): "Executive incentives and the horizon problem: An empirical investigation," *Journal of Accounting and Economics*, 14, 51 – 89.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney (1995): "Detecting Earnings Management," *The Accounting Review*, 70, pp. 193–225.

DeFond, M. L. and J. Jiambalvo (1994): "Debt covenant violation and manipulation of accruals," *Journal of Accounting and Economics*, 17, 145 – 176.

Ecker, F., J. Francis, P. Olsson, and K. Schipper (2012): "The Effect of the Amount and Configuration of News on Inferences about Firm-Specific Events," *Working Paper*.

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

FAMA, E. F. AND J. D. MACBETH (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, pp. 607–636.

FENG, L. AND M. S. SEASHOLES (2004): "Correlated Trading and Location," *The Journal of Finance*, 59, 2117–2144.

FERSON, W. AND K. KHANG (2002): "Conditional performance measurement using portfolio weights: evidence for pension funds," *Journal of Financial Economics*, 65, 249 – 282.

FERSON, W. AND R. SCHADT (1996): "Measuring Fund Strategy and Performance in Changing Economic Conditions," *The Journal of Finance*, 51, 425–461.

GASPAR, J.-M., M. MASSA, AND P. MATOS (2006): "Favoritism in Mutual Fund Families? Evidence on Strategic

Cross-Fund Subsidization," *The Journal of Finance*, 61, 73–104.

Golez, B. and J. Marin (2012): "Price Support by Bank-Affiliated Mutual Funds," *Working Paper - University of Texas at Austin.*

Griffin, J. M. and M. L. Lemmon (2002): "BooktoMarket Equity, Distress Risk, and Stock Returns," *The Journal of Finance*, 57, 2317–2336.

Grinblatt, M. and S. Titman (1993): "Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns," *The Journal of Business*, 66, 47–68.

Grossman, S. (1976): "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information," *The Journal of Finance*, 31, pp. 573–585.

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

Gu, Z., C.-W. J. Lee, and J. G. Rosett (2005): "What Determines the Variability of Accounting Accruals?" *Review of Quantitative Finance and Accounting*, 24, 313–334.

Guedj, I. and J. Papastaikoudi (2005): "Can mutual fund families affect the performance of their funds?" *Working Paper - University of Texas at Austin.*

Healy, P. M. (1985): "The effect of bonus schemes on accounting decisions," *Journal of Accounting and Economics*, 7, 85 – 107.

Hellwig, M. (1980): "On the aggregation of information in competitive markets," *Journal of Economic Theory*, 22, 477–498.

Jones, J. J. (1991): "Earnings Management During Import Relief Investigations," *Journal of Accounting Research*, 29, pp. 193–228.

Kacperczyk, M. and A. Seru (2007): "Fund Manager Use of Public Information: New Evidence on Managerial Skills," *The Journal of Finance*, 62, 485–528.

KEMPF, A. AND S. RUENZI (2008): "Tournaments in Mutual-Fund Families," *Review of Financial Studies*, 21, 1013–1036.

KHORANA, A. (1996): "Top management turnover an empirical investigation of mutual fund managers," *Journal of Financial Economics*, 40, 403 – 427.

KOTHARI, S., A. J. LEONE, AND C. E. WASLEY (2005): "Performance matched discretionary accrual measures," *Journal of Accounting and Economics*, 39, 163 – 197.

LI, K. AND N. R. PRABHALA (2005): *Self-Selection Models in Corporate Finance*, Robert H. Smith School Research Paper No. RHS 06-020 l.

LINTNER, J. (1965): "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *The Review of Economics and Statistics*, 47, pp. 13–37.

McNichols, M. F. (2000): "Research design issues in earnings management studies," *Journal of Accounting and Public Policy*, 19, 313 – 345.

Minton, B. A. and C. Schrand (1999): "The impact of cash flow volatility on discretionary investment and the costs of debt and equity financing," *Journal of Financial Economics*, 54, 423 – 460.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley (2003): "How are earnings managed? Examples from auditors," *Accounting Horizons*, 17, 17–35.

Newey, W. K. and K. D. West (1987): "Hypothesis Testing with Efficient Method of Moments Estimation," *International Economic Review*, 28, pp. 777–787.

O'Hara, M. (2003): "Presidential Address: Liquidity and Price Discovery," *The Journal of Finance*, 58, 1335–1354.

Pastor, L. and R. Stambaugh (2002a): "Investing in Equity Mutual Funds," *Journal of Financial Economics*, 63, 351–380.

——— (2002b): "Mutual Fund Performance and Seemingly Unrelated Assets," *Journal of Financial Economics*, 63, 315–349.

Pesaran, M. H. (2004): "General Diagnostic Tests for Cross Section Dependence in Panels," *SSRN eLibrary*.

Pfeifer, U., T. Poersch, and N. Fuhr (1996): "Retrieval effectiveness of proper name search methods," *Information Processing & Management*, 32, 667 – 679.

Rajgopal, S. and M. Venkatachalam (2011): "Financial reporting quality and idiosyncratic return volatility," *Journal of Accounting and Economics*, 51, 1 – 20.

Schmidt, B. and L. Goncalves-Pinto (2012): "Co-Insurance in Mutual Fund Families," *Working Paper - National University of Singapore*.

Sharpe, W. F. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, 19, pp. 425–442.

Sirri, E. R. and P. Tufano (1998): "Costly Search and Mutual Fund Flows," *The Journal of Finance*, 53, 1589–1622.

Skinner, D. and R. Sloan (2002): "Earnings Surprises, Growth Expectations, and Stock Returns or Don't Let an Earnings Torpedo Sink Your Portfolio," *Review of Accounting Studies*, 7, 289–312.

Sotes-Paladino, J. and L. Goncalves-Pinto (2011): "The Value of Cross-Trading to Mutual Fund Families in Illiquid Markets: A Portfolio Choice Approach," *Working Paper - National University of Singapore.*

Starks, L. T. (1987): "Performance Incentive Fees: An Agency Theoretic Approach," *The Journal of Financial and Quantitative Analysis*, 22, pp. 17–32.

TEOH, S. H., I. WELCH, AND T. WONG (1998): "Earnings management and the underperformance of seasoned equity offerings," *Journal of Financial Economics*, 50, 63 – 99.

TUCKER, J. W. (2012): "Selection Bias and Econometric Remedies in Accounting and Finance Research," *Journal of Accounting Literature, Forthcoming.*

WATTS, R. L. AND J. L. ZIMMERMAN (1986): *Positive Accounting Theory*, Prentice-Hall.

WERMERS, R. R. (2003): "Is Money Really 'Smart'? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence," *SSRN eLibrary.*

XIE, H. (2001): "The Mispricing of Abnormal Accruals," *The Accounting Review*, 76, 357 – 373.

YADAV, V. (2010): "Portfolio Matching by Multi-Fund Managers: Effects on Fund Performance and Flows," *Working Paper - ESSEC University.*